Universals and Particulars in Morphology:
Processing and Generalization in Native and Non-Native Speakers of Hebrew

by

Yael Farhy

Cumulative Ph.D. thesis

to obtain the degree of Doctor of Philosophy (Ph.D.)

Submitted to the Faculty of Human Sciences of the University of Potsdam

First supervisor: Prof. Dr. Harald Clahsen (University of Potsdam)

Second supervisor: Dr. João Veríssimo (University of Potsdam)

Potsdam, August 2019

Preface

The present doctoral project was part of the Potsdam Research Institute for Multilingualism (PRIM). It is written in English and is presented as a cumulative Ph.D. thesis at the University of Potsdam, Faculty of Human Sciences.

This thesis is composed of an introduction and four main chapters (3 - 6), followed by a general discussion and a conclusion section. In the first chapter, I introduce the main research topic as well as the main research questions and objectives of this work. The second chapter presents an overview of the publications. The main chapters of this thesis consist of four manuscripts of which three are first-authorship publications (Publications I, II, and III), whereas one is a single-author publication (Publication IV). Publications I, II and III have already been published, while Publication IV has been accepted for publication with an international journal of the field. A synthesis and general discussion of the main findings in the four manuscripts is given in Chapter 7, followed by general conclusions.

Acknowledgements

The last five years have felt like a never-ending roller coaster ride. I am relieved and honored to finally get off the train and excited to move on to new challenges. The PhD journey was challenging and painful at times, but I hope that in the end it made me a stronger person. Looking back, I am amazed by the enormous support I received over the years from family, friends, colleagues and even people I barely knew. All those people left me no choice but to overcome all the moments of frustration and self-doubt, and finally complete this dissertation.

First, I would like to thank Harald Clahsen for his supervision, for his consistent involvement and genuine interest in this project. I am grateful for his detailed feedback and challenging questions in every step of the way. I would also like to thank João Veríssimo for his contagious excitement about research, for being a patient teacher and a supportive mentor, and for opening me the door (unintentionally) to the world of data science. I am also grateful to all the people working at PRIM over the last five years, for their consistent encouragement and support. In addition, I want to thank the Minerva Fellowship for providing me with full financial support for three years and by that allowing me to fully dedicate myself to this research project.

This work could have never been completed without my special group of participants, who deserve a lot of credit: the non-native Hebrew speakers, aka 'my Argentinians' in Israel, were eager to help from day one, not only participating but actively reaching out to more potential participants (special thanks to Sabrina and Julieta, and Ella who even collected some of the data).

From the bottom of my heart I want to say a huge thank you for my Kiron family, especially my first team at Kiron: the very supportive and carrying Olivia (who is in charge of some of the proofreading of this work), Julian, Sofia, Steffen and Renata. They took me in when I had no idea where to go, and little by little helped me to rebuild myself. I was saved by Kiron and for that I will be forever grateful.

I also want to thank friends and acquaintances, all those who entered my life, even for a brief moment, and helped creating a magnificent net of support in the last five years. I would have fallen apart without you. Special thanks to those who have been there from the very beginning to the end of the journey: Sasha, my over-protective angel, who has always made sure I know how much I worth (and proofread a large part of this work) and Sandra, who has been the voice of reason when I could not see the end of the journey. Thank you, the 'Bib

Gang': Ana Stutter, Sasha and Assunta, with guest appearances by Mona and Toivo, for about a year of support and encouragement (full of coffee and chocolate) at the TU library. At those complicated days it felt like we only had each other. I am grateful for Michela and Ana Linnik, who shared insights from their PhD experience and their post-PhD life. I am also grateful for the support from friends abroad: Nastia, Lena, Gal and Julieta, who were always up for existential conversations about the meaning of life.

A special thanks to those with whom I shared most of my meals in the last two years: Sascha, for always making sure I have more than enough home-cooked food and high quality coffee, and Torben, my Eisbär, for teaching me to let go, taking me out of my comfort zone, and using the most ridiculous and childish humor to cheer me up.

Last but not least, I want to thank my family, particularly my parents. This work could not have been accomplished without my parents who very bravely and selflessly accepted me being far away. It has not been easy for them. I am grateful for having parents who have become more and more understanding and open-minded as the years went by. I am grateful for their frequent visits to Berlin as if Berlin was just around the corner, and not a four-hour flight away.

The last words in this section are dedicated to the strongest person I know, my mom. I cannot find the words to express my gratitude and admiration to her, because words cannot do her justice. Her magical energy can move mountains. Her warmth and care are inspiring. She, more than anyone else, is the reason that this thesis was completed. She has been my life mentor, coacher and therapist, especially in the last five years. She has always been there for me, in countless hours on the phone, preaching and guiding, determined to lift me back up after every crisis. I am the luckiest person in the world to have her, and I know that when she is by my side, I can overcome all obstacles life has to offer.

Table of Contents

Summary

For many years, psycholinguistic evidence has been predominantly based on findings from native speakers of Indo-European languages, primarily English, thus providing a rather limited perspective into the human language system. In recent years a growing body of experimental research has been devoted to broadening this picture, testing a wide range of speakers and languages, aiming to understanding the factors that lead to variability in linguistic performance. The present dissertation investigates sources of variability within the morphological domain, examining how and to what extent morphological processes and representations are shaped by specific properties of languages and speakers. Firstly, the present work focuses on a less explored language, Hebrew, to investigate how the unique non-concatenative morphological structure of Hebrew, namely a non-linear combination of consonantal roots and vowel patterns to form lexical entries (*L-M-D* + *CiCeC* = *limed* 'teach'), affects morphological processes and representations in the Hebrew lexicon. Secondly, a less investigated population was tested: late learners of a second language. We directly compare *native (L1) and non-native (L2) speakers*, specifically highly proficient and immersed late learners of Hebrew. Throughout all publications, we have focused on a morphological phenomenon of inflectional classes (called *binyanim*; singular: *binyan*), comparing productive (class *Piel*, e.g., *limed* 'teach') and unproductive (class *Paal*, e.g., *lamad* 'learn') verbal inflectional classes. By using this test case, two psycholinguistic aspects of morphology were examined: (i) how morphological structure affects *online recognition* of complex words, using masked priming (Publications I and II) and cross-modal priming (Publication III) techniques, and (ii) what type of cues are used when extending morpho-phonological patterns to novel complex forms, a process referred to as *morphological generalization*, using an elicited production task (Publication IV).

The findings obtained in the four manuscripts, either published or under review, provide significant insights into the role of productivity in Hebrew morphological processing and generalization in L1 and L2 speakers. Firstly, the present L1 data revealed a close relationship between productivity of Hebrew verbal classes and recognition process, as revealed in both priming techniques. The consonantal root was accessed only in the productive class (Piel) but not the unproductive class (Paal). Another dissociation between the two classes was revealed in the cross-modal priming, yielding a semantic relatedness effect only for Paal but not Piel primes. These findings are taken to reflect that the Hebrew mental representations display a balance between stored undecomposable unstructured stems (Paal) and decomposed structured stems (Piel), in a similar manner to a typical dual-route architecture, showing that the Hebrew mental lexicon is less unique than previously claimed in psycholinguistic research. The results of the generalization study, however, indicate that there are still substantial differences between inflectional classes of Hebrew and other Indo-European classes, particularly in the type of information they rely on in generalization to novel forms. Hebrew binyan generalization relies more on cues of argument structure and less on phonological cues.

Secondly, clear L1/L2 differences were observed in the sensitivity to abstract morphological and morpho-syntactic information during complex word recognition and generalization. While L1 Hebrew speakers were sensitive to the binyan information during recognition, expressed by the contrast in root priming, L2 speakers showed similar root priming effects for both classes, but only when the primes were presented in an infinitive form. A root priming effect was not obtained for primes in a finite form. These patterns are interpreted as evidence for a reduced sensitivity of L2 speakers to morphological information, such as information about inflectional classes, and evidence for processing costs in recognition of forms carrying complex morpho-syntactic information. Reduced reliance on structural information cues was found in production of novel verbal forms, when the L2 group displayed a weaker effect of argument structure for Piel responses, in comparison to the L1 group. Given the L2 results, we suggest that morphological and morphosyntactic information remains challenging for late bilinguals, even at high proficiency levels.

1 General Introduction

The study of experimental psycholinguistics, and cognitive science in general, requires dealing with a large amount of variability. However, for a long time, it was dominantly assumed that variability in performance indicates noise, without yielding substantial insights into the core representational properties of the linguistic system. Thus, it was ignored theoretically and overcome by a precise experimental design which attempts to reduce variability. However, recently, a different approach has emerged, which no longer views variability as a sign of experimental weakness but as an essential and informative aspect of the cognitive system, which is guided by a dynamic and flexible architecture (Amenta & Crepaldi, 2016). Thus, the variability in performance between languages and individuals has become an important aspect of experimental research, as it highlights the flexible features of the cognitive system itself (Andrews & Lo, 2013). In recent years, different sources of variability have been investigated, such as differences across languages, linguistic phenomena, contexts and tasks, as well as differences that stem from individual characteristics of the language users including vocabulary size, spelling abilities, working memory, age, age of language acquisition and hearing abilities (for a review, see Amenta & Crepaldi, 2016).

Despite the attested effects of language and speaker properties on language performance, psycholinguistic research on morphology has given relatively minor attention to variability. Most empirical evidence has been elicited from studies testing relatively homogenous groups of L1 speakers of Indo-European languages, with a strong focus on English. Such a sample is very narrow, since (i) English is considered by many as having a relatively poor morphological system with idiosyncrasies (Blevins, 2006), and (ii) nowadays a large amount of individuals worldwide speak more than one language. In fact, on a global scale, there are more English speakers that are L2 learners rather than L1 speakers (Crystal, 2012). Despite it being taken from a relatively narrow sample, the empirical evidence has provided the basis for general claims about universal properties of the human language system. Less attention has been given so far to the possible influence of specific morphological properties of languages (e.g., Arabic’s non-concatenative morphology; Boudelaa & Marslen-Wilson, 2015) and specific speaker characteristics (e.g., influence of speakers’ vocabulary size on morphological priming effects; Andrews & Lo, 2013) on complex word processing and the way morphology is represented in the human mind. However, if we wish to form a theory of language processing that captures the universal properties of a language and at the same time acknowledges the idiosyncratic characteristics of different languages, collecting evidence from different languages and speakers is essential. This will allow us (i) to attempt to validate theoretical frameworks which rely mostly on detailed but limited evidence from homogenous groups, and ultimately (ii) to develop and adapt these frameworks in order for them to account for certain language-specific and speaker-specific patterns, and at the same time address universal properties of morphology. Ideally, such theoretical models will be able to explain how the specific properties of an individual language and speaker shape the type of solutions that are adopted to optimize the mapping between form and meaning (Frost, Grainger, & Carreiras, 2008).

Therefore, the present dissertation investigates effects of variability within the morphological domain. How and to what extent are morphological processes and representations shaped by specific properties of languages and speakers? To address this question, the present work focuses on a less explored language, Hebrew, and a less investigated population: late learners of a second language. We have tested Hebrew, a language with relatively unique morphological properties. Hebrew is a Semitic language which has a very salient and pervasive non-concatenative morphological structure, namely a non-linear combination of consonantal roots and vowel patterns to form lexical entries (*L-M-D* + *CiCeC* = *limed* 'teach'). Throughout the experimental work, we have examined how this specific structure of Hebrew shapes morphological processes and representations in the Hebrew lexicon, while comparing our findings to main findings from Indo-European literature. Furthermore, we have examined how these morphological processes are affected by speaker characteristics, focusing on the comparison between native (L1) and non-native (L2) speakers, specifically those who acquired Hebrew at a later age (after the age of 7 years), but are still highly proficient and immersed. It is still highly debated whether and how acquiring a second language at a later age influences processing and the organization of the mental lexicon (see Chapter 1.2.2).

Throughout each of the publications, we have focused on a comparison between productive and unproductive morphological patterns, due to the central role productivity plays in current theoretical views and empirical findings (detailed review in Chapter 1.1). We examined two psycholinguistic aspects of morphology where productivity has been shown to play a central role: (i) how morphological structure affects online recognition of complex words (e.g., how the recognition of *walked* is affected by its structure, [*walk* + *-ed*], i.e., the base *walk* and the past tense suffix *-ed*) and (ii) what type of cues are used when extending morpho-phonological patterns to novel complex forms, a process referred to as morphological generalization (e.g., what form do speakers produce for the English past tense of the novel word *spling*? *Splang* or *splinged*?).

The topic of the dissertation will be addressed as follows. The introduction reviews relevant concepts, theories and psycholinguistic evidence. First, the role of productivity in morphology and how it corresponds to processing and generalization (Chapter 1.1) is discussed. Later, we present the debate between psycholinguistic accounts of Hebrew morphology and describe relevant concepts in Hebrew morphology (Section 1.2.1). The next part (Section 1.2.2) introduces theories and psycholinguistic evidence of L2 speakers, in order to understand unique properties of L2 morphology that were proposed in previous literature. In Chapter 1.3, the phenomenon of the Hebrew verbal inflectional classes is introduced, with an emphasis on the most common classes, *Paal* and *Piel*, which contrast on productivity level and are the center of investigation throughout the four publications. Finally, in Chapter 1.4, the aims and objectives of the dissertation will be laid out. After the introduction, an overview of the four publications is included, followed by the full publications. All publications focus on a comparison of a productive (Piel) and an unproductive (Paal) verbal class. The first three publications examine word-level real-time processing of Hebrew verbal classes and the fourth focuses on generalization properties of those classes. Publications II and IV also include an investigation of L2 speakers. Lastly, a general discussion of the main findings of the four publications is provided, attempting to explain the Hebrew findings with current theoretical frameworks, and proposing a linguistic account that explains the full pattern of results. In addition, a discussion of the L2 pattern in Publications II and IV is presented to shed light on how the present results contribute to current L2 theoretical accounts.

1.1 The Role of Productivity in Morphological Processing and Generalization

Morphological productivity has been defined as "the possibility for language users to coin, unintentionally, a number of formations which are in principle uncountable" (Schultink, 1961). More simply, the degree of productivity of morpho-phonological patterns can be referred to as the extent to which the morpho-phonological pattern is lexically unrestricted and can be extended to new forms. The exact way to quantitatively measure morphological productivity has been debated over the years. Some emphasized the involvement of type frequency in the measurement process (Aronoff, 1976), token frequency (Rumelhart & McClelland, 1986), and others stressed the central role of hapax legomena (words that appear once in the language or a corpus) along with token frequency as a better indication of productivity (Baayen, 1992). The present dissertation adopts the perspective of Baayen (1992), sharing the notion that the formation of new word forms, as reflected in hapax legomena, is the most crucial aspect of morphological productivity, rather than type frequency. Within this framework, the English suffix *-ness*, for example, is considered more productive than the suffix *-ity*, despite a larger type frequency of *-ity*.

Over the years, a close relationship has been found between properties of generalization and processing of complex forms, and productivity of morpho-phonological patterns. First, the present section will present an overview of such empirical evidence from a wide variety of techniques. Then, an overview of how different theoretical accounts attempted to interpret the findings will follow.

The degree of productivity of morpho-phonological pattern has been significantly related to specific sources of generalization; patterns with low productivity tend to be generalized by phonological similarity of the novel form to other existing forms, while highly productive patterns are not affected by phonological similarity when extended to novel forms (e.g., Bybee & Moder, 1983). A large body of evidence for productivity effects on generalization has emerged from tasks of elicited production of novel words and acceptability judgment of novel words (English: Prasada & Pinker, 1993; Greek: Stavrakaki & Clahsen, 2009; Italian: Say & Clahsen, 2002; German: Clahsen, 1997; Hahne, Mueller, & Clahsen, 2006; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995; Hebrew: Berent, Pinker, & Shimron, 1999; Japanese: Hagiwara et al., 1999). For instance, Prasada and Pinker (1993) tested the generalization properties of the English past tense and found that when presented with the novel word *spling*, which is highly similar to the existing words *sing* and *ring* (which go through an unproductive operation of vowel change in formation of their past form), L1 speakers had higher probability to produce an irregular past tense form like *splang*, compared to novel words with low similarity to existing irregular words. On the other hand, the regular past tense (the productive *-ed* suffix) was not sensitive to factors of phonological similarity and was produced in very high rates compared to the irregular past tense. Similar evidence was also found in Semitic languages. Berent et al. (1999) investigated the generalization of the plural inflection suffixes. In an acceptability judgement task of novel words, it was found that the acceptance of an irregular inflection (suffix *-ot* to masculine-sounding nouns) was affected by the degree of phonological similarity between the novel word and other existing nouns which take the irregular inflection. The acceptance for regular inflection (suffix *-im*) was much higher:

accepted in over 90% of the cases, not influenced by phonological similarity, unlike the irregular forms.

However, other studies, mainly those that manipulated phonological similarity using model simulations, have questioned the significant influence of productivity on the nature of generalization (e.g., Hahn & Nakisa, 2000). They reported that productive patterns are also generalized based on phonological similarity, when implementing either exemplar analogy-based models (*analogical model of language* [AML]; Eddington, 2000, 2002; *generalized context model*: Nosofsky, 1986), models with rules sensitive to phonological properties, such as the *minimal generalization learner* (MGL; Albright, 2002; Albright & Hayes, 2003), or *connectionist networks* (Rumelhart & McClelland 1986; Colombo, Stoianov, Pasini, & Zorzi, 2006; Hare, Elman, & Daugherty, 1995).

Further support for a dissociation between productive and unproductive morphological patterns arises from studies of complex word recognition using various experimental techniques. Productive and unproductive patterns showed contrasting priming effects, with productive ones yielding stronger priming effects (Morris & Stockoll, 2012; Neubauer & Clahsen, 2009; Rastle, Lavric, Elchlepp, & Crepaldi, 2015). Base frequency effects were stronger for productive patterns than for unproductive ones in lexical decision tasks (Arabic: Wray, 2016; English: Ford, Davis, & Marslen-Wilson, 2010; Spanish: López-Villaseñor, 2012; Dutch: Bertram, Schreuder, & Baayen, 2000; German: Clahsen, Eisenbeiss, & Sonnenstuhl, 1997). Response Times (RTs) to complex words with productive patterns were longer than RTs to monomorphemic words, but this difference did not occur for unproductive patterns (Finnish: Bertram, Laine, & Karvinen, 1999). In lexical decision tasks, pseudowords with productive affixes were more difficult to reject than pseudowords with unproductive affixes (e.g., in Italian: Laudanna, Burani, & Cermele,1994).

An asymmetry between recognition of words with productive and unproductive affixes was found also in online measures, such as ERP (German: Hahne et al., 2006; Weyerts, Münte, Smid, & Heinze, 1996; English: Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Newmann Ullman, Pancheva, Waligura, & Neville, 2007; Rastle et al., 2015; Spanish: Rodriguez-Fornell, Münte, & Clahsen, 2002, but see, for example, Justus et al., 2011 for counter evidence) and fMRI studies (e.g., Carota, Bozic, & Marslen-Wilson, 2016; Vannest, Polk, & Lewis, 2005. For a comprehensive review of online measures in morphology, see Leminen, Smolka, Dunabeitia, & Pliatsikas, 2019). For example, Newmann et al. (2007) examined the English past tense in a violation detection task and found an elicited left-lateralized anterior negativity (LAN) for violations of the productive regular past tense but not for the unproductive irregular past tense, a component that has been suggested to reflect neurocognitive activity that underlies compositional processes in morphology and syntax. The authors suggested that this shows that distinct neurocognitive tracks are involved in the recognition of the regular and irregular English past tense. In a recent fMRI study of derived forms in Italian, Carota et al. (2016) reported that productive (and semantically transparent) forms, but not unproductive ones, showed selective left perisylvian activity, a network that has been associated with combinatorial and decompositional processes in previous studies.

Out of the different measures noted above, the present thesis focuses on the priming paradigm, that is, the way existing complex forms prime their base stem during word recognition, which allows to tap into real-time word-level processing, examining *morphological priming* during word recognition. Morphological priming is defined as the amount of facilitation in target recognition after the presentation of a prime that is morphologically related to it, i.e., sharing a morphological constituent with the target word (e.g., prime: *walked* – target: *walk*). This facilitation is measured in relation to the target recognition after a prime which is semantically, phonologically and morphologically unrelated (e.g., *decide*). The priming technique is considered to tap into real-time processing of words; the presentation of the prime activates linguistic properties that are shared with the target's properties, including morphological ones. That means that the shared morphological constituent (*walk*) is already activated when the target is presented, leading to faster recognition, defined as a morphological priming effect. Under a decompositional approach, the effects are interpreted as evidence for decomposition of the constituents of the prime (e.g., *walk* + *ed*), which allows for an early activation of the shared constituent during recognition of the target (*walk*). Alternatively, morphological priming effects have been interpreted as a function of a graded overlap between form and meaning of the prime and the target (e.g., Gonnerman, Seidenberg, & Andersen, 2007; Morris, Frank, Grainger, & Holcomb, 2007).

Priming is not a uniform paradigm, but it encompasses several types of priming techniques, arguably tapping into distinct levels of word recognition processes. The present dissertation focuses on two types: masked priming and cross-modal priming. In masked priming, the primes are visually presented for 33–50 ms immediately before the targets. This technique is intended to tap into an initial modality-specific step of accessing a word. In this access phase, words are decomposed into constituents, blind to semantic compositionality (Marslen-Wilson, 2007), which is consistent with findings showing a priming effect for pairs like *department–depart*, morphologically-related pairs which do not share semantic properties, and no priming for pairs that share a semantic relationship but have no morphological relationship (*doctor–nurse*; e.g., Rastle Davis, & New, 2004). In cross-modal priming, primes and targets are presented in different modalities; usually, primes are auditory, and targets are visual. This technique is thought to tap into a later stage of recognition processes: the central representation of the lexical entries, namely the morpho-semantic representation. Therefore, in cross-modal priming and overt priming, pairs like *department–depart* tend to not display a priming effect, since their representations are not morphologically connected at the central level (English: Feldman & Soltano, 1999; Marslen-Wilson et al., 1994; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rueckl & Aicher, 2008; Polish: Reid & Marslen-Wilson, 2003; French: Longtin, Segui, & Hallé, 2003).

Dissociation of morphological priming effects between productive and unproductive patterns was found in various languages and phenomena (English past tense: Marslen-Wilson, Hare, & Older, 1993; Stanners, Neiser, Hernon, & Hall, 1979; German participles and plurals: Neubauer & Clahsen, 2009; Sonnenstuhl et al., 1999; Portuguese inflectional classes: Veríssimo & Clahsen, 2009). Specifically, the findings show that productive patterns display significantly larger priming effects than unproductive ones, which sometimes do not display a priming effect at all. For demonstration, in a cross-modal priming study, Sonnenstuhl et al. (1999) examined priming effects of German participles and found that when an auditory prime included the productive affix *-t* (*gekauft* 'bought' primes *kaufen* 'buy') a full priming effect was yielded; that is, the morphological priming effect was not different from an identical priming effect (*kaufen–kaufen*). In contrast, when an auditory prime included the unproductive affix *-en* (*geschlafen* 'slept' primes *schlafen* 'sleep'), only a partial priming effect was yielded; that is, morphological priming was found but it was significantly smaller than an identical priming effect (*kaufen–kaufen*). This is despite the fact that in both cases the form is decomposable, and the verbal stem is fully repeated in the prime and target.

How do theoretical accounts explain the contrasts in generalization properties and priming effects between productive and unproductive morphological patterns? It has been debated whether these empirical dissociations reflect two distinct routes and/or representation types for productive and unproductive patterns, or whether one route is employed for all morphological forms, flexible enough to account for different effects for patterns with differing degrees of productivity.

The dual-route approach interprets the asymmetry in findings between productive and unproductive patterns as evidence for the different types of mental representations. This view postulates that the linguistic system consists of a lexicon containing lexical entries and a computational system of rule-based combinatorial operations. A central concept of the theory is the distinction between two types of representations of complex forms: *structured forms* and *unstructured forms* (Clahsen, 1999, 2006; Pinker, 1999). The two types of representations are considered to correspond to *morphological productivity;* complex forms of productive patterns have structured representations and those of unproductive patterns have unstructured representations. According to the dual-route approach, structured representations are generated by symbolic morphological rule operations and can therefore be extended easily to novel forms. Consequentially, they are often highly productive. Unstructured representations, on the other hand, are not generated by rules, stored as wholes and are connected to their base form via associative links. They are extended to novel forms via similarity-based associations, mostly phonological similarity, and not by symbolic rules, as found in generalization studies reported above. They are therefore less easily extended to novel forms and have lower productivity. Recognition processes are also different between structured and unstructured representations. During recognition, forms of structured representations are decomposed to their morphological constituents before accessing the lexicon. Forms of unstructured representations are not decomposed during word recognition but accessed directly as whole forms. The dual-route approach interprets the contrast reported in priming effects between productive and unproductive morphological patterns as differences rooted in access process, as productive forms are decomposed and unproductive ones are not.

On the other hand, other researchers do not consider the asymmetry in generalization properties (mainly phonological similarity) or priming effects as reflecting two routes or representation types, but a single route which is able to account for the different productivity-based effects. A very prominent single-route approach is the one based on a distributed approach, including accounts such as the *network model* (Bybee, 1995) and connectionist network models (Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002; Plunkett & Marchman,1993; Plunkett & Nakisa, 1997; Rumelhart & McClelland, 1986). This view does not postulate a specific level of morphology in the mental lexicon but argues that morphology reflects a learned sensitivity to the systematic relationship between the word form and its meaning. This approach argues that generalization is based on a single mechanism involving associative networks rather than symbolic rules and found evidence for phonological-based generalization even in generalization of productive regular morpho-phonological patterns, as mentioned earlier (e.g., Albright, 2002), contradicting claims of the dual-route approach. The distributed approach does not interpret contrasts found in priming effects between productive and unproductive patterns as reflecting two different processes. Since, according to this approach, recognition is based on mapping between form and meaning, the different priming effects are argued to reflect differences in the overlap of form and meaning between prime and target. Prime-target pairs like *taught–teach* show less priming than *walked–walk*, mainly due to less form overlap between prime and target (Gonnerman, et al., 2007).

1.2 Sources of Variability

The current work investigated variability in processing and generalization of productive and unproductive morphological patterns. Here I introduce the two sources of variability that were examined in the dissertation. The first section will introduce the particular properties of Hebrew morphology and how they have been shown or suggested to affect the organization of the mental lexicon. Then, the following section will describe how the specific properties of L2 speakers might affect the way they process and represent complex words.

1.2.1 Hebrew morphology

The most dominant approach to Semitic morphology (e.g., McCarthy, 1981; Berman, 1997) entails a very robust non-concatenative morphology which applies to most forms and consists of consonantal roots and vowel patterns (alongside concatenative morphology). The notion of the consonantal root in Hebrew morphology dates back to the 11th–12th centuries, in the work of the first grammarians of Hebrew (Aronoff, 2013). According to this approach, most lexical entries (also referred as lexemes) are formed by a non-linear combination of constituents: a consonantal root and a vowel pattern; both are bound and cannot be pronounced when are separated from each other. The root is a sequence of mostly three consonants which are inserted into designated positions in the vowel pattern (which sometimes also contains consonants). For example, the word *talmid* ‘a pupil’ is formed by the root *L-M-D* and the pattern *taCCiC* (the letter ‘C’ represents the designated positions of the root consonants). Roots are considered to contribute a core meaning to the word; for instance, the root *L-M-D* is often linked to a general notion of learning. The vowel patterns assign the syntactic category of a lexical entry (*taCCiC* denotes that the word is a noun), and is usually not fully represented orthographically (e.g., the orthographic form of *talmid* is *TLMID תלמיד)*. This approach is referred to as the *root-based approach* since it postulates that the consonantal root plays a very central role in the organization of the Hebrew mental lexicon. This approach emphasizes the uniqueness of the Semitic morphology by positing a special type of morphological constituent that is very central in Semitic languages but does not exist in non-Semitic languages. This special constituent, the consonantal root, is relatively unique as it is unpronounceable by itself and combined non-concatenatively. Thus, this approach views the Hebrew morphological system as more abstract and less dependent on properties of phonology and orthography in comparison to Indo-European languages.

Psycholinguistic research in the last 20 years has closely investigated whether the unique root-and-pattern structure as described in root-based linguistic theories is reflected in the mental processes of Hebrew speakers as well as Arabic (a Semitic language with similar root-and-pattern structure) during complex word recognition and production. Overall, robust empirical evidence was found for the existence of the consonantal root as a mental constituent using various techniques, such as masked priming (Hebrew: Deutsch, Frost, & Forster, 1998; Feldman & Bentin, 1994; Frost, Forster, & Deutsch, 1997; Arabic: Boudelaa & Marslen-Wilson, 2005), cross-modal priming (Hebrew: Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000; Arabic: Boudelaa & Marslen-Wilson, 2011, 2015), picture-word interference paradigm (Deutsch, 2016; Deutsch & Meir, 2011; Kolan, Leikin, & Zwitserlood, 2011), the segment-switching task (Feldman, Frost, & Pnini, 1995), semantic judgement task (Prior & Markus, 2014), examination of pseudowords in a lexical decision task (Yablonski & Ben-Shachar, 2016), elicited production and acceptability judgement of novel words (Berent, Everett, & Shimron, 2001; Berent & Shimron, 1997), eye tracking (Deutsch, Frost, Pelleg, Pollatsek, & Rayner, 2003; Deutsch, Frost, Pollatsek, & Rayner, 2005) and online measures such as MEG (Gwilliams & Marantz, 2015; Kastner, Pylkkänen, & Marantz, 2018), fMRI (lexical related judgement: Bick, Goelman, & Frost, 2008; masked priming: Bick, Frost, & Goelman, 2010) and EEG (Boudelaa, Pulvermüller, Hauk, Shtyrov, & Marslen-Wilson, 2010). Considerably less empirical attention was given to the processing of vowel patterns in complex words. The evidence for vowel patterns as separate mental constituents was reported to be less robust for Hebrew (but not Arabic) and was found only for verbal patterns (Deutsch et al., 1998), although recent evidence indicates that perhaps nominal pattern extraction is also part of lexical access (Deutsch & Malinovitch, 2016; Deutsch, Velan, & Michaly, 2018).

The mental representation of consonantal roots was investigated in a series of studies of morphological generalization by Berent and colleagues (e.g., Berent et al., 2001; Berent & Shimron, 1997). Their goal was to investigate whether the root is indeed a mental symbolic variable in the speaker's mind. For this purpose, they tested sensitivity to a phonological constraint that does not allow an initial-root gemination; that is, two identical consonants cannot appear in the first and second positions of the Hebrew root (with a few exceptions such as *M-M-N*, *M-M-SH*). Hebrew, however, allows final-root gemination (e.g., *M-S-S*). Berent and colleagues assumed that if speakers apply this constraint to novel roots, it would support the existence of roots as mental symbolic variables. In acceptability judgement tasks, speakers have rated novel words with novel initial-root gemination as very unnatural compared to novel words with novel roots of three different consonants or with final-root gemination (Berent et al., 2001; Berent & Shimron, 1997). In elicited production tasks, they were presented with a biconsonantal novel root (e.g., *SM*) and an existing target vowel pattern (e.g., *CaCaCti*), and were required to produce a novel word with the root and pattern presented to them. The results showed that novel words were produced with final-root gemination (doubling the consonant in the final position of the root, such as *samamti*) in 46% of the responses, while only less than 0.5% of the responses included novel words with initial-root gemination (*sasamti*; Berent, 2002; Berent et al., 2001). Overall, the findings from both acceptability judgement and elicited production tasks indicated that Hebrew speakers were sensitive to the constraint on the position of the gemination. According to the authors, such sensitivity can be best explained by presupposing the existence of separate mental representations for consonantal roots to which the constraint is applied.

The Hebrew root-and-pattern structure has been examined thoroughly in lexical-access studies, in which one of the main techniques has been the priming paradigm, especially masked priming. Hebrew priming studies consistently detected morphological root effects; that is, prime words like *TaLMID*[1] ‘pupil’ led to a faster recognition of words like *LeMIDaH* ‘learning’ compared to an unrelated prime like *BiTaXON* ‘security’ (masked priming: Feldman & Bentin, 1994; Frost et al., 1997; Deutsch et al., 1998; Frost, Kugler, Deutsch, & Forster, 2005; Velan, Frost, Deutsch, & Plaut, 2005; cross-modal priming: Frost et al., 2000). At the same time, semantic priming was not detected (*nurse* did not prime *doctor*: Frost et al., 1997), and orthographic priming was also not found (i.e., an overlap between prime and target that is only orthographic but not morphological or semantic; Frost et al., 2005). Taken together, the findings were argued to reflect a morphological effect that cannot be fully explained by an overlap of semantics and form. Instead, the common interpretation to these findings is that the root (*L-M-D)* was extracted and activated from the prime, allowing a faster recognition of the target, as the root of the target has already been activated in the prime. These findings show that a morphological constituent can be extracted even when it is not displayed as a contiguous phonological unit but is disrupted by another morpheme intervening between the root consonants.

[1] For a more accurate representation of the Hebrew script, capital letters mark consonants and vowels which have an orthographic representation, while small case letters represent vowels that are not orthographically represented in the script.

A similar conclusion came from a series of studies using the picture-word interference paradigm (Deutsch, 2016; Deutsch & Meir, 2011). In Deutsch and Meir (2011), L1 Hebrew speakers were asked to produce the name of the object they saw in a picture. Around the same time, they heard a word related to the picture: either a root-related word, a semantically related word or a phonologically related word (or an unrelated word). The stimulus onset asynchrony (SOA) was manipulated so the distractor word was presented shortly before the picture (–200 ms, –100 ms), at the same time as the picture or shortly after the picture (+100 ms, +300 ms). The results showed distinct patterns to the three related conditions. A root effect was significant and stable throughout the different SOAs; that is, the root-related word facilitated the production of the object in the picture to the same extent in all SOAs. In contrast, a semantic effect was mostly inhibitory; that is, it interfered with the word production. A phonological effect was facilitatory, but unlike the root condition, it was influenced by the SOA, not significant in –200 ms and strongly increasing at +100 ms. The authors suggested the results reflect an autonomous morphological root-based process that is distinct from semantic and phonological processes.

Further support for the claim that Semitic morphology is more abstract compared to other languages and goes beyond interaction of form and meaning properties has been found in priming studies. Empirical contrasts were found between Semitic and non-Semitic languages in the way semantic and orthographic properties play a role in recognition of complex words. Regarding semantics, it was reported that prime-target pairs that shared a root but had an opaque semantic relation between them, like *MiTXaSheV* ‘considerate’–*XIShUV* ‘calculation’, showed a root-priming effect, not only in masked priming (Boudelaa & Marslen-Wilson, 2005; Deutsch et al., 1998; Frost et al., 1997), like in many previous findings of Indo-European languages, but even in cross-modal priming (Boudelaa & Marslen-Wilson, 2015; Frost et al., 2000), a very uncommon finding in the general psycholinguistic literature (but see Smolka, Komlosi, & Rösler, 2009; Smolka, Preller, & Eulitz, 2014, for similar findings in German). As described earlier, cross-modal priming arguably taps into a central level of the mental lexicon, where only entries with a morpho-semantic relation to their base are represented in a structured manner (such as *sadness*). Complex words that are not related on a morpho-semantic level to their base (like *department*) are assumed to not have structured representations and thus are not expected to prime their base in cross-modal priming (Marslen-Wilson, 2007). The finding that such priming does occur in Hebrew and Arabic posits an obstacle to the assumption that lexical access and morphological processing are universal. Similar findings in Hebrew were reported also in other paradigms (Deutsch, 2016; Prior & Markus, 2014). For example, in a picture-word interference paradigm, it was found that root-related distractor words that had an opaque semantic relation to the word represented in the picture yielded similar facilitation effects as distractors with transparent relation to it (Deutsch, 2016). In line with these types of findings, the overwhelming majority of models of lexical access (or retrieval) in Hebrew have not presupposed two separate levels (a morpho-orthographic level and a morpho-semantic level), but one morphological level, strongly evolving around roots (Deutsch et al., 1998; Deutsch & Meir, 2011; but see Kolan et al., 2011).

Regarding orthographic properties, contrast between Hebrew and Indo-European languages was found in the *transposed letter effect* (TLE), the facilitation in target recognition when the visual prime differs orthographically from the visual target by transposed letter (*pencil–pencil*). However, this effect was not found in root-based words in Hebrew but only in Hebrew words without a root-based structure (Velan & Frost, 2009, 2011). In line with this finding, in a masked priming study, English-Hebrew (balanced) bilinguals showed an orthographic priming effect in English (*freeze–free*) but not in Hebrew (*SIDUR 'an arrangement'–SIPUR 'a story'*; Frost et al., 2005). Advocates of the root-based view interpreted the results as support that the Semitic mental lexicon is organized differently than the non-Semitic lexicon. Unlike Indo-European languages, the Hebrew mental lexicon is not organized based on orthographic principles. Instead, the root-based lexical entries are primarily organized around consonantal roots. To summarize, the empirical evidence suggests that morphological structure in the Hebrew mental lexicon is more independent of semantic and orthographic principles compared to Indo-European languages.

However, not all linguists agree with the claim of the root-based approach that Hebrew and other Semitic languages have a unique constituent like the consonantal root, which does not exist in non-Semitic languages. Several accounts have challenged the root-based approach (Aronoff, 1994; Bat-El, 1994, 2002; Gafos, 1998, 2003; Ussishkin, 1999, 2005; for Arabic, see Benmamoun, 1999, 2003). They share the assumption that a unique consonantal root constituent is not required to explain Semitic morphology, and that the stem (namely, root and pattern) is the minimal morphological unit for word formation processes; thus, this approach is referred to as a *stem-based approach*. Semitic morpho-phonological patterns do seem to be unusual at first glance: non-adjacent vowels are modified as part of word formation processes and morphologically related words share a sequence of consonants which have co-occurrence restrictions. Root-based accounts regard these Semitic morpho-phonological properties as Semitic-specific and as supporting the postulation of a root unit. The stem-based approach argues that in fact none of these patterns are unique to Semitic morphology (for a review, see

Bat-El, 2002). Morphological processes that involve vowel modification occur even in a relatively morphologically poor language like English (irregular past tense like *drink–drank*). It may be that the consonants in Semitic morphology are especially salient, because the process of vowel modification as part of word formation is very common (Seidenberg & Gonnerman, 2000). Thus, rather than qualitative differences in the structure of representations that are available in the Semitic mental lexicon, it might be that Semitic morphology is quantitatively different from other languages, reflected in the statistical properties of the lexicon and application of vowel modification (Berent, Vaknin, & Marcus, 2007). Storage of stems as the minimal unit, rather than consonantal roots, still allows speakers to track co-occurrences of the stem consonants. To summarize, according to the stem-based approach, Semitic morpho-phonological regularities and patterns can be captured by a universal set of constraints and representation types.

The stem-based approach argues that the empirical evidence described above that supports the unique root-based processing of Hebrew can in fact be accommodated also by postulating the stem as the minimal unit encoded in Hebrew (Berent et al., 2007). For example, root-priming effects can also be interpreted as arising from a morphological relationship between two stems (like *grow–grew*) and not an activation of the same root in the prime and target. Recall that a series of generalization studies was conducted by Berent and colleagues, investigating the gemination constraint in Hebrew. In the most recent study of the series, Berent et al. (2007) have questioned their previous conclusion that the root is a mental symbolic representation in the Hebrew lexicon. In an acceptability judgement task, they examined the initial-root gemination constraint again by presenting novel words with initial-root gemination, final-root gemination and roots without gemination, but this time they manipulated the vowel pattern as well. Roots were presented inside a vowel pattern that was found to often include gemination (*CiCuC*), and inside a vowel pattern in which geminations are rare (*CeCeC*). They found that the strength of the constraint depended on the vowel pattern the root was embedded into; speakers accepted initial-root geminations more when the vowel pattern was *CiCuC* than *CeCeC*, although these forms were still less acceptable than forms with final-root gemination and without gemination. Given these results, the authors have raised the possibility that the constraint is applied over stems and not roots, in line with universal views arguing that lexical representations encode stems, and that a unique root representation is not necessary to describe the Hebrew mental processes.

Furthermore, the conclusion regarding the reduced dependency of the Hebrew lexicon on semantic and orthographic principles can be questioned. For instance, in the cross-modal study by Frost et al. (2000), there was still contrast between semantically opaque and transparent prime-target pairs; even though both pair types showed a significant priming effect, transparent pairs displayed larger priming effect than opaque ones. Plaut and Gonnerman (2000) even managed to mirror these results in a connectionist network model, which is based on form and meaning mapping. Thus, it appears that semantic information is accessed during processing of the prime, at least at a later, more central-level processing stage. Another example is from the TLE results in Hebrew. An interesting pattern was found in Velan and Frost (2011); unlike words with an internal structure, TLE was significant when Hebrew words did not have the typical root-and-pattern structure but were nouns without an internal structure like *AGaRTaL* ‘a vase’. In fact, even words with uncommon roots, like *TaRMIL* ‘a backpack’ (the root *R-M-L* does not appear in other words) showed a TLE in one of the two experiments. Recognition of such words was also facilitated by orthographically related pseudowords, unlike words with roots of high type frequency.

Taken together, the findings from both generalization and word recognition studies generally support a root-based view; that is, the specific Hebrew properties lead to different representational structure which is centered around consonantal roots, a unique morphological constituent that is present only in Semitic languages. Despite its important role in investigating universality in language processes, relatively small empirical attention was given to the stem-based view proposed by several theoretical linguists, suggesting that Hebrew morpho-phonological regularities can be captured by universal representations based on encoding stems, and that Hebrew morphology is not as unique as the prominent root-based view claims.

1.2.2 L2 morphology

L2 speakers, even at high proficiency levels, often show difficulties in mastering different linguistic domains of their L2. Previous psycholinguistic literature examining L2 speakers has revealed nonnative-like performance in the production and comprehension of complex morphological forms, such as morphological errors in production (e.g., **drinked*) and reduced sensitivity to morpho-syntactic errors in reading (e.g., Coughlin & Tremblay, 2013; Jiang, Hu, Chrabaszcz, & Ye, 2017; McCarthy, 2008; McDonald & Roussel, 2010; Montrul, 2011; White, Valenzuela, Kozlowska-Macgregor, & Leung, 2004). Numerous accounts have attempted to explain such contrast between L1 and L2 performance. Generally, these accounts can be divided into two views. One view emphasizes *cognitive-general* properties as the underlying source for nonnative-like performance in L2, rather than linguistic-specific properties. This approach emphasizes that language usage is strongly connected to cognitive abilities as it relies on the abilities to encode, store and retrieve information from memory, and proposes that a precise understanding of the memory processes is essential for understanding L1/L2 differences in processing. Some advocates of this approach have proposed that such differences are driven by L2 limitations in working memory capacity, namely limitations in the amount of information that an individual can hold active at a certain timepoint (Hopp, 2010; McDonald, 2006). Empirical evidence was found for this proposal in McDonald (2006), showing that general cognitive factors like memory capacity, decoding ability and processing speed are correlated with poor performance of L2 in grammaticality judgement tasks. The author also examined L1 speakers performing under different stress conditions that affect memory (remembering 7-digit strings while performing the task), decoding (hearing the sentences with an overlay of white noise) and speed (required to answer 500 ms after the end of the sentence). Under these stress conditions, the performance of L1 speakers paralleled the performance of L2 speakers in a grammaticality judgement task without stress conditions. The findings of this study were taken to argue that general cognitive sources can explain poor L2 performance, and that L2 speakers process the language under difficult cognitive conditions, which sometimes hinder them from accessing and applying their grammatical knowledge.

In a recent account, Cunnings (2017) explains the L1/L2 difference in sentence processing using memory-based terms as well, but unlike McDonland (2006), he assumes a model which does not postulate a separate working memory component (see discussion in McElree, 2006). Consequentially, he defines the source of L1/L2 differences not in terms of memory *capacity*, but by emphasizing the processes involved in memory encoding, storage and retrieval operations. He argues that what determines L1/L2 sentence processing differences is that L2 speakers have a greater susceptibility to interference in memory retrieval during online parsing. For example, Cunnings suggests that previous results (e.g., Keating, 2010) showing L2 reduced sensitivity to agreement violations in long dependencies ('the owner of the successful companies *have decided to quit') can be interpreted as resulting from a larger L2 interference in memory retrieval.

On the other hand, other accounts point to linguistic-specific sources as the underlying source of nonnative-like L2 morphological comprehension and production, arguing that cognitive-general sources are not sufficient for explaining the differences between L1 and L2 processing. A prominent example for this approach is the *shallow structure hypothesis* (SSH; Clahsen & Felser, 2006a, 2006b, 2018), which explains the L1/L2 differences as stemming from two possible sources: nonnative-like grammatical knowledge and/or a nonnative-like processing system. According to the hypothesis, L2 speakers are less efficient with computation and manipulation of abstract grammatical representations in real-time processing. They compensate by relying more on semantic, pragmatic and other types of non-grammatical surface-level information (Clahsen & Felser, 2006a, 2006b, 2018). Although the SSH has originally focused on parsing of syntactic structures, it extended its predictions to the morphological domain, arguing that the L2 system relies on whole-form lexical storage when the L1 system relies on combinatorial processes and decomposition of complex forms.

The most prominent support for this claim has emerged from masked priming studies reporting a reduced morphological priming effect for L2 speakers, a finding which was interpreted by advocates of the SSH as reflecting a whole-form access to complex forms without decomposition and access to the decomposed constituents (Clahsen, Balkhair, Schutter, & Cunnings, 2013; Clahsen & Neubauer, 2010; Jacob, Heyer, & Veríssimo, 2018; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008). For example, Jacob et al. (2018) tested masked priming effects in highly proficient L2 German speakers (L1 Russian) and examined whether facilitation in target recognition occurs when primes are regular German participles with suffix *-t* (*geändert* 'changed') or derived nouns with productive *-ung* suffix (*Änderung* 'a change') that share a stem with the target (*ändern* 'to change'). Derived forms clearly facilitated the recognition of the target, but inflected forms did not. RTs of the inflected condition were not different from the unrelated condition and were even significantly slower than those of the derived condition. The results are consistent with the SSH claim that L2 access to inflected words like *walked* does not involve decomposition to [*walk* + *ed*], but they are accessed directly as whole forms. According to the SSH derivational priming does not necessarily involve decomposition. Since priming of derived forms is lexically mediated (namely, involved two lexical entries), the priming effect can stem from activation of the partial overlap between the lexical entries (*Änderung* and *ändern* have separate entries) rather than from decomposition (Kirkici & Clahsen, 2013). Inflectional priming, on the other hand, cannot be lexically mediated, as the prime and target belong to the same entry, and priming can be arguably explained only by morphological decomposition. Since morphological parsing is not fully operational in L2, reduced priming effects occurred mainly in inflected but less so in derived forms.

The *declarative/procedural* (D/P) model of Ullman (2004), and specifically its application to L2 acquisition (Ullman, 2005), makes similar predictions about L2 higher reliance on whole storage and less on rule-based computation of complex words. Yet, this is a neurocognitive model, which claims that the basis of the difference lies in maturation factors. The account is based on the distinction between two brain memory systems: *procedural* *memory* and *declarative memory*. Combinatorial rule-based computation and decomposition are operated at the procedural memory system, and the storage of words in the lexicon is part of the declarative memory. According to the model, maturational changes during childhood or adolescence lead to reduction of the procedural and enhancement of the declarative memory system. Therefore, individuals who acquire a second language at a later age will tend to over-rely on the declarative memory in L2 processing, also when processing complex word forms, which typically involves the procedural memory in L1 speakers. This leads to higher reliance on whole storage and less on rule-based computation of complex words. Findings of reduced morphological priming for L2 are thus also consistent with the D/P model.

Other linguistic accounts disagree with the claim that L2 speakers rely less on decomposition during processing of complex words. Several masked priming studies found similar priming effects for inflected forms in L1 and L2 speakers (Coughlin & Tremblay, 2015; Feldman, Kostić, Basnight-Brown, Đurđević, & Pastizzo, 2010; Foote, 2017; Voga, Anastassiadis-Symeonidis, & Giraudo, 2014), questioning the reliability of the findings showing reduced morphological priming in L2. Even when assuming such findings are reliable, the reduced morphological priming can stem from other sources, such as cognitive overload due to the brief presentation of the prime in masked priming, and not necessarily from lack of decomposition. Gor, Chrabaszcz and Cook (2017) have suggested that L2 lexical access involves decomposition, also of inflected forms, but it is rather inefficient compared to L1 decomposition. In an auditory lexical decision task, they tested late L2 Russian learners on noun targets with nominative and oblique cases, which had either null affixation (nominative: *zavod* 'factory', oblique: *bumag* 'paper') or overt affixation (oblique: *zavoda*, nominative: *bumaga*). While L1 speakers had longer RTs for the oblique forms, both with overt and null affixation, L2 did not show longer RTs for the oblique case, for either overt or null affixation (in experiment 1). They explained that L2 speakers did not process the complex morpho-syntactic information of the oblique case, regardless of whether it had a null or overt affixation, and therefore did not show processing costs like the L1 speakers. They simply stripped the overt affixes to reach the meaning of the stem. In experiment 2, when a modification of the design forced them process the whole form (including pseudowords with real stem and affixes, which did not allow to reach a correct answer only based on stem detection), L2 speakers with higher proficiency showed a native-like pattern. To summarize, they suggest that L2 morphological processing involves affix stripping, focusing on accessing the meaning of the stem, but underuses recombination of the affix and checking mechanism to access the morpho-syntactic information.

In a recent account, Hopp (2016, 2018) presented the *lexical bottleneck hypothesis* (LBH) and highlighted that lexical processing can affect L2 syntactic and morpho-syntactic parsing. L2 speakers possess diffused lexical representations and relatively small vocabulary size. Thus, their lexical retrieval and access is slower, greatly affected by frequency of the words (e.g., Bowden, Gelfand, Sanz, & Ullman, 2010; Clahsen & Neubauer, 2010; Neubauer & Clahsen, 2009). According to the LBH, difficulty in lexical retrieval can account for a large part of the L2 speakers' tendency for nonnative-like syntactic parsing that was reported in many sentence processing studies that examined grammatical gender (e.g., Sabourin & Stowe, 2008) and syntactic structure building (e.g., Roberst & Felser, 2011). For example, Hopp (2013) found that L2 German speakers (L1 English) who had a native-like gender assignment performance, also employed gender for prediction indistinguishably from the L1 speakers. In contrast, L2 speakers who had a partially native-like gender assignment, did not use gender for prediction, even though they could assign correctly the gender to the nouns in the specific task. According to Hopp, these results demonstrate that L2 speakers tend to rely less on grammatical gender in sentence prediction because their lexical representations are not firm enough, reflected by a higher level of errors in gender assignment. Since wrong gender assignment will lead to a prediction error, L2 speakers reduce their utility of gender for prediction. Once they have acquired native-like gender representations, they utilize gender for prediction as L1 speakers. Although the LBH does not directly rely on studies of word-level morphological processing, but on sentence processing evidence, its implications are relevant for morphological processing and can posit an alternative explanation for the reduced masked priming effects as stemming from inefficient and slow lexical retrieval that strongly hinder morphological processing of a prime that is presented for such a brief moment.

Overall, the picture that emerges from the empirical evidence is not consistent. This can be the result of high variance in characteristics of L2 speakers, larger than in typical L1 populations. Several of these characteristics have been previously shown to affect L2 processing, such as age of language acquisition (e.g., Veríssimo, Heyer, Jacob, & Clahsen, 2018), L1 properties (e.g., Ionin & Montrul, 2010), proficiency level (e.g., Gor et al., 2017) and level of language exposure (e.g., Dussias & Sagarra, 2007). Furthermore, high variance in L2 responses is common in general, even when the group is relatively homogenous (McDonald, 2006). Yet the inconsistency can be also attributed to the type of morphological phenomena tested, which mostly involve affix stripping, either processes of derivation (i.e., lexeme/lexical entry formation, such as suffix *-ness*) or inflection (spell-out of morpho-syntactic properties like person and gender, such as suffix *-ed*). The recognition of such forms largely overlaps with orthographic properties and might reflect L2 reading properties rather than morphological processing. Growing evidence has shown that L2 speakers are more prone to elicit orthographic priming effects, namely facilitation in target recognition when the prime and target overlap only in orthographic properties but not in morphology and semantics (e.g., *freeze–free*; Diependaele, Duñabeitia, Morris, & Keuleers, 2011; Feldman et al., 2010; Heyer & Clahsen, 2015; Li, Taft, & Xu, 2017; Qiao & Forster, 2017). It is possible that L2 speakers rely more on orthographic processes during visual target recognition, and thus the source of L2 morphological priming (at least when the prime and targets are visually presented) is orthographic in nature. Perhaps more fine-grained linguistic distinctions, which are less orthographic dependent, are needed to pinpoint the exact nature of L2 morphological processing and enable us to understand the subtle differences in the mental lexicon processes between L1 and L2 speakers. Recent studies attempted to examine more fine-grained morpho-syntactic features, either by testing the recognition of complex forms with null affixation, which allows to examine effects of morphological complexity independent of orthographic complexity (as described earlier in Gor et al., 2017), or by comparing recognition of inflected affixed forms that encode different morpho-syntactic features (Bosch & Clahsen, 2016). For example, in a masked priming study, Bosch and Clahsen (2016) examined three types of inflected German adjectives: *-e, -s,* and *-m*. Each type of inflected form was displayed as a prime and as a target (e.g., *geheimem–geheime* and *geheime–geheimem*) in a fully crossed design (3 x 3 = 9 conditions). Morphological priming was compared to identical priming (*geheime–geheime*) in L1 and highly proficient L2 speakers (L1 Russian). The results showed that unlike L1 speakers, prime type and target type did not significantly interact in the L2 group. While in the L1 group *-s* primes (*geheimes*) did not differ from the identity condition (*geheime*) in facilitating *-e* targets, but *-m* (*geheimem*) primes were slower than the identity condition, in the L2 group this kind of asymmetry did not exist. They concluded that the asymmetry in morphological priming patterns, despite similar orthographic overlap, implies that morpho-syntactic features of the affixes are accessible for L1, and the lack of such asymmetry suggests that they are not accessible for L2 in the initial form-based lexical access. This line of research is likely to lead to more precise understanding of L2 processing and representations of morphologically complex words, which goes beyond affix-stripping and surface orthographic properties.

Besides real-time recognition processes, another way to better understand the principles by which the L2 mental lexicon is organized is to investigate morphological generalization. Such L2 literature has focused on production of *existing* complex words. This body of research has shown that L2 speakers perform various morphological errors when producing complex word forms (e.g., McDonald & Roussel, 2010; Montrul, 2011; Parodi, Schwartz, & Clahsen, 2004), which reflect an attempt to extend morphological knowledge to unfamiliar (or less familiar) words, resulting in an error. Typically, these errors occur in spontaneous speech, even in highly proficient speakers. These errors are inconsistent; speakers can produce different morphological forms to the same word in different experimental sessions or within the same session, or judge the same morphological form differently in different time points (Johnson, Shenkman, Newport, & Medin, 1996). Errors include the use of bare forms when an inflected form is required (e.g., 'yesterday I *talk with my brother'), default forms (e.g., **fishes* as plural form of *fish* instead of *fish*) and over-regularization of regular forms (**drinked*; e.g., McCarthy, 2008; White et al., 2004). According to the *missing surface inflection hypothesis* (Prévost & White, 2000), the locus of these errors is in the realization of a particular form rather than in a deficit in abstract representations. The errors were claimed to be a result of a temporary breakdown in accessing the surface inflected form, despite fully internalized L2 morphological representations and complete acquisition.

Regarding the domain of morphological generalization to novel words, research is relatively scarce. A few studies have reported that L2 speakers are influenced by phonological similarity when producing or rating novel complex words (English: Cuskley et al., 2015; German: Hahne et al., 2006; Neubauer & Clahsen, 2009; Greek: Agathopoulou & Papadopoulou, 2009; Clahsen, Martzoukou, & Stavrakaki, 2010), but it is still not clear if they rely on phonological similarity and associations more than L1 speakers do. In an elicited production study, Cuskley et al. (2015) reported that production of irregular patterns of English past tense was affected by the phonological distance between the novel base word and other existing words which take an irregular past tense in both L1 and L2 speakers, failing to find a significant difference between the groups in the degree of phonological effect, concluding that L1 and L2 speakers rely to the same extent on phonological information in morphological generalization. On the other hand, in a study by Neubauer and Clahsen (2009), L2 speakers were shown to rely more on phonological properties in morphological generalization than L1 speakers. The authors examined generalization of both regular and irregular German inflection in an acceptability judgement task, focusing on participle forms of denominal verbs which are homonymous to existing irregular participles (which take the irregular suffix *-n* instead of the regular *-t*). For example, the denominal verb *verwachsen* derives from the noun *Wachs* 'wax' and is homonymous to *verwachsen* 'to grow' which takes the irregular suffix *-n* to form a participle. Denominal verbs have been previously shown to prefer regular participles in German (Marcus et al., 1995) and English (e.g., Kim, Marcus, Pinker, Hollander, & Coppola, 1994).

The regular inflection preference has been attributed by the authors to grammatical structure; in denominal verbs, the information about the verbal root, where irregularity is specified, is not accessed due to the structure of the denominal verb (see Selkirk, 1982, for further clarification), and therefore the application of the regular rule is not blocked. On the other hand, if an irregular inflection is preferred, the underlying assumption for this preference is that it arises from reliance on phonological associations, since it is phonologically identical to an irregular participle. The results showed that while L1 speakers preferred the regular form (*verwachst*) in cases of denominal verbs (semantic context was presented to imply the meaning of the verb), L2 speakers showed no preference, showing an overall similar rating for -*n* (*verwachsen*) and -*t* forms (*verwachst*). They speculated that for the L2 group two factors were competing, structural (preferring *verwachst*) and phonological (preferring *verwachsen*), resulting in no preference between regular and irregular inflections. They interpreted this pattern as indicating that L2 speakers rely less on morphological structure and more on phonological associations in morphological generalization.

The present work aims to deepen the understanding of L2 morphological processing and generalization. It seeks to reveal whether the acquisition of a second language at a later age, even at high proficiency levels, has certain implications for the structure of the mental representations and the processing of complex words, or whether L2 language users are able to employ the same language processes that are used for L1. At high levels of L2 proficiency, differences between L1 and L2 are very subtle, mainly found in the early phase of processing of complex and abstract syntactic and morphological phenomena, which do not fully overlap with semantic and orthographic/phonological properties. The next section introduces the phenomenon investigated in the present thesis: inflectional classes. Previous research on the processing of complex words has focused on a restricted set of morpheme-based phenomena, especially in L2 research (e.g., English past tense suffixation -*ed* or derivational suffixation -*ness*), strongly dependent on semantics and/or syntactic properties. The present work, on the other hand, focuses on a phenomenon that is more ‘purely’ morphological, inflectional classes, which grants the opportunity to test ‘morphology by itself’ (Aronoff, 1994). Furthermore, specifically for Hebrew, the morphological structure of inflectional classes is considered more abstract than in other non-Semitic languages, formed by a non-linear combination of constituents which are otherwise unpronounceable. Due to its genuine morphological and abstract nature, this phenomenon is a strong candidate to yield L1/L2 differences.

1.3 The Phenomenon Under Investigation: Inflectional Classes

Inflectional classes are common in a wide range of languages, such as Romance, Slavic and Semitic languages. According to Aronoff (1994), whose theoretical framework is adopted here, an inflectional class is “a set of lexemes whose members each select the same set of inflectional realizations” (p. 64); that is, lexemes (or lexical entries) that belong to the same inflectional class share an inflectional paradigm. Inflectional classes are argued to be abstract morphological effects of lexeme realization rules, which assign membership in the class. Once an abstract class is assigned, the class determines the inflectional paradigm, without encoding any syntactic or semantic information, unlike other morphological operations, such as plural or tense inflection.

To demonstrate Aronoff’s theoretical account of inflectional classes, we first consider Italian inflectional classes of verbs, also referred to as conjugation classes. In Italian, each verb belongs to one of three inflectional classes, which are identified by theme vowels in their infinitive form: theme vowel -*a*- for class I (e.g., *parlar* ‘to speak’), -*e*- for class II (e.g., *sparger* ‘to scatter’) and -*i*- for class III (e.g., *dormir* ‘to sleep’). The three inflectional classes are assigned as “a direct effect of lexeme realization rules” (Aronoff, 1994, p. 128). Following that, the inflectional class assigns a phonological marker, which is a specific theme vowel (e.g., [*parl+a*]). The theme vowel does not encode any syntactic or semantic information but is simply the phonological effect of the inflectional class and serves to identify its inflectional paradigm. To conclude, Aronoff (1994) proposes that inflectional classes are the morphological realization of lexeme formation rules and that they determine the inflectional paradigm of the verb. In the following section, we describe how Aronoff’s account fits also the case of Hebrew *binyanim*, inflectional classes with non-concatenative properties.

1.3.1 Binyanim

Hebrew, like all Semitic languages, presents verbal classes, called *binyanim* (singular form: *binyan;* see Arad [2005] for a detailed review). The term is often translated into English as *conjugations* or *verbal patterns*. Every verb belongs to one of seven binyanim; each binyan is identified by a specific vowel pattern (see Table 1 for an overview). Typically, for a Semitic language, the vowel pattern is combined non-concatenatively with a root to create a verb in its basic form (without affixation), which encodes the morpho-syntactic features of past tense, third person, singular and masculine (*CiCeC* [*Piel*] + *L-M-D* = *limed* ‘teach’). Each binyan has a full paradigm of prefixed and suffixed forms, inflected for tense, person, number and gender.

Table 1

Overview on the seven Hebrew binyanim

Class name	Vowel pattern	Pattern in Hebrew script	Example (L-M-D, R-G-S)	Example translation
Paal	CaCaC	_ _ _	LaMaD למד	(he) learned
Piel	CiCeC	_ _י_	LIMeD לימד	(he) taught
Hitpael	hitCaCeC	_ _ _הת	HiTLaMeD התלמד	(he) interned
Hifil	hiCCiC	_י _ _ה	HiRGIS הרגיש	(he) felt
Nifal	niCCaC	_ _ _נ	NiLMaD נלמד	(he) was learned
Pual	CuCaC	_ _ו_	LUMaD לומד	(he) was taught
Hufal	huCCaC	_ _ _הו	HURGaS הורגש	(he) was felt

Note: capital letters in the example column represent the consonants and vowels that are orthographically represented in the Hebrew script, while others are not represented in the Hebrew script, as the Hebrew script includes a very partial vowel representation.

Throughout the years, the linguistic nature of binyanim has been debated. According to some views, the binyanim are derivational categories, categories of word formation, since the non-linear combination of a root and a vowel pattern of a binyan necessarily forms a new lexical entry. A basic property of a derivation process is that its output is a new lexical entry, unlike inflection processes, which map grammatical properties to forms within the same lexical entry (Anderson, 1992; Clahsen, Sonnenstuhl, & Blevins, 2003). However, defining them as only derivational is problematic, because the binyanim have characteristics of an inflectional system as well. Unlike a derivational system, binyanim are obligatory – characteristic of inflectional system – in the sense that each verb belongs to one binyan. Therefore, the binyanim can best be described as both derivational and inflectional (Aronoff, 1994; see Figure 1): lexeme-formation rules (derivation) form a lexeme (lexical entry) and assign an abstract morphological property to the lexeme, namely membership in a class. At the same time, the binyan, as any inflectional class, determines the inflectional form of the verb. But unlike the Romance classes, the class does not assign a phonological marker for each class, expressed in the form of a theme vowel.

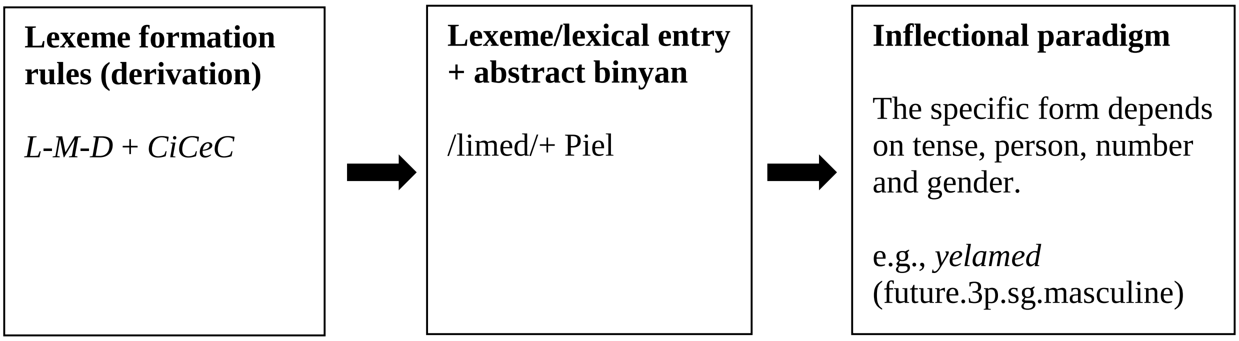

Figure 1. The derivational and inflectional roles of binyanim according to Aronoff (1994). The figure demonstrates that the lexical entry *limed* ‘teach’ is formed by derivational rules, which assign the binyan (i.e., the inflectional class) Piel to the entry. The binyan, Piel in this example, determines the inflectional paradigm of the entry.

The dual role of the binyanim as both derivational and inflectional classes[2] makes them a relatively unique phenomenon, similar to but different from the Romance inflectional classes. First, unlike the Romance verbal roots, the Semitic root does not carry the locus of all semantic, syntactic and morpho-syntactic information. Only after it is combined with a vowel pattern to form a lexeme, a full meaning can be formed. Therefore, the same root, unlike in typical inflectional classes, can appear in more than one binyan, more than one lexical entry. For example, the lexemes *LaMaD* ‘learn’ and *LIMeD* ‘teach’ have the same root *L-M-D* but have different entries and belong to different binyanim.

Second, the binyanim are not completely semantically and syntactically arbitrary like the typical inflectional classes. In the Hebrew linguistic literature (e.g., Arad, 2005; Berman, 1997), the binyanim are described as being semantically and syntactically compositional to some extent (i.e., the verb meaning and argument structure can be reconstructed from the individual constituents), while it is usually acknowledged that a large portion of the verbs is assigned

[2] Two of the seven (Pual and Hufal) are exclusively the passive form of another two binyanim (Piel and Hifil) and are not referred to as inflectional classes since they are fully predictable (Aronoff, 1994).

rather arbitrarily to binyanim. Most prominently, the lack of a blind and arbitrary class assignment is reflected in an argument structure bias for most binyanim (also referred to as a subcategorization bias) or even a clear-cut constraint. The binyanim *Nifaal* and *Hitpael* never appear with a direct object (i.e., they are never transitive). *Piel* and *Hifil* have a bias to appear with a direct object, and *Paal* is the only binyan which does not carry a specific argument structure bias, equally allowing both transitive and intransitive verbs to join (Berman, 1997).

Paal and Piel

Throughout all the experiments in the present dissertation, the binyanim *Paal* and *Piel* have been the center of investigation. The decision to focus on these classes was guided by certain unusual traits of these classes, especially Paal. The present section will clarify what those traits are and how a comparison of the two classes will contribute to highlighting the role of productivity and morphological structure in the organization and processing of lexical entries in the human mind.

The dissertation emphasizes two differences between Paal and Piel (see Table 2). The first relates to their differences in productivity, while maintaining both high type frequency and relatively similar phonological form. The vowel pattern *CaCaC* characterizes Paal verbs, and *CiCeC* characterizes Piel verbs, two vowel patterns that do not contain prefixes, unlike other binyanim. Looking at frequency in biblical Hebrew, Paal was the most common class, when 71% of Hebrew roots occurred in Paal (the second was Hifil with 24% of roots, then Piel with 23%; Aronoff, 1994). In Modern Hebrew, however, the status of Paal as the most common binyan is uncertain. In a similar analysis we conducted, this time based on a large corpus of Modern Hebrew (Itai & Wintner, 2008), 51.2% of all three-letter roots (n = 1563) occurred in Paal, and 45.3% in Piel. When including also longer roots (which cannot appear in Paal), the distribution shifts toward Piel: 53.4% of all roots (n = 1780) appear in Piel, and 42.8% in Paal. This shift between Piel and Paal can be attributed to the unproductive nature of Paal in Modern Hebrew. New verbs are not entered to Paal but rather to Piel (and Hitpael). For illustration, in a comparison between Hebrew dictionaries from 1963, 1972 and 1982, Bolozky (1999a) found that only two Paal verbs were added to the more recent dictionaries which demonstrates it is a closed class. For comparison, 23 Piel verbs and 22 Hitpael verbs were added in the same time frame, confirming their status as open classes. This is a clear example that high type frequency does not always imply high productivity, as also claimed by Baayen (1992; see Chapter 1.1). A few explanations were suggested in the literature for the odd property of Paal as a highly frequent class which is at the same time not productive in Modern Hebrew. First, all Paal verbs are constrained to roots of three consonants, whereas Piel verbs, as well as Hitpael ones, can accommodate roots of more than three consonants. This of course makes Piel and Hitpael highly attractive to accommodate novel verbs, which often originate from foreign languages like English or German and have at least four consonants (e.g., *TiLFeN* ‘to call (on the telephone)’ is a Piel verb with root *T-L-F-N*). Second, the Paal stem template is characterized by a large variation (along with Nifal, which is also not productive in Modern Hebrew), unlike the stable template of Piel and Hitpael. Paal has different stem templates for prefixed (CCVC) and unprefixed stems (CVCVC). In addition, the vowel pattern of these templates is not consistent; six different verb groups can be identified within Paal, each with different vowel pattern properties (Aronoff, 1994; Waltke & Occoner, 1990).

Table 2

A comparison between the properties of Paal and Piel

Class name	Vowel pattern	Argument structure bias	Productivity
Paal	*CaCaC*	no bias	restricted class
Piel	*CiCeC*	+direct object (79%)	unrestricted class

The second difference between Paal and Piel lies in their argument structure bias. While Piel verbs are mostly transitive, occurring with a direct object in their subcategorization frame (it is estimated that 79% of Piel verbs are transitive, based a large corpus [Itai & Wintner, 2008]), Paal verbs can equally be transitive or intransitive. To what extent are Hebrew speakers influenced by argument structure biases in the domain of morphological generalization? Previous studies examined the sensitivity to syntactic and semantic tendencies of the binyanim in generalization of novel verbs (Berman, 1993; Bolozky, 1999a). The results suggested that adults and eight-year-old children are sensitive to the semantic and syntactic tendencies of the binyanim when asked to produce a novel verb. For example, they produced more Piel verbs (and Hifil) verbs when the novel verb was supposed to express a causative action, as those classes are inclined to include more causative verbs. However, these studies did not examine the influence of argument structure alone. Since the meaning of the novel verb was always described when asked to produce the novel verb, argument structure was strongly correlated with semantics. Whether speakers are sensitive to the relationship between binyanim and argument structure itself remains unclear.

1.3.2 Processing and generalization of inflectional classes

A relatively small body of research has examined the processing and generalization of inflectional classes. Since the largest body of research in this topic addresses verbal inflectional classes of Romance languages, the current section focuses on them, in an attempt to understand the processing and generalization of inflectional classes in general.

Two different views have been proposed for the nature of processing and generalization of the Romance inflectional classes. In an extended version of the dual-route approach for Romance inflectional classes (Veríssimo & Clahsen [2009] for Portuguese verbs and Say and Clahsen [2002] for Italian verbs), it was suggested that word recognition and generalization of productive and unproductive classes involve different processes, reflecting different types of representations. Stems of the productive class I were suggested to be created by stem formation rules over variables (root + theme vowel -*a*-), so they can be applied unrestrictedly to any verbal root. This explains the high productivity of class I in the language. In contrast, stems of the unproductive classes II and III were claimed to not have an internal morphological structure, block the rule application and be stored as wholes. Their generalization is very restricted, based on similarity-driven mechanism, sensitive to phonological associations, which is the basis for the productivity contrast between class I and classes II and III.

Support for this approach was found in a cross-modal priming study on Portuguese verbs (Veríssimo & Clahsen, 2009). Morphologically related primes were either verbs in the infinitive form of class I (*limitar* ’to limit’) or class III (*adquirir* ‘to acquire’), and targets were their corresponding first-person present tense form (*limito* and *adquiro*), which does not include a theme vowel. The results have shown a full morphological priming effect for primes from the productive class I (facilitation that was identical to the facilitation yielded from primes identical to the target, *limito–limito*). In contrast, only a partial morphological priming effect was detected for the unproductive class III (facilitation that was smaller than the facilitation of the identical prime *adquiro–adquiro* but larger than with the unrelated prime). Given these findings, the authors concluded that the priming contrast arises from the different lexical representations of the classes (structured vs. unstructured). In the case of verbs from class I (*limitar* ‘to limit’), the full priming effect was argued to arise from a repeated activation of the same lexical representation in prime and target, the verbal root (*limit*-). On the other hand, the partial priming effect found for class III was considered to indicate an activation of a related but distinct representation, a stem-based representation (*adquiri*).

Additional support for the dual-route view of Romance classes was found in generalization studies (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014) which addressed the question of how language users expand their morphological knowledge about inflectional classes to novel forms, specifically examining how the production and/or acceptability judgment of each inflectional class is affected by the phonological similarity between existing roots and novel roots. For example, Veríssimo and Clahsen (2014) conducted an elicited production task of novel verbs in Portuguese. Participants were given a sentence containing a novel verb in the first-person present tense (a form that does not include a theme vowel), such as *acuo*, and in the next sentence they were asked to fill in the gap. The syntactic context of the second sentence required the production of the same verb but in the infinitive form. The novel forms in the study had a range of different degrees of phonological similarity to the three verbal classes, assessed by a computational model (the MGL; Albright, 2002). The algorithm of the model extracts context-based morpho-phonological rules and computes scores of reliability expressing phonological similarity to inflectional classes. Each novel root had a certain score for classes I, II and III. The results showed a clear phonological effect of the unproductive verbal classes II and III; that is, speakers were more likely to produce verbs from class II or III when the novel verbal root had a higher phonological similarity score to class II or III, accordingly. In contrast, class I, which is the only highly productive class, did not display any similarity effects. These results were interpreted by the authors as showing a clear contrast in the generalization properties of productive and unproductive Romance classes; generalization of unproductive classes is restricted to novel items that are similar to the existing items that take the same morphological pattern. The lack of effect for the productive class was interpreted as an indication that the productive class is not generalized by phonological similarity but rather by an application of symbolic rules, a default class that is applied regardless of similarity.

Nonetheless, the dual-route approach for inflectional classes has been criticized by single-route advocates. According to single-mechanism approaches, phonological similarity alone can explain generalization of inflectional classes, even for the productive class I. Supporting this claim are several studies of Italian inflectional class generalization (Albright, 2002; Colombo et al., 2006; Eddington, 2002), which found that the generalization patterns of Italian speakers can be predicted by several computational models, which involve phonological-based generalization. For instance, Albright (2002) tested Italian speakers in an acceptability judgement task of novel infinitives from different inflectional classes and compared the results to the outcome of a simulation of a computational algorithm, the MGL (same algorithm used in Veríssimo & Clahsen, 2014; see above). The outcome of the model simulation, reflected by scores of reliability (essentially phonological similarity), strongly correlated with the pattern of responses of the participants. Given these results, the author argued that even for the highly productive class I, speakers have internalized phonological regularities and are sensitive to the phonological context when extending inflectional classes to novel forms, and thus contradicting the dual-route approach for inflectional classes.

Taking together the different findings, it seems clear that phonological similarity is a central factor in inflectional class generalization in unproductive classes, a claim that is agreed upon among both the dual- and the single-route approaches. However, the role of phonology for highly productive classes is still under an intense debate between the two approaches.

1.4 Aims and Objectives of the Present Work

The present work addresses the question of whether the morphological processor and organization of morphological representations are universal, or whether languages and speakers differ fundamentally on the way complex words are represented and processed. To answer this question, the present project conducted a close examination of recognition and generalization of Hebrew *productive (Piel) and unproductive (Paal) inflectional classes* in L1 and L2 speakers. Through an examination of Piel and Paal throughout the four publications, the present work aims at assessing and understanding how productivity mediates the way complex words are represented and processed in the Hebrew mental lexicon, and whether the answer to it is shaped by the specific properties of Hebrew morphology and/or the language user characteristic as an L1 or L2 speaker.

The following research questions were targeted:

1. Hebrew morphological properties
    a. Compared to previously reported findings about Indo-European languages, does *recognition* of Hebrew complex words
        i. rely more on structured representations? (Publication I)
        ii. rely less on semantic transparency of morphological constituents? (Publication III)
    b. Does *generalization* of Hebrew inflectional classes (binyanim)
        i. rely on phonological similarity, a central source of information in morphological generalization of Indo-European languages? (Publication IV, experiment 1)

ii. rely on argument structure information, a more specific source of information in binyanim generalization? (Publication IV, experiment 2)

2. L1 speakers vs. L2 speakers
   a. Do L2 speakers have a reduced sensitivity to subtle morphological differences during complex word *recognition* compared to L1 speakers? (Publication II)
   b. Does complex word *generalization* in L2 speakers rely less on structural cues compared to L1 speakers? (Publication IV)

All questions were examined by a close investigation of the two most common Hebrew inflectional classes: Paal and Piel. Two properties of Paal and Piel are particularly relevant for the experimental work. First, as described earlier, Piel is a highly productive inflectional class, and Paal is an unproductive inflectional class. Within the framework of the dual-route approach (e.g., Pinker, 1999), productivity goes hand in hand with representation type in the mental lexicon: structured representations for productive morphological patterns and unstructured representations for unproductive ones. On the other hand, the most dominant view of the Hebrew mental lexicon argues that all root-based words in Hebrew have structured representations (e.g., Deutsch, 2016). Comparing Paal and Piel on recognition and generalization allows to reach a better understanding with regard to the degree to which Hebrew processing and generalization are unique, as argued by the root-based approach, or perhaps more similar to Indo-European languages, as suggested by the stem-based approach to Hebrew morphology (e.g., Ussishkin, 2005). Examination of Paal and Piel can even lead to more precise conclusions about the role of productivity than can be reached in examining the Romance inflectional classes, where class I is the most frequent and most productive in extending to novel forms, while class II and III are less frequent and less productive. Thus, in the Romance classes, it is not possible to disentangle the effect of frequency from productivity. In contrast, Paal and Piel, despite their robust productivity contrast, both have high type frequency, allowing a rare chance to disentangle frequency from productivity in the psycholinguistic field. Second, Paal and Piel are characterized by different argument structure properties. While most Piel verbs have an argument structure bias (+ direct object, i.e., transitive), Paal verbs do not have a bias and can be equally transitive or intransitive. These characteristics would assist in determining whether generalization of inflectional classes in Hebrew is affected by argument structure information, a source of information that to the author's knowledge does not play a role in inflectional class generalization in non-Semitic languages.

The questions in (1) investigate effects of language-specific properties, particularly effects of Hebrew-specific properties of inflectional classes, mainly the non-concatenative verb formation (root + verbal pattern: *L-M-D* + *CiCeC* = *LIMeD* 'teach') and argument structure bias that are associated with the binyanim. These properties do not exist in the typical inflectional classes, like in the Romance classes, which are based on concatenative verbal stem formation and on classes that are completely arbitrary in any semantic or syntactic terms. How do Hebrew-specific properties shape recognition (1a) and generalization (1b) processes of inflectional classes, in comparison to previous findings on Indo-European languages? Does verb recognition of Hebrew involve the same dual architecture that was proposed for the Romance classes, including both structured stems and unstructured stems (e.g., Say & Clahsen, 2002)? Or do the Hebrew verbal stems, as claimed by the root-based approach, always have root-based structured representations that are independent of semantic relatedness (Deutsch, 2016; Frost et al., 1997)?

To examine whether and to what extent recognition of Hebrew complex words relies more on morphological decomposition and less on a semantic relation between words (1a), lexical processing of Paal and Piel verbs was compared using the priming technique, with a focus on sensitivity to morphological roots during an initial stage of recognition, tapping into access representations and a later stage, tapping into central representations (Marslen-Wilson, 2007). In addition, effects of semantic transparency between primes and targets were examined. If Hebrew strictly relied on morphological decomposition independently from semantics (Deutsch, 2016), this should be apparent not only in the productive class, Piel (e.g., *LIMeD* 'teach' decomposed to *L-M-D* and *CiCeC*), but also in the unproductive class, Paal (e.g., *LaMaD* 'learn' decomposed to *L-M-D* and *CaCaC*). To examine the sources of information used during generalization (1b), production of novel Paal and Piel verbs was tested, examining whether phonological similarity (a common source for morphological generalization in Indo-European languages; e.g., Prasada & Pinker, 1993) is applied in the generalization of binyanim, that is, whether novel roots that are similar to existing roots from a certain class are more likely to be assigned to that class during novel verb formation. A second type of information was examined, argument structure, which correlates with the binyanim and is relatively specific to the Semitic classes. This was carried out to examine whether the specific properties of the binyanim – argument structure tendencies – affect the way generalization of inflectional classes occurs in Hebrew. If this is the case, Piel production is hypothesized to increase in the context of a direct object, as most Piel verbs are transitive.

The questions in (2) investigate effects of speaker-specific characteristics, particularly how L2 speakers recognize and generalize inflectional classes in comparison to L1 speakers. They address previous accounts which have argued that L2 speakers have a reduced sensitivity to abstract and structural information during complex word processing (e.g., the SSH; Clahsen & Felser, 2006a, 2006b, 2018). To answer question (2a) we examined whether during word recognition L2 speakers are sensitive to the subtle and abstract morphological difference between Paal and Piel, which are otherwise very similar. To answer question (2b), we examined how the argument structure bias of Piel, and the lack of bias of Paal, affect the way these verbs are generalized, given the hypothesized reduced sensitivity of L2 speakers to abstract information. The L2 groups throughout the project have been homogenous, restricted to late learners of Hebrew with very high proficiency levels, who had emigrated from South America and were fully immersed in the Israeli society. Only highly proficient L2 speakers were tested, since the purpose of the research project was to target end-state representations and processing of L2 speakers and not an ongoing acquisition of the L2 language. Furthermore, the L2 groups were Spanish L1 speakers, a language that contains verbal inflectional classes (with a few cases of L1 Portuguese, which is also a Romance language with verbal inflectional classes). Thus, they were all familiar with the phenomenon under investigation already from their L1 language. This has restricted possible variance in L2 group performance which stems from different L1 effects.

The findings from the four manuscripts for publication reported here contribute to a better understanding of the mental representations and processes involved in computing morphologically complex words, particularly verbs of productive and unproductive inflectional classes. The present dissertation thus extends the available research on theoretical models of L1 and L2 morphological processing and generalization by exploring (i) inflectional classes, a previously understudied linguistic phenomenon, (ii) in Hebrew, a language from a less-explored language family, (iii) collecting data from both L1 and L2 populations, and by (iv) examining both recognition of existing words and the formation of novel complex words.

2 Overview of Publications

In the following chapters, the results of the present work are presented in the form of four manuscripts, either published or under review. A summary of these four manuscripts is provided in the following section.

Publication I (first author; published in *The Quarterly Journal of Experimental Psychology*, 2018, doi: 10.1080/17470218.2017.1310917)

Universal and Particular in Morphological Processing: Evidence From Hebrew

Authors: Yael Farhy[1], João Veríssimo[1] and Harald Clahsen[1]

Summary: This study reexamined the prevalent claim that Semitic lexical access is more ‘morphological’ in nature and involves full morphological parsing (decomposition ‘down to the root’; e.g., Boudelaa & Marslen-Wilson, 2015) by investigating recognition of verbs following morphologically-related (namely, root-sharing) primes. Specifically, we directly compared the effects of primes from a productive inflectional class (*Piel*) and an unproductive inflectional class (*Paal*) in L1 Hebrew speakers. We tested whether these forms facilitated the recognition of targets sharing the same root (belonging to class *Hitpael*), which indicates morphological decomposition ‘down to the root’ in lexical access. There were two Form Type conditions (1singular.past, Infinitive), each with three Prime Types (Paal, Piel and Unrelated), presented for 50 ms immediately before target words. The results revealed different priming patterns for Piel and Paal: a significant root-priming effect for Piel (Le*XaLeK–HiTXaLeK*), but not for Paal (La*XLOK–HiTXaLeK)*, regardless of Form Type. Our results indicate that access to inflected verbs in Hebrew is modulated by abstract morphological information, namely class membership; access engages fully decomposed representations for the productive Piel class, but access to Paal verbs appears to be mediated by undecomposed stems from which the root is less directly available. In contrast to previous claims about the Semitic processor, we propose that processing in the Semitic mental lexicon relies on a division of labor between decomposition to separate constituents and a direct whole-stem access, similar to morphological processing in other languages.

Personal contribution: I prepared the experimental set-up, conducted all experimental sessions with participants, and performed the statistical analyses. I was involved in data interpretation. In addition, I wrote the methods and results, plus the first version of the introduction, and was involved in modification of the script following peer reviews.

Contribution of co-authors: Harald Clahsen was involved in data interpretation, wrote the final version of introduction and edited and finalized the manuscript. João Veríssimo was involved in planning the experimental set-up, analyzing and interpreting the data, he wrote the discussion section, edited and finalized the manuscript.

Publication II (first author; published in *Bilingualism: Language and Cognition,* 2018, doi: 10.1017/S1366728918000032)

Do Late Bilinguals Access Pure Morphology During Word Recognition? A Masked-Priming Study on Hebrew as a Second Language

Authors: Yael Farhy[1], João Veríssimo[1] and Harald Clahsen[1]

Summary: Previous research has proposed that the mechanisms for accessing morphological and morpho-syntactic information from visually presented words under masked-priming conditions are not fully operational in adult second language (L2) learners relative to native (L1) speakers (e.g., Silva & Clahsen, 2008, and much subsequent work). The study extends these findings to a less explored language type, Semitic, by reporting results from the first masked-priming study comparing L1 and L2 Hebrew speakers. We tested a group of highly proficient, fully immersed L2 Hebrew speakers, in a design and procedure identical to the one in Publication I, which allowed a direct comparison with the L1 group from Publication I. The results revealed a significant three-way interaction of Form Type, Prime Type, and Participant Group (L1, L2) indicating L1/L2 differences in masked morphological priming. While L1 speakers showed the same priming pattern for both Form Types (1sg.past, infinitive), with a significant root-priming effect for Piel but not for Paal, the L2 group showed a different pattern, with significant priming in the Infinitive condition for both Piel and Paal (e.g., Le*XaLeK*–*HiTXaLeK*), but not in the 1sg.past condition (e.g., *NiSaKTI*–*HiTNaSeK*). Our results indicate that (i) L2 speakers of Hebrew are less sensitive to the subtle morphological contrasts between Piel and Paal, at least under masked-priming conditions, and (ii) their access to forms which carry complex morpho-syntactic information (1sg.past) was more demanding compared to non-finite forms. Overall, the results suggest that non-native morphological processing has a reduced sensitivity to morphological and morpho-syntactic information.

Personal contribution: I prepared the experimental set-up, conducted most experimental sessions with participants, and performed the statistical analyses. In addition, I was involved in data interpretation and wrote the methods and results, plus the first version of the introduction.
Contribution of co-authors: Harald Clahsen was involved in data interpretation. He wrote the final version of introduction, edited and finalized the manuscript. João Veríssimo was involved in planning the experimental set-up, analyzing and interpreting the data, writing the discussion section, and editing and finalizing the manuscript.

Publication III (first author; published in *Language and Speech*, 2018, doi: 10.1177/0023830918811863)

Semantic Effects in Morphological Priming: The Case of Hebrew Stems

Authors: Yael Farhy[1] and João Veríssimo[1]

Summary: Previous research on Semitic lexical organization has argued that morphological effects are independent of semantic relatedness. The present study reexamined this claim by conducting a cross-modal priming experiment with a group of Hebrew L1 speakers, testing (i) whether verbs from a productive class (Piel) and an unproductive class (Paal) elicit root-priming effects (whether the auditory presentation of these primes facilitate the visual recognition of targets sharing a root, relatively to an unrelated word: /lexalek/–*HTXLK*, /laxlok/–*HTXLK*), and (ii) whether morphological priming effects are modulated by the degree of semantic relatedness between primes and targets. The results revealed significant root-priming effects for Piel and Paal. However, class type interacted with semantic relatedness; a larger degree of semantic relatedness yielded faster target recognition following Paal, but not Piel primes. We propose that stems of unproductive Hebrew classes (Paal) are stored as wholes leading to root priming which is mediated by semantics. In contrast, stems of productive classes (Piel) constitute structured representations and activate their roots directly. Our results challenge accounts in which all Semitic morphological effects are independent of semantic relatedness, as well as accounts that dismiss structured representations altogether. Instead, they support a dual-morphology system in which constituent structure is closely aligned with productivity.

Personal contribution: I prepared the experimental set-up, conducted all experimental sessions with participants, and performed all statistical analyses. In addition, I was involved in data interpretation and wrote the introduction, methods, results, and first version of discussion, and in modification of the script following peer reviews.

Contribution of co-author: João Veríssimo was involved in planning the experimental set-up, analyzing and interpreting the data, writing the final discussion section, and editing and finalizing the manuscript.

Publication IV (single author; accepted for publication in *The Mental Lexicon*)

Morphological Generalization of Hebrew Verb Classes: An Elicited Production Study in Native and Non-Native Speakers

Author: Yael Farhy[1]

Summary: This study examined properties of morphological generalization in Hebrew inflectional classes (particularly the common classes Paal and Piel), focusing on the role of two information sources in formation of novel verbs: (i) phonological similarity, found to be a very robust source of generalization across languages, and specifically in inflectional classes of non-Semitic languages (e.g., Albright, 2002), and (ii) argument structure, which is correlated with inflectional classes in Hebrew, but not other inflectional classes, like the Romance classes. In addition to L1 speakers of Hebrew, proficient non-native (L2) speakers of Hebrew were also examined to determine how morphological generalization is affected by specific characteristics of language users. In two elicited production tasks, participants were asked to fill in the blank with a novel verb from the novel noun presented in a previous sentence (e.g., novel noun *desel* (root *D-S-L*), possible answers: *dasalti* (Paal), *disalti* (Piel) and so on).

Experiment 1 examined phonological similarity, that is, whether phonological similarity of a novel root to existing roots affect verbal class generalization. Based on the implementation of an analogy-based model (*Analogical Modeling of Language*; Skousen et al., 2002), three phonological similarity conditions were included: (i) novel roots similar to Paal, (ii) novel roots similar to Piel and (iii) novel roots not similar to any class. The results showed, for both groups, a small but significant increase in Piel responses in the Piel similarity condition compared to the no-similarity condition. No further differences were found between the similarity conditions in Piel and Paal responses. In experiment 2, argument structure was manipulated in two conditions, one with a Direct Object (+DO) after the blank for the novel verb form, and one without (–DO). The results of experiment 2 showed a very robust increase in Piel responses in the +DO condition compared to the –DO condition for both the L1 and L2 groups, although the argument structure effect was significantly weaker for the L2 compared to the L1 group. Paal responses remained constant across conditions. The results mirrored the bias of Piel in the language towards an argument structure with a direct object, as well as the lack of one for Paal. The findings show that phonological similarity effects are less ubiquitous for morphological generalization than previously thought. For Hebrew inflectional class generalization, they do

not play a crucial role. Instead, at least in this case, argument structure is a more relevant predictor of both L1 and L2 speakers' morphological generalization. We conclude that the sources of morphological generalization are (at least in part) language-specific.

Personal contribution: I conceived, set up, programmed and ran the experiments. I also acquired, analyzed and interpreted the data. Finally, I wrote and edited the submitted manuscript.

Author's Affiliations:

1. Potsdam Research Institute for Multilingualism, University of Potsdam, 14467, Potsdam, Germany

3 Publication I

Published in The *Quarterly Journal of Experimental Psychology*, 2018, doi: 10.1080/17470218.2017.1310917

Universal and Particular in Morphological Processing: Evidence From Hebrew

Yael Farhy, João Veríssimo and Harald Clahsen
Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam

Abstract

Do properties of individual languages shape the mechanisms by which they are processed? By virtue of their non-concatenative morphological structure, the recognition of complex words in Semitic languages has been argued to rely strongly on morphological information and on decomposition into root and pattern constituents. Here, we report results from a masked priming experiment in Hebrew in which we contrasted verb forms belonging to two morphological classes, Paal and Piel, which display similar properties, but crucially differ on whether they are extended to novel verbs. Verbs from the open-class Piel elicited familiar root priming effects, but verbs from the closed-class Paal did not. Our findings indicate that, similarly to other (e.g., Indo-European) languages, down-to-the-root decomposition in Hebrew does not apply to stems of non-productive verbal classes. We conclude that the Semitic word processor is less unique than previously thought: Although it operates on morphological units that are combined in a non-linear way, it engages the same universal mechanisms of storage and computation as those seen in other languages.

How do properties of individual languages shape supposedly universal mechanisms of language processing? In the current study, this question is investigated with respect to morphology, a common source of cross-linguistic variability. Languages differ considerably in the ways in which they express morphosyntactic information—for example, via concatenative structures (e.g., *walk* + *ed*), marked stems (e.g., Portuguese: *fiz-* "did"), or periphrasis (e.g., Vietnamese: *đã đi* "went"). For experimental psycholinguistics, such variability raises the question of how the mental representation and processing of complex words are affected by properties of individual languages (e.g., Bick, Goelman, & Frost, 2011; Frost, Forster, & Deutsch, 1997). Specifically, mechanisms of word recognition might differ across languages, directly reflecting this cross-linguistic variability, or instead, might be abstract and general enough to apply to different kinds of morphological encoding.

Experimental studies of Semitic languages, like Hebrew and Arabic, have featured prominently in this debate, by virtue of their salient and pervasive *non-concatenative* morphology. That is, besides the linear combination of stems and affixes that is common in many languages, the formation of stems in Semitic languages involves the non-linear combination of consonantal *roots*, carrying core meaning, and vowel *patterns*, which may also express grammatical information (e.g., Hebrew: *L-M-D* + *taCCiC* = *talmid* "pupil"). It has been proposed that this property drives the Semitic lexical processor to be primarily "morphological" in nature, designed to extract a complex word's abstract structure (root and word pattern), irrespective of meaning or surface form (Boudelaa & Marslen-Wilson, 2015; Frost et al., 1997). By contrast, the word recognition system of Indo-European languages, such as English, is thought to be less purely driven by morphology, but instead more affected by factors such as semantic transparency and orthographic similarity (Velan & Frost, 2011). For example, while semantically opaque forms in English (e.g., *business–busy*) typically do not produce morphological facilitation effects in overt priming experiments (e.g., Gonnerman, Seidenberg, & Andersen, 2007; Marslen-Wilson, Tyler, Waksler, & Older, 1994), in Hebrew and Arabic, morphological priming is also obtained between opaque forms that share a root (Boudelaa & Marslen-Wilson, 2005; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000).

According to Bick et al. (2011), the contrast between experimental effects in English and Semitic arises because word forms in English often cannot be straightforwardly mapped onto morphemes (e.g., {business} ≠ {busy} + {-ness}), whereas in Semitic languages almost all words are morphologically structured. Therefore, the extraction of root information provided by full morphological parsing ("down-to-the-root") is thought to be a main priority of the

Semitic lexical processor (e.g., Boudelaa & Marslen-Wilson, 2011). Alternatively, in distributed connectionist accounts, morphological knowledge is represented as associations between full forms and their meanings (Gonnerman et al., 2007), such that languages with more inconsistent form-to-meaning mappings (i.e., languages with a "richer" morphology, such as Semitic) are supposed to show stronger morphological effects that are less dependent on semantics (Plaut & Gonnerman, 2000).

Against this background, the present study examines inflected verb forms in Hebrew, which have previously been claimed to be morphologically represented and fully decomposed into roots and patterns in lexical access (e.g., Deutsch, Frost, & Forster, 1998). Following much previous research, we employed the masked priming paradigm, a technique that has been found to be particularly sensitive to morphological structure (e.g., Frost et al., 1997; Marslen-Wilson, 2007). We specifically investigated root-priming effects from inflected forms of different Hebrew verbal classes called *binyanim* (singular: *binyan*). Although morphemic decomposition (as revealed by root-priming effects) has been argued to be the primary step of Hebrew word recognition, no study has examined potential differences between the various Hebrew verbal classes in the process of word recognition.

To preview, our findings provide only partial support for a down-to-the-root-parsing approach in Hebrew. Whilst complex forms that belong to an "open class" (i.e., that contain a productive word pattern) were indeed found to be fully decomposed, this was not the case for inflected verb forms that belong to a "closed class" that does not extend to new verbs. We will conclude that the Semitic morphological processing system is less unique than previously thought. Although Semitic languages employ non-concatenative root-and-pattern combinations, they nevertheless show the same alignment between productivity and decomposition that is seen in many other languages.

Hebrew verbal morphology

There are seven verbal classes or binyanim in Hebrew, defined by their particular vowel patterns. As mentioned above, these vowel patterns combine non-linearly with the consonants of the root, such that both root and pattern surface in a discontinuous way. Furthermore, the same root sometimes appears in more than one verbal class, creating different verbs. For example, verbs of the binyan "Paal" display the pattern *CaCaC*,[3]which when combined with the root *L-M-D* yields the verb *lamad* "he learned". In another binyan, "Piel", the same root is combined with the pattern *CiCeC*, to yield a verb with a different meaning, *limed* "he taught".

[3] We use *C* to represent root consonants.

Although the Hebrew binyanim display certain regularities that suggest abstract syntactic and semantic properties, only two binyanim (Pual and Hufal) are completely predictable (as passive analogues of two other binyanim). The other five show only tendencies that are far from deterministic (for review, see Arad, 2005). In fact, many verbs have highly lexicalized meanings, which are generally difficult or impossible to compute compositionally on the basis of individual roots and patterns. Furthermore, the system is “filled with holes” (Aronoff, 1994, p. 124), with almost no roots occurring in all binyanim. Such properties invite a treatment of the Hebrew binyanim as an example of derivational morphology (Waltke & O’Connor, 1990), but Aronoff notes that binyanim assignment differs from derivation by being obligatory. That is, while underived lexemes exist in every language, Hebrew verbs cannot exist just as roots and have to be assigned to a binyan in order to be properly inflected. Since binyan membership determines the particular shape of every form of a verb, but does not clearly encode specific syntactic or semantic properties, the Hebrew binyanim are not morphemes in any meaningful sense, but are better conceived of as abstract morphological categories.

The present study

The current study contrasts priming effects produced by verbs of the Paal and Piel binyanim, two verb classes that display comparable type frequencies and that, as a whole, are similar in terms of the general syntactic and semantic properties of their verbs (see Table 1). Nevertheless, there is a striking difference between the two: The Paal binyan constitutes a closed class, which “plays no role at all in the formation of new verbs”, whereas Piel is readily extended and, in fact, “the most important binyan for forming new verbs” (Aronoff, 1994, p. 130). This difference has been demonstrated, for example, by the longitudinal examination of neologisms, as well as in elicited production experiments (Bolozky, 1999a).

Table 1

Properties of the Paal and Piel verb classes

Binyan	Phonological base form	Example	Semantic properties	Type frequency
Paal	*CaCaC*	*lamad* ‘learned’	active	19.4 %
Piel	*CiCeC*	*limed* ‘taught’	active	17.1 %

Note. The type frequency percentages were calculated from a corpus containing 4,131 verbs (Itai & Wintner, 2008).

From a dual-morphology perspective (e.g., Clahsen, 1999; Pinker, 1999), the discrepancy in the productivity of Paal and Piel points to possible representational differences, particularly with regard to the contrast between structured and undecomposed stems. Highly productive morphological operations—which extend readily to novel items—are likely to be *rule based*. That is, they are the result of operations over variables, placeholders that stand for whole grammatical categories like "verbal root" (Marcus, 2001; Pinker & Ullman, 2002; Veríssimo & Clahsen, 2014). At the same time, rules are combinatorial operations that generate structured representations and that can be employed in processing to (de)compose stems and word forms from (or into) their morphological constituents. The link between productivity and constituent structure can be seen across language families and morphological operations. A case in point is regular inflection in Germanic languages (e.g., the English *-ed* past tense), which generalizes widely to novel verbs and produces experimental effects that are indicative of structured representations, such as robust priming on the recognition of their bases (e.g., Marslen-Wilson & Tyler, 1998; Newman, Ullman, Pancheva, Waligura, & Neville, 2007; Stanners, Neiser, Hernon, & Hall, 1979). Another example, closer to the Hebrew binyanim, comes from languages with conjugation classes. In Portuguese, for example, the class that extends to novel verbs also displays priming effects that indicate down-to-the-root decomposition (Veríssimo & Clahsen, 2009). In contrast, members of unproductive classes—which only rarely welcome new members—are predicted by dual-morphology accounts to be lexically stored as exceptions to general morphological rules. Therefore, they do not activate their roots "directly", via decomposition into morphological constituents, but are instead argued to involve whole-form access and processing. Accordingly, such forms typically produce reduced priming effects on their bases, even when they are phonologically and semantically transparent (e.g., Sonnenstuhl, Eisenbeiss, & Clahsen, 1999; Stanners et al., 1979; Veríssimo & Clahsen, 2009).[4] Within Semitic languages, further support for the relationship between pattern productivity and morphological constituency comes from a recent study by Wray (2016) with Arabic speakers. In a series of auditory lexical decision experiments, response times for forms belonging to both productive and non-productive binyanim were found to be predicted by word-form frequency, whereas only the recognition of productive binyanim was predicted by the frequency of the root.

[4] The studies mentioned here have employed overt priming paradigms, which (unlike the masked priming technique that is used in the present study) may be susceptible to strong effects of semantic relatedness. Nevertheless, all studies contrasted verb classes that were perfectly matched with regard to semantic transparency, indicating a morphological (rather than semantic) source for the contrast between productive and non-productive operations.

If the rationale that we have laid out applies to Hebrew verb forms, then genuine root-priming effects (signalling full morphemic decomposition) should be restricted to verbs of open classes, like the Piel binyan. Conversely, we hypothesized that verbal stems belonging to a closed class, like Paal, are not related to their roots by rule, but are accessed via an unstructured stem representation. If that is the case, they should fail to produce the typical root-priming effect that has been observed in previous priming studies in Semitic (e.g., Boudelaa & Marslen-Wilson, 2015; Deutsch et al., 1998).

Since the same root can appear in different binyanim, Hebrew allows the opportunity to compare priming effects elicited by different verbal classes (Paal and Piel) on the very same target words. In the current study, targets were verbs belonging to the Hitpael binyan that share a root with their morphologically related (Paal or Piel) primes. Hitpael verbs were used as targets, because (a) they display a pattern that is productive in new word formation (Bolozky, 1999a), (b) their forms are not homographic with other verbal forms, and (c) Hitpael verbs do not have a systematic or predictable semantic relation to either Paal or Piel verbs.[5]

Two sets of items were included in the present experiment, one set in which targets were preceded by primes in the first-person singular past form (1sg past), and another for which primes were presented in the infinitive (infinitive condition). This allowed us to assess the replicability of priming effects across items, by examining whether the same contrasts were obtained with another set of target words, as well as with primes presented in a different verbal form. In addition, these specific verbal forms were selected to control for possible orthographic effects. All primes–target triplets shared three consonant letters (the root), but 1sg past forms of Piel verbs contained an additional letter (a vowel) that was not present in Hitpael targets; for primes presented in the infinitive, it was instead the Paal forms that contained an additional letter.

Method

Participants

Thirty native speakers of Hebrew (20 females, 3 left-handed) between the ages of 18 and 37 years (mean: 28.75 years) participated in the experiment. All participants were born in Israel, had completed at least 12 years of education, and used spoken and written Hebrew on a daily

[5] In a small number of roots ($n = 4$), however, the *phonological* similarity between Piel and Hitpael verbs is slightly larger than is normally the case, because their patterns show a predictable phonemic alternation between stops and fricatives. We return to this issue in the Discussion section and present an analysis without the items in which this phonemic change occurs.

basis. They all had normal or corrected-to-normal vision, and none had been diagnosed with any language disorders.

Materials

Table 2 displays the experimental design, including an example stimulus set. Experimental targets consisted of 42 Hitpael verb forms in the 3sg past, a form that does not display inflectional affixes (i.e., it is constituted only by root + pattern). Each target word was paired with three types of primes: (a) a prime belonging to the Paal class, based on the same verbal root as the target, (b) a prime belonging to the Piel class, also based on the same root, and (c) an unrelated prime. Half (21) of the targets were preceded by primes presented in the 1sg past form (1sg past condition), and the other half of the targets were preceded by primes presented in the infinitive form (infinitive condition). Neither primes nor targets were homographic with any other form in Hebrew. A list of all experimental primes and targets employed in this study is presented in the Supplemental Material.

Table 2

Experimental conditions, with an example stimulus set

	Prime			
Form	Unrelated	Paal	Piel	Target (Hitpael)
1sg Past	טיפסתי TIPaSTI ‘I climbed’	נשקתי NaShaKTI ‘I kissed/touched’	נישקתי NIShaKTI ‘I kissed’	התנשק HiTNaSheK ‘he kissed’ (reciprocal)
Infinitive	לבחור LiVXOR ‘to choose’	ללמוד LiLMOD ‘to learn’	ללמד LeLaMeD ‘to teach’	התלמד HiTLaMeD ‘he did an internship’

Note. Examples include both Hebrew orthographic forms and their phonological form in Latin script (upper case letters represent letters that are present in the Hebrew orthographic form, in which vowels are typically omitted).

Table 3 displays means and standard deviations of different stimuli properties, for each experimental condition. Frequency values were based on a corpus of over 130 million tokens (Itai & Wintner, 2008) and are expressed in the Zipf scale (i.e., log10 of frequency per billion; van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Semantic relatedness scores for each prime–target pair were obtained in a pretest conducted with 26 native speakers of Hebrew, who were asked to rate semantic relations between infinitive forms on a scale from 1 ("Not at all related") to 7 ("Very related").

As can be seen in Table 3, targets in the 1sg past and infinitive conditions were very closely matched in their mean values of lemma frequency and number of letters. Within each condition, Paal and Piel primes were also closely matched in mean lemma frequency and in their semantic relatedness to Hitpael verbs. Unrelated primes were based on different verbal roots (i.e., morphologically unrelated to the target) and had no orthographic, phonological, or semantic overlap with their corresponding target forms. Half of the unrelated primes belonged to the Paal binyan, and half belonged to the Piel binyan. Unrelated primes were matched in mean lemma frequency to both Paal and Piel primes, in both the 1sg past and the infinitive conditions.

Table 3

Means and SDs (in parenthesis) of stimulus properties, for all conditions

Condition	Lemma frequency (Zipf)	Semantic relatedness (1–7)	Length (in letters)
1sg Past			
Unrelated	4.27 (0.74)	1.40 (0.33)	5.52 (0.51)
Paal	4.21 (0.97)	3.78 (0.78)	5.05 (0.22)
Piel	4.37 (0.64)	3.97 (0.84)	6.05 (0.21)
Target	3.31 (0.83)		5.00 (0.00)
Infinitive			
Unrelated	4.43 (0.58)	1.42 (0.25)	4.57 (0.51)
Paal	4.33 (0.61)	3.57 (0.94)	5.00 (0.00)
Piel	4.52 (0.70)	3.78 (0.80)	4.00 (0.00)
Target	3.29 (0.92)		5.00 (0.00)

One reviewer expressed the concern that spelling in Hebrew is often inconsistent, particularly with regard to the omission and redundant insertion of vowel letters (see, e.g., Ravid & Kubi, 2003). As mentioned above, the orthographic forms in the different conditions differed in vowel letters (see Table 2), and this may conceivably have processing consequences. In particular, forms that are potentially subject to inconsistent spelling (i.e., forms in the Paal 1sg past and the Piel infinitive conditions, in which vowels are absent in non-pointed script) may take more time to process, perhaps reducing priming effects. In order to ensure that the particular materials that we have employed are spelled in a consistent way by adult speakers, we have conducted a spelling experiment with 20 native Hebrew speakers. All prime words of the Paal and Piel conditions were included in this experiment. Prime words were divided into two lists, so that Piel and Paal verbs with the same root would not be presented to the same participant. Words were presented auditorily, and participants were asked to type every word. In 96.43% of the responses, spelling of Paal and Piel primes was accurate. Vowel omissions (e.g., HPXTI instead of HIPXTI after hearing the word /hipaxti/) occurred in only 0.83% of the responses (7 responses), and there were no vowel insertions. From these results, we conclude that the prime words used in the present masked priming experiment are spelled in a stable way by native speakers. Therefore, any potential reduction of priming effects in the Paal 1sg past and Piel infinitive conditions cannot plausibly be attributed to differences in the consistency of their spelling.

A set of 294 filler prime-target pairs were also included in the experiment (126 word-word pairs and 168 word-nonword pairs). Therefore, every participant saw 336 targets, half of them words and half pseudowords. Critical prime–target pairs were distributed over three experimental lists, so that each target appeared only once in each list. Participants were randomly assigned to one of the three lists, such that each list was presented to 10 participants.

Procedure

Participants were tested individually in a quiet room. They were asked to perform a lexical decision task on visual targets, as quickly and accurately as possible. Specifically, they were instructed to press a gamepad button (labelled “Yes”), using their dominant hand, when they recognized an existing word in Hebrew, and to press another button (labelled “No”), using their non-dominant hand, when they were presented with pseudowords. The DMDX software (Forster & Forster, 2003) was used for stimulus presentation and data collection. The experiment started with a practice phase including 12 trials (6 words and 6 pseudowords). The 336 experimental trials were then presented in a pseudo-randomized order, with three breaks provided during the experiment. Every trial consisted of the following events, in immediate succession: a fixation cross (500 ms), a blank screen (500 ms), a row of hash marks (500 ms), a prime word (50 ms), and the target (presented until a response was made, up to a timeout of 2000 ms). Therefore, the stimulus onset asynchrony (SOA) between prime and target was 50 ms. Response times (RTs) were measured from the onset of target presentation. Primes and targets were presented in Arial font, primes in size 20 and targets in size 24. After the experiment, participants were asked a set of questions that probed for awareness of prime words. The whole session lasted approximately 30 minutes.

Data analysis

Two items from the 1sg past condition were removed from subsequent analyses due to very low accuracy (below 50%): *hidama* “was similar” and *hishtamer* “was preserved”. All other items had accuracy rates of at least 80%. No participants were excluded. Incorrect responses (4.0%) and timeouts (0.4%) were removed from the dataset. In order to reduce the influence of outliers, extremely slow RTs (above 1500 ms) were discarded (1.3% of the remaining data). Finally, the distribution of RTs was normalized by applying a reciprocal transformation (i.e., −1000/RT; Baayen & Milin, 2010).

Reciprocal RTs were analysed using mixed-effects linear regression, with crossed random effects for participants and items. The following fixed predictors were included: (a) prime type (unrelated, Paal, Piel), (b) form type (1sg past, infinitive), (c) the prime type by form type interaction, and (d) trial (the position of each item in the experiment, centred). The factors prime type and form type were assigned treatment contrasts. Therefore, model estimates reflected comparisons against reference levels, and the statistical comparisons of interest were obtained by relevelling one or both factors and refitting the model.

In order to reduce the probability of Type I errors without sacrificing statistical power, we followed the recommendation of Matuschek, Kliegl, Vasishth, Baayen, and Bates (2015) and included random slopes if they improved model fit (as measured by Akaike information criterion, AIC). All possible random structures of prime type, form type, and their interaction were assessed. The best model (i.e., the one with the lowest AIC) contained no random slopes and is reported below.

Results

Table 4 displays mean RTs, standard errors (*SE*s), and accuracy rates in all conditions. Means and *SE*s were calculated from reciprocal RTs and were back-transformed. Accuracy rates were very high and were comparable across conditions. Therefore, they were not further analysed.

The results of the mixed-effects regression analysis are presented in Table 5. The two item sets were examined separately by changing the reference level of the form type variable. First, the effect of prime type on RTs was examined for the set of 1sg past targets, by comparing Paal and Piel primes against the unrelated baseline. The results show that the previous presentation of morphologically related Paal forms did not facilitate target recognition, as they elicited comparable RTs to those for primes that had no morphological relation to the target. In contrast, RTs were significantly faster after the presentation of Piel forms than after unrelated primes. Secondly, the effect of prime type on RTs was examined for the infinitive set, by relevelling the form type factor. As was the case in the 1sg past set of items, Paal primes presented in the infinitive also failed to facilitate lexical decision responses, but RTs after Piel primes were significantly faster than after unrelated primes. Consistent with this pattern, there were no interactions between form type and the levels of prime type—that is, magnitudes of priming elicited by Paal and Piel verbs were not modulated by whether they were presented as infinitives or as 1sg forms.

Table 4

Back-transformed means and SEs, and accuracy rates, for each condition

Form Type	Unrelated		Paal		Piel	
	RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
1sg Past	634 (10.94)	94%	629 (10.34)	97%	609 (9.46)	97%
Infinitive	641 (10.08)	94%	639 (10.91)	94%	620 (10.81)	97%

Because no interactions were present, Paal and Piel priming was also assessed across all items in both the infinitive and 1sg past conditions (by assigning "main effect" contrasts to Form Type, i.e., converting the factor to a numeric format and centring it). This model also showed significantly shorter RTs after Piel primes than after unrelated primes ($b = -0.0611$, $t = -3.14$, $p = .002$), but no priming effect from Paal verbs ($b = -0.0068$, $t = -0.35$, $p = .728$). Furthermore, Piel primes elicited significantly faster RTs than Paal primes ($b = -0.0543$, $t = -2.80$, $p = .006$)—that is, larger root priming was obtained from verbs belonging to the Piel binyan than from verbs of the Paal binyan.

Discussion

The main finding of the present study is that forms of Hebrew verbs belonging to the Paal and Piel verb classes (or *binyanim*) produce different masked priming effects on the recognition of verbs that share the same root. A clear dissociation between the two classes was obtained: The presentation of Piel verbs produced robust priming effects, while Paal verbs did not facilitate target recognition. Importantly, this pattern of results was replicated in two different sets of primes and targets, one for which all primes were presented in the 1sg past form and another for which all primes were presented in the infinitive form.

Table 5

Results from a mixed-effects regression on reciprocal RTs, with Infinitive and 1sg Past as reference levels

Fixed effect	Estimate (*b*)	SE	*t*-value	*p*-value
Reference for Form Type:1sg Past				
Intercept	-1.5696	0.0536	-29.28	<.001
Prime Type: Paal (vs. Unrelated, 1sg Past)	-0.0167	0.0281	-0.59	.554
Prime Type: Piel (vs. Unrelated, 1sg Past)	-0.0634	0.0283	-2.24	.024*
Form Type (in Unrelated)	0.0220	0.0509	0.43	.666
Prime Type: Paal X Form Type	0.0190	0.0391	0.49	.626
Prime Type: Piel X Form Type	0.0045	0.0390	0.12	.908
Trial (centered)	-0.0002	<0.0001	-2.10	.036*
Reference for Form Type: Infinitive				
Intercept	-1.5475	0.0524	-29.53	<.001
Prime Type: Paal (vs. Unrelated, Infinitive)	0.0023	0.0271	0.08	.932
Prime Type: Piel (vs. Unrelated, Infinitive)	-0.0590	0.0269	-2.20	.028*

**p*<.05

Note. Redundant coefficients for the Infinitive reference level are not shown (i.e., Form Type, Prime Type X Form Type, and Trial), as these are identical to the ones with 1sg Past reference.

We interpret these results as evidence that the early stages of visual lexical access in Hebrew are modulated by abstract morphological information—namely, binyanim membership. Accounts that instead attribute morphological effects to prime–target overlap in form and meaning (e.g., Gonnerman et al., 2007), or to the consistency of mappings between orthographic and semantic representations across the language (Plaut & Gonnerman, 2000), cannot easily explain our results, for several reasons. First, Paal and Piel primes were closely matched with respect to their semantic transparency to the Hitpael targets. Secondly, Paal and Piel primes were also very closely matched in their orthographic overlap with the targets, such that exactly the same three letters in the target (the root) were activated by all morphologically related primes. A small orthographic difference between Paal and Piel primes was indeed present within each item set: Piel primes presented in the 1sg past contained an additional letter that was not present in the target, but this difference was reversed for primes presented in the infinitive (in which Paal primes contained an additional letter); nevertheless, exactly the same priming pattern was obtained in both item sets. Thirdly, the masked priming paradigm typically produces reduced semantic effects at short SOAs, as well as small word-to-word orthographic effects (and especially so in Hebrew; Velan & Frost, 2007, 2009). Finally, the Paal and Piel classes are, as a whole, remarkably similar in a range of morphological and non-morphological properties: (a) They form structured stems, constituted by vowel patterns and productive consonantal roots; (b) their vowel pattern morphemes do not encode specific syntactic or semantic information; and (c) they have comparable type frequencies in the language. Therefore, we conclude that the priming pattern that we obtained cannot be explained by differences in formal or semantic overlap or in the consistency and strength of form-to-meaning mappings across the language.

It is true, however, that in certain verbs the second root consonant of the Hitpael targets is phonologically more similar to its counterpart in Piel, than in Paal. In particular, certain root consonants in the second position (*K*, *B*, *P*) surface as stops in Piel and Hitpael verbs (e.g., איבדתי /i**b**adti/ → התאבד /hitʔa**b**ed/), but may be fricatives in Paal verbs (e.g., אבדתי /a**v**adti/ → התאבד /hitʔa**b**ed/). Because this change is not salient in Modern Hebrew (Adam, 2002), is restricted to these three letters, and is not accompanied by any orthographic changes, it is unlikely to play a role in morphological priming effects. Nevertheless, an additional analysis was carried out, in which we excluded the four items for which the second root consonant could be considered phonologically more similar between Piel primes and Hitpael targets (3 items from the 1sg past condition and 1 from the infinitive condition). This analysis produced exactly the same pattern of results—that is, a significant priming effect for Piel ($b = -0.0671$, $t = -3.20$), but not for Paal primes ($b = -0.0180$, $t = -0.85$), as well as greater priming for Piel than for Paal ($b = -0.0491$, $t = -2.34$). Furthermore, even if this alternation of the second root consonant was conceived of as an abstract underlying phonological feature that is present in all Piel and Hitpael verbs (rather than in only those with *K*, *B*, and *P* in the second root position), it is hard to see how this would explain the full pattern of our results. This additional phonological feature would amount to a very small difference in overlap, especially when compared to the large orthographic and morphological overlap that exists for *both* Paal and Piel primes. Nevertheless, despite the large overlap between Paal and Hitpael forms, a clear dissociation between binyanim was obtained, such that only Piel–but not Paal–forms elicited a robust root priming effect.

In contrast to distributed accounts, which invoke the convergence of orthographic and semantic codes to explain morphological priming effects, decompositional accounts propose that word recognition in Semitic languages is achieved via rapid decomposition into morphological constituents, such that the consonants of the root are the targets of lexical search (Boudelaa & Marslen-Wilson, 2015; Velan & Frost, 2011). These proposals also fail to explain the full pattern of results in the present study, because they predict that root extraction underlies the recognition of all structured forms in Hebrew (Deutsch et al., 1998). In contrast, our results demonstrate that morpho-lexical representation in Hebrew includes entries (in our case, Paal verb stems) for which lexical access does not involve down-to-the-root parsing. While it is true that simple "non-Semitic" Hebrew words (Velan & Frost, 2011) and irregularly inflected nouns (Vaknin-Nussbaum & Shimron, 2011) are also thought to be accessed via their stems or full forms (rather than by decomposition), our study is the first to identify a word class that displays a prototypical root and pattern Semitic structure, but fails to produce the familiar root priming effect.

Common to both morphological (rule-based) and non-morphological (distributed) accounts of the Semitic mental lexicon is the notion that the pervasive internal structure of Semitic stems shapes the language processing system (Bick et al., 2011; Boudelaa & Marslen-Wilson, 2011, 2015; Plaut & Gonnerman, 2000). The results from the current study suggest that such proposals might benefit from qualification. Specifically, our study has revealed similarities between word recognition in Semitic and in Indo-European languages, with respect to the distinction between lexical storage and grammatical computation. The striking difference between Paal and Piel verb classes is that despite their comparable frequencies, the Paal binyan constitutes a closed class, whereas the Piel binyan is an open class that can be extended to new verbs. One straightforward way of explaining how distinct priming effects coincide with productivity differences is by postulating that Piel, but not Paal, stems are rule based. Rules are operations over variables (Marcus, 2001)—that is, they readily apply to whole categories (e.g., "verbal root"). At the same time, rules are combinatorial operations, which means that they can be employed to decompose stems into morphological constituents. If the processing of Paal verbs is not mediated by a "Paal rule", then their recognition will necessarily depend on access to whole (undecomposed) stems. As such, under this perspective, the "Semitic mental lexicon" is not fully decompositional but, instead, shows a division of labour between structured and undecomposed stems. This distinction between stored entries and combinatorial operations is also present in the processing of complex forms and stems in a range of Indo-European languages (e.g., Sonnenstuhl et al., 1999; Veríssimo & Clahsen, 2009) and has been argued to be a fundamental feature of language (e.g., Clahsen, 1999; Pinker, 1999; Pinker & Ullman, 2002).

It is true that the morphology of Semitic languages has “special characteristics” (Frost, 2006, p. 440), in that morphological units (roots and patterns) surface in a discontinuous way. This means that rules of stem formation must be able to manipulate representations that can be non-linearly combined—for example, verbal patterns that contain “open slots” for root consonants. In addition, mechanisms of lexical access in Hebrew need to be flexible enough to extract constituents that are interleaved, rather than being dependent on the concatenation of surface strings or on the identification of stand-alone semantic units. Nonetheless, our results suggest that when it comes to the fundamental mechanisms that morphology depends on, Hebrew complex words show the same division of labour between storage and computation as that seen in many other languages. We conclude that there is no “Hebrew brain” or “English brain” (Bick et al., 2011, p. 2280). Rather, in both languages, the processing system reflects the dual nature of the language faculty and makes use of the same universal architecture.

Acknowledgements

We thank Alon Itai for assisting with the design of the materials and Aya Meltzer-Asscher for help with the recruitment of participants. We also thank Shmuel Bolozky, Dorit Ravid, Harald Baayen, Sachiko Kinoshita, and an anonymous reviewer for helpful comments on a previous version of this manuscript.

Appendices

Appendix 1

Experimental items employed in the present study for the 1sg Past and Infinitive condition

1SG PAST CONDITION

Unrelated	Paal	Piel	Target
משכתי mSkty /maʃaxti/ pulled	אבדתי ʔbdty /avadti/ was lost	איבדתי ʔybdty /ibadti/ lost	התאבד htʔbd /hitʔabed/ committed suicide
ישבתי ySbty /yaʃavti/ sat	בצעתי bcʕty /batsaʕti/ ripped	ביצעתי bycʕty /bitsaʕati/ performed	התבצע htbcʕ /hitbatseʕ/ was performed
אימנתי ʔymnty /imanti/ trained	דמיתי dmyty /damiti/ was similar to	דימיתי dymyty /dimiti/ simulated	הידמה hydmh /hidama/ became similar
דיברתי dybrty /dibarti/ spoke	הפכתי hpkty /hafaxti/ turned/ became	היפכתי hypxty /hipaxti/ changed side	התהפך hthpk /hithapex/ turned over one's self
בהיתי bhyty /bahiti/ stared	חבקתי xbqty /xavakti/ embraced	חיבקתי xybqty /xibakti/ hugged	התחבק htxbq /hitxabek/ hugged with someone
עיכבתי ʕykbty /ʕikavti/ hindered	חנכתי xnkty /xanaxti/ inaugurated	חינכתי xynkty /xinaxti/ educated	התחנך htxnk /hitxanex/ was educated
חזרתי xzrty /xazarti/ returned	ידעתי ydʕty /yadaʕti/ knew	יידעתי yydʕty /yideʕti/ notified	התוודע htwwdʕ /hitvadeʕ/ was introduced
דרסתי drsty /darasti/ ran over	יעצתי ycʕty /yaʕatsti/ advised	ייעצתי yycʕty /yiʕatsti/ advised	התייעץ htyyʕc /hityaʕets/ consulted with someone
בישלתי bySlty	לוויתי lwwyty	ליוויתי lywwyty	התלווה htlwwh

/biʃalti/	/laviti/	/liviti/	/hitlava/
cooked	borrowed	accompanied	joined
שכחתי	מניתי	מיניתי	התמנה
Skxty	mnyty	mynyty	htmnh
/ʃaxaxti/	/maniti/	/miniti/	/hitmana/
forgot	counted	appointed	was appointed
שרטתי	נגחתי	ניגחתי	התנגח
SrTty	ngxty	nygxty	htngx
/saratti/	/nagaxti/	/nigaxti/	/hitnagex/
scratched	butted	bashed	clashed
טיפסתי	נשקתי	נישקתי	התנשק
Typsty	nSqty	nySqty	htnSq
/tipasti/	/naʃakti/	/niʃakti/	/hitnaʃek/
climbed	kissed	kissed	kissed with someone
אישרתי	עניתי	עיניתי	התענה
ʔySrty	ʕnyty	ʕynyty	htʕnh
/iʃarti/	/ʕaniti/	/ʕiniti/	/hitʕana/
approved	answered	tortured	was tortured
ניצחתי	פניתי	פיניתי	התפנה
nycxty	pnyty	pynyty	htpnh
/nitsaxti/	/paniti/	/piniti/	/hitpana/
won	turned to	cleared	was evacuated/ had time
זרקתי	פקדתי	פיקדתי	התפקד
zrqty	pqdty	pyqdty	htpkd
/zarakti/	/pakadti/	/pikadti/	/hitpaked/
threw	ordered	commanded	officially joined
שתלתי	פקחתי	פיקחתי	התפקח
Stlty	pqxty	pyqxty	htpqx
/ʃatalti/	/pakaxti/	/pikaxti/	/hitpakex/
planted	opened	supervised	became clever
הידקתי	קלפתי	קילפתי	התקלף
hydqty	qlpty	qylpty	htqlp
/hidakti/	/kalafti/	/kilafti/	/hitkalef/
fastened	peeled	peeled	was peeled
סחטתי	רגשתי	ריגשתי	התרגש
sxTty	rgSty	rygSty	htrgS
/saxatti/	/ragaʃti/	/rigaʃti/	/hitrageʃ/
squeezed	was not calm	moved/excited	was excited/moved
ביקרתי	רציתי	ריציתי	התרצה
byqrty	rcyty	rycyty	htrch
/bikarti/	/ratsiti/	/ritsiti/	/hitratsa/
visited	wanted	served/pleased	agreed
עיבדתי	שמרתי	שימרתי	השתמר
ʕybdty	Smrty	Symrty	hStmr

/ʕibadti/ processed	/ʃamarti/ guarded/saved	/ʃimarti/ preserved	/hiʃtamer/ was preserved
הימרתי hymrty /himarti/ gambled	שתקתי Stqty /ʃatakti/ was silent	שיתקתי Sytqty /ʃitakti/ paralyzed	השתתק hSttq /hiʃtatek/ became silent

INFINITIVE CONDITION

Unrelated	Paal	Piel	Target
לטרוף ltrwp /litrof/ to devour	לבקוע lbkwʕ /livkoʕ/ to be hatched	לבקע lbkʕ /levakeʕ/ to cleave	התבקע htbkʕ /hitbakeʕ/ cracked
לנגן lngn /lenagen/ to play (music)	לברור lbrwr /livror/ to select	לברר lbrr /levarer/ to check	התברר htbrr /hitbarer/ turned out
להגות lhgwt /lahgot/ to pronounce	לחבור lxbwr /laxvor/ to join	לחבר lxbr /lexaber/ to connect/ compose	התחבר htxbr /hitxaber/ connected one's self
לשפר lSpr /leSaper/ to improve	לחלוק lxlwq /laxlok/ to share	לחלק lxlq /lexalek/ to divide	התחלק htxlq /hitxalek/ shared/ was divided
לזהם lzhm /lezahem/ to pollute	לחפוף lxpwp /laxfof/ to overlap	לחפף lxpp /lexafef/ to do a sloppy job	התחפף htxpp /hitxafef/ took off
לסיים lsyym /lesayem/ to finish	לחשוב lxSwb /laxSov/ to think	לחשב lxSb /lexaSev/ to calculate	התחשב htxSb /hitxaSev/ considered
לסנן lsnn /lesanen/ to filter	ללכוד llkwd /lilkod/ to capture	ללכד llkd /lelaked/ to unify	התלכד htlkd /hitlaked/ was unified
לבחור lbxwr /livxor/ to choose	ללמוד llmwd /lilmod/ to learn	ללמד llmd /lelamed/ to teach	התלמד htlmd /hitlamed/ interned
לאשש lʔSS	לספוח lspwx	לספח lspx	הסתפח hstpx

/leʔaSeS/	/lispox/	/lesapex/	/histapex/
to confirm	to absorb	to annex	was annexed
לגזול	לסרוק	לסרק	הסתרק
lgzwl	lsrwq	lsrq	hstrq
/ligzol/	/lisrok/	/lesarek/	/histarek/
to steal	to scan	to comb	combed one's hair
לאכול	לפטור	לפטר	התפטר
lʔkwl	lpTwr	lpTr	htpTr
/leʔexol/	/liftor/	/lefater/	/hitpater/
to eat	to exempt	to fire	resigned
לחזק	לפעום	לפעם	התפעם
lxzq	lpʕwm	lpʕm	htpʕm
/lexazek/	/lifʕom/	/lefaʕem/	/hitpaʕem/
to strengthen	to beat	to beat	was amazed
לרכוב	לפרוק	לפרק	התפרק
lrkwb	lprwq	lprq	htprq
/lirkov/	/lifrok/	/lefarek/	/hitparek/
to ride	to unload	to disassemble	was disassembled
לבלוט	לפרוש	לפרש	התפרש
lblwT	lprwS	lprS	htprS
/livlot/	/lifroʃ/	/lefareʃ/	/hitpareʃ/
to stand out	to retire	to interpret	was interpreted
לחתן	לפשוט	לפשט	התפשט
lxtn	lpSwT	lpST	htpST
/lexaten/	/lifʃot/	/lefaʃet/	/hitpaʃet/
to give in marriage	to raid	to simplify	spread
לספק	לקבוע	לקבע	התקבע
lspq	lqbwʕ	lqbʕ	htqbʕ
/lesapek/	/likboʕ/	/lekabeʕ/	/hitkabeʕ/
to provide	to determine	to fixate	became fixed
לגדול	לקשור	לקשר	התקשר
lgdwl	lqSwr	lqSr	htqSr
/ligdol/	/likʃor/	/lekaʃer/	/hitkaʃer/
to grow	to tie	to connect	called
לפלוש	לשטוח	לשטח	השתטח
lplwS	lSTwx	lSTx	hStTx
/lifloʃ/	/liʃtox/	/leʃatex/	/hiʃtatex/
to invade	to lay out	to flatten	prostrated
לנמק	לשכון	לשכן	השתכן
lnmq	lSkwn	lSkn	hStkn
/lenamek/	/liʃkon/	/leʃaken/	/hiʃtaken/
to explain	to reside	to house	settled
לסתום	לשלוח	לשלח	השתלח
lstwm	lSlwx	lSlx	hStlx

/listom/	/liʃlox/	/leʃalex/	/hiʃtalex/
to shut	to send	to dismiss	insulted
לחמוק	לשקוע	לשקע	השתקע
lxmwq	lSqwʕ	lSqʕ	hStqʕ
/laxmok/	/liʃkoʕ/	/leʃakeʕ/	/hiʃtakeʕ/
to escape	to sink	to cause to sink	settled

4 Publication II

Published in *Bilingualism: Language and Cognition*, 2018, doi: 10.1017/S1366728918000032

Do Late Bilinguals Access Pure Morphology During Word Recognition? A Masked-Priming Study on Hebrew as a Second Language

Yael Farhy, João Veríssimo and Harald Clahsen

Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam

Abstract

This study extends research on morphological processing in late bilinguals to a rarely examined language type, Semitic, by reporting results from a masked-priming experiment with 58 non-native, advanced, second-language (L2) speakers of Hebrew in comparison with native (L1) speakers. We took advantage of a case of 'pure morphology' in Hebrew, the so-called binyanim, which represent (essentially arbitrary) morphological classes for verbs. Our results revealed a non-native priming pattern for the L2 group, with root-priming effects restricted to non-finite prime words irrespective of binyanim type. We conclude that root extraction in L2 Hebrew word recognition is less sensitive to both morphological and morphosyntactic cues than in the L1, in line with the Shallow-Structure Hypothesis of L2 processing.

Introduction

Do non-native, late bilinguals make use of morphological and morphosyntactic information during online word recognition in the same way as native speakers? Although this question has received a lot of attention in recent experimental research, the matter is still controversial. Some researchers have claimed that native (L1) and non-native (L2) speakers apply the same mechanisms for processing morphologically complex words, but that L2 processing may be negatively affected by difficulties with lexical access or retrieval, working memory limitations, and/or slower processing speed (e.g., Cunnings, 2017; Hopp, 2016; McDonald, 2006). Alternatively, more substantial L1/L2 differences have been posited by the Shallow-Structure Hypothesis (SSH), originally for sentence processing (Clahsen & Felser, 2006a, 2006b) and later extended to morphological processing (e.g., Clahsen, Felser, Neubauer, Sato & Silva, 2010; Clahsen, Balkhair, Schutter & Cunnings, 2013; Clahsen, Gerth, Heyer & Schott, 2015). The SSH holds that even proficient L2 speakers tend to have problems building or manipulating abstract grammatical representations in real time, and that relative to native speakers, L2 processing of morphologically complex words relies more heavily on storage of complex forms and less on morphological structure and computation. In experimental research, priming experiments (specifically, masked priming) have been widely used as a technique to provide insight into the kinds of cues the word recognition system relies on during morphological processing (see Marslen-Wilson, 2007, for a review). Previous L2 priming studies, however, have produced mixed results. For inflection, for example, some studies reported L1-like morphological priming effects for English past-tense forms and for French -*er* verbs in groups of late bilinguals, even under masked-priming conditions (e.g., Coughlin & Tremblay, 2015; Feldman, Kostić, Basnight-Brown, Đurđević & Pastizzo, 2010; Voga, Anastassiadis-Symeonidis & Giraudo, 2014), whereas other studies found morphological facilitation effects for inflectional phenomena in English, German, and Turkish for their L1 control groups, but not for groups of (highly proficient) late bilinguals (e.g., Jacob, Heyer & Veríssimo, 2018; Kirkici & Clahsen, 2013; Neubauer & Clahsen, 2009; Silva & Clahsen, 2008). The question of whether an inflected word's morphological structure and its morphosyntactic features (e.g., finiteness features) are underused in L2 processing or whether the L2 system employs these information sources in the same way as the L1 system remains to be answered.

Against this background, this study reports the results from a masked-priming experiment on late bilinguals’ processing of Hebrew inflectional morphology, the first L2 study of its kind on a previously unconsidered language type (Semitic). Due to its non-concatenative properties, the Semitic lexical processor has been claimed to be primarily driven by morphology, designed to extract a complex word's abstract structure (root and word pattern), irrespective of meaning or surface form (Boudelaa & Marslen-Wilson, 2015; Frost, Forster & Deutsch, 1997). With those properties, Hebrew morphology should be an ideal test case to test the role of morphology during L2 word recognition.

Background: Hebrew morphology

There is an extensive linguistic and psycholinguistic literature on this topic (see, e.g., Arad, 2005; Aronoff, 1994; Frost et al., 1997; Plaut & Gonnermann, 2000), which will not be discussed here. Rather, the following remarks are meant as background information for those unfamiliar with Hebrew morphology and the corresponding experimental research.

Hebrew composes most words non-concatenatively, by a non-linear combination of a consonantal root and a vowel pattern, which together constitutes stems to which inflectional affixes may be added. The consonantal root typically contains three consonants and carries the core meaning of the word. The same root is commonly assigned to more than one pattern. For verbs there are seven distinct vowel patterns called *binyanim*, which provide designated positions for inserting a root's consonants. For example, the verb *katav* '(he) wrote' consists of the root *K-T-V* and the binyan Paal, which is expressed by the vowel pattern *CaCaC* (with 'C' representing the root consonants). While two binyanim (Pual and Hufal) are fully predictable in their morphosyntactic function (encoding passive voice), the other five classes show only general semantic and syntactic tendencies (for review, see Arad, 2005). In fact, for many Hebrew verbs, it is generally difficult or impossible to compute their meaning compositionally on the basis of roots and patterns. Furthermore, binyanim assignment is obligatory, with every verb having to be assigned to a binyan before it can be inflected. Aronoff (1994, Chapter 5) points out that with these properties the Hebrew binyanim may be conceived as a system of inflectional classes, akin to the conjugational classes in the Romance languages. As such, the binyanim are not morphemic – in the sense of reliably encoding particular syntactic or semantic properties – but instead serve as an abstract morphological mark of a given verb's inflectional class, which dictates the phonological shape of its different forms (Aronoff, 1994, p. 127).

For the present study, we examined the two most common binyanim, Paal and Piel. While both classes have relatively high type frequencies (calculated as percentages of verbs of each class in a corpus containing 4,131 verbs; Itai & Wintner, 2008), they differ in their productivity (see Table 1). New verbs are typically assigned to Piel, whereas Paal forms represent a restricted class of lexical items in Modern Hebrew (see Bolozky, 1999a).

Table 1

Properties of the Paal and Piel classes

Class	Phonological base form	Example	Type frequency
Paal	*CaCaC*	*lamad* 'learned'	19.4 %
Piel	*CiCeC*	*limed* 'taught'	17.1 %

Hebrew (and Arabic) morphology has been subject to a large number of experimental studies, albeit almost exclusively with native speakers (Boudelaa & Marslen-Wilson, 2005; Deutsch, Frost & Forster, 1998; Frost et al., 1997; Frost, Kugler, Deutsch & Forster, 2005). One familiar finding from this research is a ROOT-PRIMING EFFECT, that is, faster lexical decision times for a target word (e.g., *mixtav* 'a letter') when the prime shares the same consonantal root, relative to an unrelated control prime (e.g., *hixtiv* '(he) dictated' → *mixtav* vs. *hirgish* '(he) felt' → *mixtav*). Interestingly, a root-priming effect was obtained even when prime and target were not transparently related in meaning (e.g., *rasham* '(he) registered' → *hirshim* '(he) impressed'). How to interpret root-priming effects is controversial. Root priming may signal full morphological decomposition of every complex Hebrew or Arabic word form into root and pattern (e.g., Boudelaa & Marslen-Wilson, 2011, 2015). Alternatively, root-priming effects have been interpreted in terms of additive or interactive effects of prime-target overlap in form and meaning, even under masked-priming conditions (e.g., Plaut & Gonnerman, 2000; see also Gonnerman, Seidenberg & Andersen, 2007).

Further insight into the nature of morphological priming in Hebrew comes from a recent masked-priming study that showed reliable root-priming effects from Piel primes, but not from Paal primes, with a group of L1 Hebrew speakers (Farhy, Veríssimo & Clahsen, 2018a). The authors attributed this contrast to distinct morpho-lexical representations for the two binyanim, structured fully decomposable stems for productive classes (like Piel) versus unstructured stems that are stored as wholes for unproductive classes (like Paal).

Little is known about late bilinguals' processing of Semitic morphology. Two recent studies provide insight into how L2 speakers handle its non-concatenative morphology. Norman, Degani and Peleg (2016) found that in L1 and L2 Hebrew speakers (the latter with either Arabic or English as L1), pseudowords composed from existing roots and vowel patterns yielded longer lexical-decision times than pseudowords composed from non-existing roots or vowel patterns. The authors interpreted this finding as signalling native-like decomposition of Hebrew words into roots and patterns in L2 Hebrew. Alternatively, however, it is possible that pseudowords with existing roots and vowel patterns take longer to reject (relative to pseudowords constructed from non-existing component parts), because they are phonologically more similar to real words. In addition, in a cross-modal study (Freynik, Gor & O'Rourke, 2017), root-priming effects of similar magnitudes were obtained in both L2 and L1 speakers of Arabic, which the authors interpret as signalling the use of native-like morphological decomposition into roots and patterns in L2 Arabic speakers. Alternatively, however, one may conceive of this finding as the result of the combined effects of overlap of form and meaning, given that cross-modal priming may be particularly sensitive to semantic effects (e.g., Gonnerman et al., 2007; Marslen-Wilson, Tyler, Waksler & Older, 1994).

In the present study, we examine whether late bilinguals access pure morphology during word recognition in the same way as native speakers, by testing late bilinguals' sensitivity to the Hebrew binyanim – a case of pure morphology that is less directly derivable from form-meaning associations than the commonly studied grammatical morphemes. Specifically, we tested whether non-native speakers showed root priming from verbs belonging to the productive Piel class and from the unproductive Paal class, presented in two different verbal forms (1sg Past and Infinitive). In addition, we directly compared our results to a group of native speakers who underwent the same experiment (Farhy et al., 2018a) and showed robust priming from Piel, but not from Paal forms. If non-native speakers are sensitive to the distinction between verbal classes, they should show the same priming pattern as the L1 control group. However, if L2 speakers underuse morphological information, they should show similar effects following Paal and Piel primes.

Method

Participants

Fifty-eight non-native Hebrew speakers (35 males, 7 left-handed, mean age: 28.69, SD: 4.46) participated in the experiment. They all lived in Israel and had emigrated from South America, with Spanish (n = 55) or Portuguese (n = 3) as their native language. All participants were late learners of Hebrew (mean age of onset: 14.0 years, SD: 4.61) and achieved a mean score of 96% (SD = 0.08) in one section of the YAEL proficiency test for university candidates (including sentence completion, sentence rephrasing, and reading comprehension), with every participant achieving a 70% score or higher. Participants also estimated (in percentages) their relative use of Hebrew, both overall, and in four specific domains (speaking, hearing, writing, and reading). Hebrew had a mean overall usage of 60.96% (SD = 18.94). Similar usage ratings were reported for the four specific domains (speaking: 61.56%; hearing: 61.84%; writing: 67.39%; reading: 63.18%). Participants used their native languages (Spanish or Portuguese) less often, with a mean overall usage of 28.95% (SD: 16.73). All participants had normal or corrected-to-normal vision and none had been diagnosed with any language disorders.

Materials

Table 2 displays the experimental design, including an example stimulus set. Experimental targets consisted of 42 Hitpael verbs, presented in the 3sg past tense, a form constituted by a root together with the Hitpael verbal pattern, with no inflectional suffixes. There were two sets of 21 targets, one for each of two Form Type conditions, 1sg Past and Infinitive; as described below, Form Type refers to the particular form in which prime words were presented. Note that the experimental targets were always presented in their 3sg past-tense form. Each target word was paired with three types of primes: (a) one prime belonging to the Paal class, based on the same verbal root as the target, (b) one prime belonging to the Piel class, also based on the same root, and (c) one unrelated prime, which had no orthographic, phonological, or semantic overlap with its corresponding target form. The materials were the same as in Farhy et al. (2018a).

Each target was preceded by Paal, Piel, and Unrelated primes presented either in a finite form, the 1sg past tense (21 targets), or in the infinitive form (21 targets)[6]. Conditions were matched for length, semantic relatedness, and lemma frequency (Itai & Wintner, 2008). We also included 294 filler prime–target pairs, 126 word–word pairs and 168 word–nonword pairs, yielding a total of 336 targets, half of them words and half pseudowords (for additional descriptions, see Farhy et al., 2018a).

Procedure

Participants were asked to perform a lexical decision task on visual targets, as quickly and accurately as possible. Every trial consisted of a fixation cross (500ms), a blank screen (500ms), a row of hash marks (500ms), a prime word (50ms), and the target (presented until a response was made, up to a timeout of 2000ms). After a further 500ms, the next trial started. Response times (RTs) were measured from the onset of target presentation.

[6] The same targets could not be preceded by both forms because several primes could only appear as 1sg past-tense forms due to Paal/Piel homography in the infinitive.

Table 2

Experimental conditions, with an example stimulus set

Form type	Target (Hitpael)	Prime type: Unrelated	Prime type: Paal	Prime type: Piel
1sg Past	התאבד htʔbd /hitʔabed/ ‘committed suicide’	משכתי mSkty /maʃaxti/ ‘pulled’	אבדתי ʔbdty /avadti/ ‘was lost’	איבדתי ʔybdty /ibadti/ ‘lost’
Infinitive	התחלק htxlq /hitxalek/ ‘(was) shared/divided’	לשפר lSpr /leʃaper/ ‘to improve’	לחלוק lxlwq /laxlok/ ‘to share’	לחלק lxlq /lexalek/ ‘to divide’

Note: Examples are shown in Hebrew orthography together with transliteration to Latin letters, phonological form, and English translation.

Data analysis

Two items with extremely low accuracy (below 50%) were removed (*hidama* ‘was similar’ and *hishtamer* ‘was preserved’, both from the 1sg Past condition), in addition to incorrect responses and timeouts (i.e., when no response was made during the 2000ms window). In addition, trials with extremely slow RTs (longer than 1,500ms) were removed (6.84%). In order to compare the current priming effects to an L1 group, Farhy et al.’s (2018a) data from 30 native Hebrew speakers (who underwent the same task and procedure) were added to the present dataset.

The RT data were analysed with generalised linear mixed-effects regression, with crossed random effects for participants and items (Baayen, Davidson & Bates, 2008). RTs were analysed without any transformation, but with the assumption that the data follows an inverse Gaussian distribution (with an identity link function), as recommended by Lo and Andrews (2015). A detailed description of this kind of regression model is provided in Appendix S1 (Supplementary material), together with an assessment of the models’ random structure. The following factorial predictors were included: (a) Prime Type (Paal, Piel, Unrelated), (b) Form Type (1sg Past, Infinitive), and (c) Group (L1, L2). In the presence of interactions, treatment contrasts were used and the statistical comparisons of interest were obtained by relevelling factors and refitting the model. In the cases where ‘main effects’ are reported, these were obtained by converting factors to numeric variables and centering them (e.g., Fraundorf & Jaeger, 2016). Accuracy data were also analysed with generalised mixed-effects regression (binomial family, logit link function; Jaeger, 2008).

Results

Table 3 displays mean RTs, standard errors (SEs), and accuracy rates in all conditions, for the L2 speaker group in the present study, as well as for the L1 control group.

Table 3

Mean RTs and SEs (in parenthesis) and accuracy rates

Group	Form type	Unrelated		Paal		Piel	
		RT (ms)	Acc.	RT (ms)	Acc.	RT (ms)	Acc.
L1	1sg Past	634 (10.94)	94%	629 (10.34)	97%	609 (9.46)	97%
	Infinitive	641 (10.08)	94%	639 (10.91)	94%	620 (10.81)	97%
L2	1sg Past	807 (13.74)	85%	820 (15.68)	86%	818 (14.72)	88%
	Infinitive	819 (13.77)	80%	790 (14.90)	79%	794 (13.55)	82%

Note: The L1 data are from Farhy et al. (2018a). Displayed means and SEs are back-transformed from a reciprocal transformation (-1000/RT), as this was the transformation used in the analysis of the L1 group reported in Farhy et al.

In the accuracy data, a main effect of Group was obtained (b = 1.24, z = 3.88, p<.001), indicating higher accuracy rates for the L1 group, across conditions, and an effect of Prime Type, indicating higher accuracy rates in the Piel prime condition, across groups (Unrelated vs. Piel: b = 0.45, z = 2.74, p = .006; Paal vs. Piel: b = 0.34, z = 2.00, p = .045). Significant interactions between the predictors were not found (all |z|s<1.13, all ps>.262).

With regard to the RTs, the analysis yielded significant three-way interactions between Prime Type, Form Type and Group, both for Paal priming (b = −36.3, t = −2.47, p = .013) and Piel priming (b = −34.7, t = −2.38, p = .017). These analyses were followed by separate within-group analyses. For the L2 group, interactions between Prime Type and Form Type were significant, for both Paal priming (b = −35.4, t = −2.34, p = .019) and Piel priming (b = −37.1, t = −2.24, p = .025), but not for the comparison between Paal and Piel (b = 1.7, t = 0.11, p = .913), indicating that the two binyanim show similar effects across the two forms. Further analyses revealed different priming patterns in the Infinitive and the 1sg Past conditions.

In the Infinitive condition, Piel and Paal primes yielded significantly shorter RTs than Unrelated primes (Unrelated vs. Piel: $b = 26.3$, $t = 2.15$, $p = .032$; Unrelated vs. Paal: $b = 22.0$, $t = 2.01$, $p = .045$). By contrast, primes presented in the 1sg Past condition did not produce any facilitation, neither for Piel ($b = -10.78$, $t = -0.93$, $p = .350$), nor for Paal primes ($b = -13.40$, $t = -1.09$, $p = .275$)[7].

In the L1 group, similar priming effects were obtained in the Infinitive and the 1sg Past conditions, with no interactions between Prime Type and Form Type, neither for Paal ($b = 2.1$, $t = 0.13$, p = .895), nor for Piel priming ($b = -2.5$, $t = -0.15$, $p = .880$). Across both Form Type conditions, Piel primes yielded significantly shorter RTs than Unrelated primes ($b = 23.9$, $t = 2.93$, $p = .003$), whereas Paal primes did not produce any reliable facilitation ($b = 2.07$, $t = 0.25$, $p = .805$), with a significant difference between RTs following Paal and Piel primes ($b = -21.8$, $t = -2.68$, $p = .007$). In addition, an examination of each Form Type condition (despite the absence of interactions) showed significant priming from Piel in both the Infinitive and 1sg Past (both *p*s<.049), but no priming effects from Paal in neither of the two Form Types (both *p*s>.786).

Discussion

In the current masked-priming study, late L2 learners of Hebrew showed similar morphological root-priming effects for infinitive forms, regardless of whether verbs belonged to the productive verbal class Piel or to the unproductive class Paal. However, when verbs were presented in a finite form (1sg past tense), no priming effects were obtained, neither for Paal nor for Piel verbs. Both of these findings stand in sharp contrast with the results obtained with the L1 group, for which root-priming effects were elicited by Piel verbs only, both when they were presented as infinitives and as finite forms. In other words, whereas for native speakers priming effects were modulated by binyan but not by finiteness, L2 speakers showed the opposite pattern, with morphological priming being crucially dependent on finiteness but not on binyan membership.

The Hebrew binyanim have been argued to constitute a system of 'pure morphology', because they determine the phonological shape of verbal stems and are essentially arbitrary with regards to the syntactic or semantic properties that they express (Aronoff, 1994). The fact that L1 root-priming effects are modulated by binyan indicates that native speakers distinguish

[7] Following a reviewer's concern that the higher L2 error rate could affect the results, we conducted a combined RT/accuracy analysis and found parallel results (see Appendix S2 in Supplementary material).

between these abstract morphological categories and use this information during visual word recognition, specifically, by decomposing stems of Piel verbs down to the level of the root or accessing whole stems in the case of the unproductive class Paal (see Farhy et al., 2018a, for discussion). However, this account cannot be extended to the L2 data. The equivalent L2 effects for Paal and Piel – priming for infinitives, and no priming for finite forms – suggests instead that late-learners of Hebrew represent verbs from productive and unproductive classes in a similar way and, at least under masked-priming conditions, do not show sensitivity to purely morphological cues like binyan membership during the recognition of complex forms.

How can the L2 root-priming effects for infinitives be explained? One possibility is that L2 speakers of Hebrew recognise and extract root constituents from non-finite forms (Freynik et al., 2017), for example, due to the root's salient role in Hebrew as a structural unit (Frost et al., 1997). Alternatively, the source of these facilitation effects may be that L2 speakers learn form-to-meaning lexical regularities, such as the co-occurrence of roots with certain semantic features (Plaut & Gonnerman, 2000). Crucially, however, the process by which roots are pre-activated in L2 masked priming is (a) equally applicable to verbs from all binyanim and (b) can be 'blocked' or made more difficult in the case of forms that contain a tense and agreement morpheme, as revealed by the lack of root priming from forms presented in the 1sg past tense. Although not initially hypothesized, the contrast between facilitated decomposition of infinitives, but lack of priming from forms with finite morphemes, is reminiscent of production studies in which late learners show considerable difficulties with the overt expression of morphosyntactic features, instead producing infinitives and unmarked forms (e.g., Blom, Polišenská & Weerman, 2006; White, 2003). Prévost and White (2000), for example, suggested that (in production) access to finite forms is sometimes blocked, resulting in underspecified non-finite forms 'winning' the competition for lexical insertion, especially under pressure. Furthermore, the lack of priming from 1sg past-tense forms in our study is broadly consistent with a number of previous priming studies, in which late bilinguals show reduced facilitation from inflected forms and marked stems (e.g., Jacob et al., 2018; Veríssimo, Heyer, Jacob & Clahsen, 2018; Krause, Bosch & Clahsen, 2015; Silva & Clahsen, 2008). The present findings therefore suggest that word recognition in advanced L2 learners is less sensitive to B O T H morphological (binyan membership) and morphosyntactic cues (finiteness) than in native speakers.

Other proposed sources of L1 versus L2 contrasts, such as limited exposure and proficiency or native-language influence cannot easily account for our results. Firstly, our L2 participants were all highly proficient and fully immersed speakers, who used Hebrew on a daily basis. Secondly, they were native speakers of Spanish or Portuguese, languages that also display verbal conjugation classes. Moreover, there is evidence that L1 speakers of Romance languages employ this purely morphological information during word recognition (Say & Clahsen, 2002; Veríssimo & Clahsen, 2009).

We conclude that even advanced late-learners are likely to show processing differences relative to L1 speakers, specifically in the domains of pure morphology and morphosyntax. More generally, our results are consistent with the Shallow-Structure Hypothesis (Clahsen & Felser, 2006a, 2006b; Clahsen et al., 2010), according to which late bilinguals underuse grammatical information and analysis in the course of linguistic processing.

Acknowledgements

This work was supported by a Ph.D. scholarship awarded to YF by the Minerva Fellowship Program and an Alexander-von-Humboldt Professorship awarded to HC. We are grateful to Natalie Uziel and Ella Avner for their help with testing and recruiting participants. We also thank two anonymous reviewers for helpful comments on a previous version of this manuscript.

Appendices

Appendix S1

The following text provides a description of (i) the statistical models that were employed (viz., generalised linear mixed-effects regression on untransformed response times), (ii) the assessment of the models' random structure, and (iii) the software versions used for data analysis, as well as the specific arguments to the function call.

1. Analysis of untransformed RTs with generalised linear models

As recommended by Lo and Andrews (2015), in the current paper we analysed the response time (RT) data with generalised linear mixed-effects regression, by including in the models the assumption that the data followed an 'RT-like' distribution—specifically, an inverse Gaussian.

As RT distributions are heavily skewed, it is common in experimental psycholinguistics to analyse the logarithm or the reciprocal of RTs, so that assumptions regarding the normality of residuals can be satisfied. This was also the approach taken in Farhy, Veríssimo, and Clahsen

(2018a), in which we analysed the reciprocal of RTs (-1000/RT), rather than the actual times produced in the experiment. However, concerns have been raised about the analysis of transformed RTs being associated with serious problems (Balota, Aschenbrenner, & Yap, 2013; Lo & Andrews, 2015; O'Malley & Besner, 2013). In particular, Balota et al. have demonstrated that nonlinear transformations (such as the reciprocal) may give rise to spurious interactions by distorting purely additive relationships. A solution, recently proposed by Lo and Andrews, is to employ generalised linear mixed-effects regression. In this approach, raw (untransformed) RTs are directly analysed, but at the same time, it is possible to include the assumption that the data follows a skewed distribution (see Lo & Andrews, for further details; for recent examples of this type of analysis, see, e.g., Masson, Rabe, & Kliegl, 2017; Medeiros & Duñabeitia, 2016). An important additional benefit of this approach is that effects can be readily interpreted in their true scale, that is, every estimate is expressed as a difference in milliseconds.

2. Random structure of the statistical model

In the current study, the data were analysed with regression models with crossed random effects for participants and items (Baayen, Davidson, & Bates, 2008). As recommended by Matuschek, Kliegl, Vasishth, Baayen, and Bates (2017), random slopes for the different predictors were tested for inclusion on the basis of the models' AIC, a measure of goodness of fit. Against a simple, intercept- only between-group regression model (with categorical fixed effects Prime Type, Form Type, and Group), we tested all possible random slopes individually and obtained the AIC of the resulting models. A random by-item slope for Group (L1, L2) improved fit the most (i.e., led to the lowest AIC), for both the RT model and the accuracy model. Further inclusion of additional random slopes led to models that did not converge or did not improve fit. Additional follow-up analyses (within each group) were conducted with intercept-only models, as Group was not a predictor in these models.

3. Software versions and function call

Data were analysed using the *lme4* package (version 1.1-12; Bates, Maechler, Bolker, & Walker, 2015) for the R language (version 3.4.1). Specifically, in the case of the RT analysis, models were fit using the following function call:

```
glmer(..., family=inverse.gaussian(link="identity"),
glmerControl(optimizer="bobyqa", optCtrl=list(maxfun=30000)))
```

In the case of the accuracy analysis, the following function call was used:

```
glmer(..., family=binomial, glmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=30000)))
```

Appendix S2

The following text reports an analysis on a combined RT/accuracy rate measure.

1. Analysis of combined RT/accuracy measure

In order to ensure that the higher error rates in the L2 group (a contrast commonly found in studies comparing L1 and L2 speakers) did not affect the pattern of priming effects that we obtained, a reviewer proposed a combined RT and accuracy analysis (see, e.g., Bruyer & Brysbaert, 2011).

Here, we report an analysis of L2 responses using a combined by-participant RT/accuracy adjustment. In order to calculate this measure, each RT was divided by the accuracy rate of the participant who provided the response. Because the adjustment produced a very skewed distribution (even more so than a normal RT distribution), the data was reciprocally transformed (i.e., -1000/(RT/accuracy)). The (back-transformed) means of the combined measure in each condition are shown in Table S1 below.

Table S1

Means (back-transformed) of the combined RT/accuracy measure

	Unrelated	Paal	Piel
1sg Past	812	814	806
Infinitive	816	798	792

As the means in Table S1 indicate, we found parallel results to the ones reported in the paper. In the Infinitive condition, RTs following both Paal primes (t = -2.12, p = .035) and Piel primes (t = -2.08, p = .038) were shorter than following Unrelated primes. In the 1sg Past condition, no priming was obtained for either Paal primes (t = 0.66, p = .507) or Piel primes (t = 0.23, p = .822). The interaction of Prime Type and Form Type was significant for Paal priming (t = 1.96, p = .050), but failed to reach statistical significance for Piel priming (t = 1.62, p = .105).

Given the same pattern of priming for both Prime Types (Paal and Piel) in the Infinitive condition, as well as the absence of priming for either Prime Type in the 1sg Past condition, we conclude that the pattern of effects that we report in the main manuscript is robust to participant's error rates.

5 Publication III

Published in *Language and Speech*, 2018, doi: 10.1177/0023830918811863

Semantic Effects in Morphological Priming: The Case of Hebrew Stems

Yael Farhy and João Veríssimo
Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam

Abstract

To what extent is morphological representation in different languages dependent on semantic information? Unlike Indo-European languages, the Semitic mental lexicon has been argued to be purely "morphologically driven", with complex stems represented in a decomposed format (root + vowel pattern) irrespectively of their semantic properties. We have examined this claim by comparing cross-modal root-priming effects elicited by Hebrew verbs of a productive, open-ended class (Piel) and verbs of a closed-class (Paal). Morphological priming effects were obtained for both verb types, but prime-target semantic relatedness interacted with class, and only modulated responses following Paal, but not Piel primes. We explain these results by postulating different types of morpho-lexical representation for the different classes: structured stems, in the case of Piel, and whole-stems (which lack internal morphological structure), in the case of Paal. We conclude that semantic effects in morphological priming are also obtained in Semitic languages, but they are crucially dependent on type of morpho-lexical representation.

Introduction

Morphologically derived words show considerable variation in the way they relate to their bases, in particular with regard to whether the semantic properties of a complex word can be compositionally determined from its constituents. As such, the degree of semantic overlap between morphologically related words can assume a range of values, from highly transparent (e.g., *teacher-teach*) to less transparent (e.g., *cracker-crack*), or even completely idiosyncratic (e.g., *whisker-whisk*, a "pseudo-derivation" that displays the form of a complex word but no overlap in meaning). This raises the question of whether (and at what point) the activation and retrieval of a complex word's constituents depends on semantic transparency.

In experimental psycholinguistics this question has been commonly addressed with experiments on morphological priming, that is, the finding that target words are recognized (or named) faster following the prior presentation of a morphological relative—arguably due to a process of decomposition that leads to the pre-activation of a shared constituent in the target, for example, [*teach+er*]-[*teach*] (for review, see Marslen-Wilson, 2007; Amenta & Crepaldi, 2012). Morphological priming has been reported in both cross-modal priming, in which primes are presented auditorily and targets visually, and in masked priming paradigms, in which primes are visually presented, but for such a brief duration that they are not consciously perceived. In the cross-modal paradigm, morphologically related primes with a semantically transparent relation to the target generally facilitate recognition, whereas those with opaque meanings do not (e.g., English: Gonnerman, Seidenberg, & Andersen, 2007; Marslen-Wilson, Tyler, Waksler, & Older, 1994; French: Longtin, Segui, & Hallé, 2003; but see Smolka, Preller, & Eulitz, 2014, for counterevidence from German). By contrast, in masked priming, numerous studies have reported priming effects for both transparent and opaque prime-target pairs (e.g., Beyersmann, Ziegler, Castles, Coltheart, Kezilas, & Grainger, 2016; Feldman, Soltano, Pastizzo, & Francis, 2004; Kazanina, 2011; Kazanina, Dukova-Zheleva, Geber, Kharlamov, & Tonciulescu, 2008; Lavric, Klapp, & Rastle, 2007; Longtin et al., 2003; Rastle, Davis, Marslen-Wilson, & Tyler, 2000; Rastle, Davis, & New, 2004). The distinct experimental patterns obtained in cross-modal and masked priming have been proposed to arise from different levels of representation and stages of processing within the lexical system (e.g., Marslen-Wilson, 2007; Marslen-Wilson, Bozic, & Randall, 2008; Rastle & Davis, 2003). In this proposal, cross-modal priming taps into representations at the "central lexical level", the modality-independent store of abstract phonological, syntactic and semantic information (Marslen-Wilson et al., 1994). At this level, only truly morphologically complex forms, that is, those closely related in meaning to their bases, are represented with internal constituent structure, for example, [teach+er]; thus, only transparent forms activate their constituents during word recognition (Marslen-Wilson et al., 1994). In contrast, masked visual priming has been argued to tap into an earlier morpho-orthographic level of representation and processing in which all potentially complex forms are decomposed into constituents (e.g., Rastle & Davis, 2003, 2008). According to this "decompositional" view, morpho-orthographic segmentation is blind to semantic properties and based only on the forms of morphemes; thus, priming effects are obtained even when prime words express opaque, idiosyncratic meanings (e.g., *whisker–whisk*; Rastle et al., 2004).

Against these results, other masked priming studies have yielded facilitation effects that were significantly stronger for transparent than for opaque conditions (Feldman, Milin, Cho, Moscoso del Prado Martín, & O'Connor, 2015; Feldman, O'Connor, & Moscoso del Prado Martín, 2009; Jared, Jouravlev, & Joanisse, 2017; Milin, Feldman, Ramscar, Hendrix, & Baayen, 2017; Schmidtke, Matsuki, & Kuperman, 2017). Such findings have been argued to support parallel distributed accounts, in which semantic information comes into play even at the early stages of visual word recognition, and morphology is not explicitly represented (arising instead from the statistical co-occurrences between form and meaning; see Gonnerman et al., 2007; Milin, Smolka, & Feldman, 2017; Plaut & Gonnerman, 2000).

Research on Semitic languages, such as Hebrew and Arabic, has presented a challenge to both decompositional and distributed accounts. In Semitic languages, word forms are based on stems, and most stems are formed by the non-linear combination of two bound morphemes: consonantal roots and vowel patterns. For example, the Hebrew word form *maxshev (מחשב)* "computer" is a stem formed by the combination of the root *X-SH-V (ח-ש-ב)* and the vowel pattern maCCeC.[8] Unlike the typical findings in Indo-European languages, root-priming effects in Semitic have been obtained for semantically opaque prime-target pairs as long as they shared a root, for example, *maxshev* (*מחשב*) "computer"–*mitxashev* (*מתחשב*) "considerate", not only in the masked priming paradigm (Boudelaa & Marslen-Wilson, 2005; Frost, Forster, & Deutsch, 1997), but also in cross-modal priming (Boudelaa & Marslen-Wilson, 2004, 2015; Frost, Deutsch, Gilboa, Tannenbaum, & Marslen-Wilson, 2000). Such results indicated that, in contrast to Indo-European languages, the recognition of Semitic words always depends on the extraction and activation of morphological constituents (viz., roots), even for forms with highly idiosyncratic meanings. In turn, this led to the suggestion that the Semitic lexical space is morphologically organized and "root-based", independently of semantic properties (Bick,

[8] We use "C" to represent root consonants.

Goelman, & Frost, 2011; Boudelaa & Marslen-Wilson, 2015; Deutsch, Frost, & Forster, 1998). We note, however, that a first indication that this view may require qualification is that semantic transparency effects have also been detected in cross-modal morphological priming in Hebrew (Frost et al., 2000, Experiment 2). In particular, even though morphological priming was obtained between opaque prime-target pairs, the effect was larger for transparently related pairs.

In sum, the findings regarding the existence of semantic transparency effects in morphological priming are mixed and complex. At the early stages of visual word recognition (as revealed by the masked priming paradigm), there is considerable disagreement about whether semantic effects exist (Heyer & Kornishova, 2018). There is less controversy that at the level of the "central lexicon" (as revealed by the cross-modal priming paradigm), morphological representation depends on semantic relatedness. On the other hand, semantic effects in morphological processing are not uniform across languages. In Semitic languages in particular, the finding that opaque (but morphologically related) forms prime each other in the cross-modal paradigm appears to undermine the view that morphological structure at the central lexical level goes hand-in-hand with relatedness in meaning.

In the present work, we propose that a more precise picture of the relation between morphological structure and semantic transparency across languages can be achieved if we consider that stems of lexical items can have different morphological representations. Specifically, stems (and word forms, more generally) may be structured and decomposable into their morphemic constituents; or alternatively, they may be represented as whole-stems and stored as single unstructured units (Clahsen, Sonnenstuhl, & Blevins, 2003). This representational distinction—a core feature of dual-morphology approaches (Clahsen, 1999; Pinker, 1999)—enables us to explain some of the linguistic contrasts and psycholinguistic findings obtained in Semitic languages, and at the same time, it generates a new prediction regarding the effects of semantic transparency.

Structured and unstructured stems in the Hebrew lexicon

Every Hebrew verb belongs to one of seven verbal classes or binyanim, each defined by a specific vowel pattern. For example, the verbs *lamad* *למד* "(he) learned" and *limed* *לימד* "(he) taught" both contain the root *L-M-D* *ל-מ-ד*, but the former belongs to the Paal class (displaying the vowel pattern *CaCaC*) and the latter belongs to the Piel class (displaying the vowel pattern *CiCeC*). We have previously proposed that stems of Paal and Piel (the most common classes) have different morpho-lexical representations in the Hebrew mental lexicon, structured [root +

pattern] representations for Piel verbs, and undecomposed whole-stems in the case of Paal verbs (Farhy, Veríssimo, & Clahsen, 2018a).

A first line of evidence for this proposal is that the Paal and Piel classes display a striking difference in their productivity. Despite comparably high type-frequency counts (19% Paal, 17% Piel; Itai & Wintner, 2008) and similar syntactic and semantic properties[9] (Berman, 1997), the two classes behave very differently in their propensity to being extended to new verbs. The Piel class has an "important productive function" as a source of new verbs (e.g., *fikes* *פִּיקֵס* from English *focus*; *fishel* *פִּישֵׁל* "mess up" from Arabic *fashla* "failure"), whereas Paal is "the least productive pattern, with almost no new verbs formed from denominal or loan sources" (Berman, 1997, p. 320). Indeed, a longitudinal examination of neologisms has shown that Piel readily welcomes new coinings, but new Paal verbs are almost never added to dictionaries (96 Piel vs. 2 Paal, out of 174 new verbs; Bolozky, 1999a). Within formal linguistic work, Paal's limited productivity has often been remarked upon and it has been linked to this class's greater morphological variability, greater variation in prosodic structure, and specific phonological restrictions (e.g., Aronoff, 1994; Bat-El, 2002; Laks, 2011). Finally, in elicited production tasks, native Hebrew speakers rarely form novel Paal verbs, whereas Piel verbs can be frequently elicited (Bolozky, 1978, 1982, 1999a). We have proposed that such a contrast in productivity can be straightforwardly explained if Piel, but not Paal stems are generated by a rule-based combinatorial operation (Farhy et al., 2018a); specifically, because rules contain variables or placeholders like "verbal root" (Marcus, 2001), they can be readily applied to novel tokens to create stems, namely Piel stem → [root V + *CiCeC*]. In contrast, if stems of the Paal class are stored as wholes and no such "Paal rule" is available to Hebrew speakers, then the class's ability to extend beyond its current set of verbs is necessarily reduced.

We have also found that inflected forms of Paal and Piel verbs yielded distinct masked priming effects (Farhy et al., 2018a). As has been claimed to be the general case for Semitic morphological processing (e.g., Deutsch et al., 1998), priming effects were elicited by Piel forms on targets with which they shared a root, for example, *limadti* *לימדתי* "(I) taught"—*hitlamed* *התלמד* "(he) interned". However, no root-priming effects were produced by well-matched Paal primes on the very same targets, for example, *lamadti* *למדתי* "(I) learned"—*hitlamed* *התלמד* "(he) interned", which indicated that, during early lexical access, verbal roots can only be automatically extracted from Piel, but not Paal stems. We have argued that this

[9] The Paal and Piel classes both contain, in the vast majority of cases, "activity verbs", in their active voice. However, most Piel verbs are transitive, while the Paal class is equally open to transitive and intransitive verbs (Berman, 1997).

contrast also suggests a representational difference between the two classes, with Piel stems being accessed as [root + pattern] structured representations, but Paal stems accessed as undecomposed wholes, due to the unavailability of rule-based decomposition.

The notion that Semitic stems or forms can be unstructured and represented as wholes is not new. Several theoretical accounts of Hebrew morphology question the necessity of the consonantal root as a separate constituent and explain a range of morphological phenomena by appealing only to stem-based representations (e.g., Bat-El, 1994; Ussishkin, 2005). Our specific proposal is that both structured (root-based) and unstructured (stem-based) representations exist within the Hebrew verbal system and that this representational distinction aligns with differences in productivity: structured stems in the case of productive, open-ended classes like Piel, and whole-stems without internal morphological structure, in the case of the unproductive Paal class (Farhy et al., 2018a; see Wray, 2016, for a similar proposal in Arabic).

In the current study, we hypothesized that this structural difference may also be reflected in a contrast between the two verbs classes in their susceptibility to semantic transparency effects in morphological priming. Specifically, if Piel stems are structured and contain the verbal root as a morphological constituent, then priming effects on other root-sharing verbs will primarily arise via structural overlap, that is, via the extraction and activation of shared roots—even between pairs that are not semantically related, as has been claimed to be the general case for Semitic languages (Boudelaa & Marslen-Wilson, 2005; Frost et al., 2000). If, by contrast, Paal verbs are processed as unstructured whole-stems, then the way they relate to other verbs with the same root cannot be via a shared morphological constituent. Rather, in tasks that are particularly sensitive to semantic properties, priming effects should be primarily based on semantic, not structural overlap (contrary to many previous claims about the processing of Semitic morphology; e.g., Boudelaa & MarslenWilson, 2015).

We tested cross-modal priming effects from inflected verb forms in Hebrew. In particular, we compared priming effects from Paal and Piel verbs on the recognition of root-sharing verbs, and examined whether such priming effects were modulated by the degree of prime-target semantic relatedness. Unlike most previous cross-modal morphological studies, which have factorized semantic relatedness into categories like “transparent” or “opaque”, we assessed these semantic effects as a continuous scale (following Heyer & Kornishova, 2018). If the recognition of Paal verb forms activates unstructured stems at the central lexical level, then semantic relatedness effects are expected to be larger for Paal primes. In contrast, if the processing of Piel forms always involves the activation of verbal roots (regardless of semantic overlap), then Piel primes are expected to show a reduced semantic transparency effect.

Method

Participants

Thirty native speakers of Hebrew (16 females, four left-handed) between the ages of 18 and 39 (mean: 28.63 years, SD: 4.79) participated in the experiment. All participants were born in Israel, had completed at least 12 years of education, used spoken and written Hebrew on a daily basis and had not been diagnosed with any language disorders.

Materials

The materials were taken from Farhy et al. (2018a; see Table 1). Experimental targets consisted of 42 Hitpael verb forms in the third person singular past tense, a form constituted by a three-consonant root combined with the Hitpael pattern (i.e., *hitCaCeC*). Each target word was paired with three types of primes: a prime belonging to the Paal class, a prime belonging to the Piel class (both primes shared a root with the target), and an unrelated prime with a different root, which was a Paal verb for half of the targets, and a Piel verb for the other half. Unrelated primes were neither semantically related (as assessed by a pre-test, described below), nor phonologically related to their targets (at most, one root consonant appeared in both prime and target, but never in the same position). Half (21) of the targets were preceded by primes presented in the first person singular past form, in which the verbal stem (with the vocalic structure *CaCaC* in Paal and *CiCaC* in Piel) is followed by the suffix *-ti* (1sg Past condition); the other half of the targets were preceded by primes presented in the infinitive form, which contained the prefix *l-* attached to the Paal and Piel stems (with the vocalic structure *CCoC* in Paal and *CaCeC* in Piel; Infinitive condition). Prime words were recorded by a female Hebrew native speaker, who was not aware of the purpose of the experiment.

The degree of semantic relatedness between primes and targets was assessed in a pre-test, conducted with a different group of 26 native Hebrew speakers (mean age: 30.77, SD: 12.80) using *SoSci Survey* (Leiner, 2014). Participants were instructed to rate “the degree to which each verb pair is related in meaning” on a scale of 1 (“very small degree”) to 7 (“very high degree”). Each of 52 potential Hitpael targets was paired with (a) a Paal verb with the same root, (b) a Piel verb with the same root, (c) an unrelated Paal verb, and (d) an unrelated Piel verb. Two versions of the questionnaire were created, so that Paal–Hitpael and Piel–Hitpael pairs that contained the same root would not be presented to the same participant. Each version was presented to half of the participants (in one of two orders to counterbalance for any effects of item position). Out of the 52 potential targets tested, 42 were selected for the priming experiment (see Farhy et al., 2018a).

Table 1

Experimental conditions, with an example stimulus set

Form	Prime			Target (Hitpael)
	Unrelated	Paal	Piel	
1sg Past	משכתי mSkty /maʃaxti/ pulled	אבדתי ʔbdty /avadti/ was lost	איבדתי ʔybdty /ibadti/ lost	התאבד htʔbd /hitʔabed/ committed suicide
Infinitive	לשפר lSpr /leʃaper/ to improve	לחלוק lxlwq /laxlok/ to share	לחלק Lxlq /lexalek/ to divide	התחלק htxlq /hitxalek/ shared / was divided

Note: Examples include Hebrew orthography, transliteration to Latin letters, phonological representation, and English translation (from Farhy et al., 2018a).

Table 2 displays the descriptive statistics for the (by-item) semantic relatedness ratings, in each of the experimental conditions. Unrelated primes were selected so that their semantic relatedness with the corresponding targets was uniformly low: mean scores were close to the low-end of the scale and displayed a small range and SD across items. In contrast, morphologically related primes (i.e., those in the Paal and Piel conditions) encompassed a wider range of semantic relatedness ratings, from closely related (ratings around 5), such as *xibakti חיבקתי* "(I) hugged"—*hitxabek התחבק* "(he) hugged (with)" to opaque (ratings around 2), such as *pikaxti פיקחתי* "(I) controlled"— *hitpake'ax התפקח* "(he) became clever". This variation allowed testing whether responses times (RTs) following Paal and Piel primes were predicted by semantic relatedness ratings. As can be seen in Table 2, means, SDs and ranges of semantic relatedness were well matched between Paal and Piel priming conditions. A mixed-effects model (with a by-participant random slope and a by-item random slope) showed that ratings of unrelated primes were significantly different from ratings given to both Paal (b = 2.28, SE = 0.25, t = 8.96) and Piel primes (b = 2.49, SE = 0.26, t = 9.47), with no significant difference between these two (b = 0.20, SE = 0.17, t = 1.20). Furthermore, probability density plots of Paal and Piel semantic relatedness scores (included in Figure 1 below) show similar distributions.

Table 2

Means, SDs and ranges of prime-target semantic relatedness ratings in each condition

	Unrelated			Paal			Piel		
Form	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
1sg Past	1.40	0.33	1.00–2.38	3.78	0.78	2.08–4.77	3.97	0.84	2.15–5.54
Infinitive	1.42	0.25	1.08–1.92	3.57	0.94	1.77–5.15	3.78	0.80	2.35–5.00

SD: standard deviation

Lemma frequency was also matched across conditions. Frequencies were obtained from a large Hebrew corpus (over 130 million tokens; Itai & Wintner, 2008) and converted to the scale (i.e., log10 of frequency per billion; van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Mean lemma frequencies for the 1st Past conditions were 4.27 (Unrelated), 4.21 (Paal), and 4.37 (Piel); in the Infinitive conditions, mean lemma frequencies were 4.43 (Unrelated), 4.33 (Paal), and 4.52 (Piel).

In addition to the 42 critical items, 294 fillers (126 words and 168 pseudowords) were included, so that every participant was presented with 336 targets. Three experimental lists were created according to a Latin-square design, with every critical target appearing once in each list. Every participant was randomly assigned to one list and therefore saw each critical target only once.

Procedure

Participants were asked to perform a lexical decision task on visual targets, as quickly and accurately as possible. They were instructed to press a gamepad button labeled “Yes” when they recognized an existing word in Hebrew (using their dominant hand) and to press another button labeled “No” when they were presented with pseudowords (using their non-dominant hand). The DMDX software (Forster & Forster, 2003) was used for stimulus presentation and data collection. There were 336 experimental trials in the experiment, presented in a pseudo-randomized order. Every trial consisted of the following events, in immediate succession: a fixation cross (500 ms), a blank screen (500 ms), an auditory prime word, and a visual target (which always started at the prime offset and was presented until a response was made, up to a timeout of 2000 ms). RTs were measured from the onset of target presentation.

Data analysis

One item (*hidama הידמה* "was similar") was removed from subsequent analyses due to very low accuracy (below 50%). All other items had accuracy rates of at least 80%. Incorrect responses (4.06%) and timeouts (0.32%) were removed from the dataset. The distribution of RTs was normalized by applying a reciprocal transformation (i.e., –1000/RT; Baayen & Milin, 2010).

The results were analyzed with mixed-effects regression with crossed random effects for subjects and items (e.g., Baayen, Davidson, & Bates, 2008). To examine priming effects from Paal and Piel forms, a first mixed-effects model was fit with Prime Type (Unrelated, Paal, Piel) as a fixed effect (interactions with Form Type, i.e., 1sg Past and Infinitive, were also considered, but were found not to improve model fit; see below). To assess semantic relatedness effects, a second model was fit with Prime Type (Paal, Piel) and Semantic Relatedness (centered) as fixed predictors, as well as their interaction. Finally, in order to compare semantic relatedness effects in the current experiment and in a masked priming experiment with the same materials (Farhy et al., 2018a), the fixed-effect experiment (Cross-modal, Masked) was also included in the model (as well as its interactions with Prime Type and Semantic Relatedness).

In order to reduce the probability of detecting spurious effects, it is commonly recommended for "random slopes" to be included in mixed-effects regression models, which allows variation in effects across participants and items to be modeled and allows rates of Type I errors to be controlled (e.g., Barr, Levy, Scheepers, & Tily, 2013). Here, we followed the recommendations of Matuschek, Kliegl, Vasishth, Baayen, and Bates (2017) and tested random slopes for inclusion on the basis of whether they improved goodness-of-fit (as measured by Akaike information criterion). For each regression model that we conducted, all possible random structures were assessed. In every case, the model with the lowest Akaike information criterion (i.e., with the best fit) was achieved with an intercept-only random structure, that is, without random slopes. These are the models reported below.

Results

Mean RTs and accuracy rates in each condition were computed (Table 3). Given that all accuracy rates were above 90%, no further analyses of accuracy were conducted.

Interactions of Prime Type and Form Type were not significant (with Paal priming, t = 0.75; with Piel priming: t = −0.85), that is, magnitudes of priming elicited by Paal and Piel verbs were not modulated by whether they were presented as infinitives or as 1sg past forms.

In addition, including Form Type as a fixed effect (together with its interaction with Prime Type) did not improve model fit, $\chi^2(3) = 2.97$, $p = .396$. Therefore, priming effects were assessed taking the two Form conditions together.

Table 3

Mean RTs, standard errors (in parenthesis), and accuracy rates in each condition

	Unrelated		*Paal*		*Piel*	
Form	RT (ms)	Accuracy	RT (ms)	Accuracy	RT (ms)	Accuracy
1sg Past	600 (9.98)	95%	574 (8.92)	95.5%	587 (9.58)	95%
Infinitive	617 (10.56)	92.4%	595 (9.70)	96.7%	584 (9.49)	99%
Both	609 (7.66)	93.7%	585 (6.60)	96.1%	585 (6.74)	97.1%

Note: Means and SEs were calculated from reciprocal RTs and back-transformed. RTs: response times; ms: milliseconds; SEs: standard errors.

Priming effects were obtained for both Paal ($b = -0.0713$, $SE = 0.0228$, $t = -3.13$) and Piel primes ($b = -0.0751$, $SE = 0.0238$, $t = -3.30$), with significantly shorter RTs after their presentation than after Unrelated primes (but no difference between the RTs in the Paal and Piel conditions; $b = 0.0038$, $SE = 0.0226$, $t = 0.17$).

A significant interaction between Semantic Relatedness and Prime Type was obtained ($b = 0.0769$, $SE = 0.0308$, $t = 2.50$), which is depicted in Figure 1(a). For Paal primes, larger semantic relatedness scores were associated with shorter RTs ($b = -0.0611$, $SE = 0.0247$, $t = -2.47$). In contrast, Piel RTs were not modulated by semantic relatedness ($b = 0.0158$, $SE = 0.0255$, $t = 0.62$). As shown in Figure 1, the semantic relatedness effect for Paal was sizable, 72 ms (in back-transformed RTs) between the minimum and the maximum relatedness score (1.77–5.15, in a 7-point scale).

Finally, semantic relatedness effects in the current cross-modal priming experiment were compared against the previously obtained masked priming dataset (Farhy et al., 2018a; effects of semantic relatedness were not examined in that study).[10] A significant three-way interaction among Experiment, Prime Type and Semantic Relatedness was obtained ($b = 0.0773$, $SE = 0.0363$, $t = 2.13$), indicating different modulations of morphological priming effects by semantic transparency in the two experiments (see Figure 1). In contrast to the current cross-modal priming experiment, there were no significant interactions between Semantic

[10] In order for this comparison to be appropriate, the reanalysis of the masked priming dataset was conducted in a perfectly parallel way to the analysis of the cross-modal priming experiment and with exactly the same items included.

Relatedness and Prime Type in our previous masked priming experiment (b = 0.0061, SE = 0.0281, t = 0.22). In addition, semantic relatedness did not have an effect on RTs following Paal (b = 0.0059, SE = 0.0235, t = 0.25) nor Piel primes (b = −0.0002, SE = 0.0238, t = −0.01).

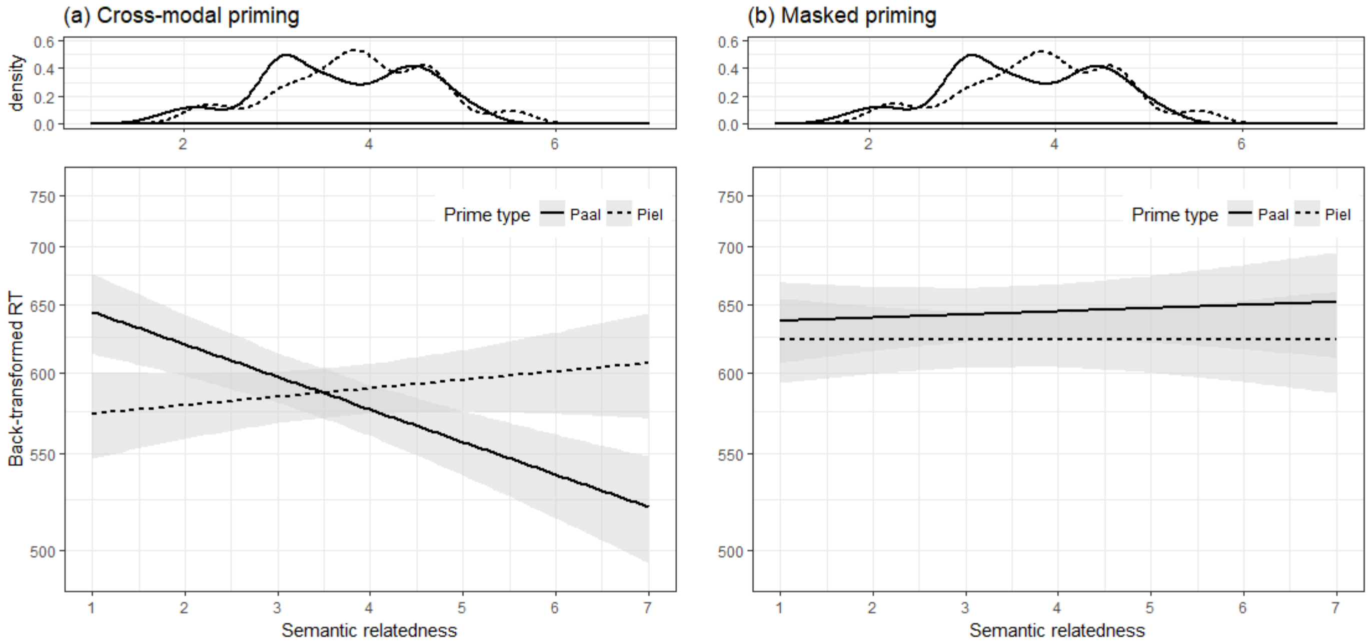

Figure 1. RTs following Paal and Piel primes as a function of semantic relatedness, (a) in the current cross-modal priming experiment and (b) in the masked priming experiment of Farhy et al. (2018a). The y axis displays back-transformed RTs (from reciprocal RTs). Shaded bands represent pointwise standard errors. Each panel includes probability density plots of Paal and Piel semantic relatedness scores. RTs: response times.

Discussion

The main finding of the present cross-modal priming study was that morphological (root-priming) effects between Hebrew verbs were modulated by semantic transparency only when primes were members of the unproductive class Paal, but not when they belonged to the productive, open-ended class Piel. More specifically, whereas both Paal and Piel primes facilitated responses to targets with which they shared a root, a greater degree of prime-target semantic relatedness was only associated with faster responses in the case of Paal (but not Piel) verbs. This dissociation between verb classes complements our previous finding that, under masked priming conditions, only Piel (but not Paal) verbs produced root-priming effects (Farhy et al., 2018a). When taken together, the two sets of findings have several implications for the role of morphology in lexical representation, as well as for the generality of previous claims regarding the Semitic mental lexicon.

Our results present a challenge to both distributed models and to full-decompositional models of the Semitic lexicon. In distributed models, morphological priming effects are attributed to the graded statistical correlations between form and meaning (Gonnerman et al., 2007), and thus, semantic transparency effects (as obtained for Paal verbs in the present study) are well-accommodated in these approaches. However, it is hard to see why such effects should differ between verbal classes in a distributed model, especially when the Paal and Piel classes as a whole have similar type frequencies and semantic properties. Conversely, models in which all verbs in Hebrew are represented as structured [root + pattern] combinations regardless of meaning properties (e.g., Deutsch et al., 1998) can account for the lack of processing differences between Piel verbs with low and high semantic relatedness to their targets. However, such proposals cannot accommodate the contrast in experimental effects elicited by verbs of the Paal and Piel classes, because both are assumed to be represented in the same way. In contrast to both of these accounts, we propose that our results can be explained by postulating different morpho-lexical representations for stems of different classes: (a) *structured stems*, which display internal constituent structure, in the case of productive classes like Piel, and (b) *whole stems*, which are represented as whole undecomposable units, in the case of an unproductive class like Paal (see Clahsen et al., 2003; Wray, 2016; for similar proposals in German and Arabic).

The starting point for our account is that constituent structure aligns with productivity, a link that is seen across morphological phenomena and across languages (e.g., Ford, Davis, & Marslen-Wilson, 2010; Newman, Ullman, Pancheva, Waligura, & Neville, 2007; Sonnenstuhl, Eisenbeiss, & Clahsen, 1999; Veríssimo & Clahsen, 2009). Productive classes like Piel can easily extend to novel roots because they are associated with morphological stem-formation rules: combinatorial operations that interleave roots and vowel patterns to generate stem representations with internal morphological constituents. For a closed-class like Paal, we have proposed that this stem-formation operation is not available in the grammar of Hebrew speakers, thereby explaining this class's restricted productivity (Farhy et al., 2018a).

Whereas structured stems can be decomposed into their morphemic constituents in the course of lexical processing, whole-stem representations lack the internal structure that supports the segmentation and extraction of morphemic constituents. Thus, access to undecomposed Paal stems does not activate separate representations of their constituents, explaining the lack of Paal priming in the masked priming paradigm (Farhy et al., 2018a)—a task which is arguably more sensitive to the role of pure morphological structure (Marslen-Wilson, 2007). In the current study, using an experimental paradigm that is sensitive to both morphological structure at the central lexical level and to the semantic overlap between primes and targets (Frost et al., 2000; Marslen-Wilson et al., 1994), we see a priming effect emerge for Paal verbs, but one that is modulated by semantic relatedness. We believe this to be a consequence of the whole-stem processing of Paal verbs: Unstructured stems do not structurally contain roots as morphological constituents, so that any effect on targets that display the same root cannot be due to the pre-activation of structural elements. Rather, this effect is likely to be semantic (or conceptual) in nature, a reflection of the graded overlap in meaning between Paal verbs and targets (cf. Gonnerman et al., 2007); as such, it is present in the cross-modal paradigm, but absent in masked priming (where semantic processing is typically found to be reduced; e.g., Rastle et al., 2000, 2004).

By contrast, the imperviousness of the Piel priming effect to semantic relatedness shows that Piel roots are extracted and activated independently of the semantic properties of particular lexical entries. That is, in the case of productive classes like Piel, opaque forms and transparent forms are processed in the same way, with the verbal root functioning as a purely morphological constituent that has a structural significance that goes beyond its role in the computation of meaning. These results are similar to those previously obtained on Semitic morphological processing, and have been used to claim that, unlike Indo-European languages like English (Marslen-Wilson et al., 1994; Gonnerman et al., 2007) or French (Longtin et al., 2003), the Semitic mental lexicon is primarily organized along morphological lines (Bick et al., 2011; Boudelaa & Marslen-Wilson, 2004, 2015; Deutsch et al., 1998). However, the full pattern of our results suggests that such claims may require qualification, as they do not apply to the whole of the Semitic lexicon. Rather, for a specific type of morpho-lexical representation—whole unstructured stems, as in Paal verbs—Semitic languages may be less unique than previously thought, showing morphological priming effects that are strongly dependent on meaning. Thus, our results further restrict the predominant view regarding the special status of non-concatenative morphological systems, and contribute to the body of work that has highlighted the similarities between Indo-European and Semitic morphology, both in psycholinguistics (Perea, Mallouh, Mohammed, Khalifa, & Carreiras, 2018; Vaknin-Nussbaum & Shimron, 2011; Velan, Deutsch, & Frost, 2013; Velan & Frost, 2011) and in formal linguistics (Bat-El, 1994; Ussishkin, 2005). More generally, we believe that the dissociation we obtained in the present study is to be expected when one considers that the contrast between internally structured representations and those that are stored as wholes may be a universal design feature of language (Clahsen, 1999; Pinker, 1999).

Acknowledgements

We thank Mark Gelber for help with the recruitment of participants. We are also grateful to Harald Clahsen, Vera Heyer, and two anonymous reviewers for helpful comments on an earlier version of this paper.

Funding

This work was supported by a PhD scholarship awarded to Yael Farhy by the Minerva Fellowship Program. Our work on this article has also been supported by an Alexander-von-Humboldt Professorship awarded to Harald Clahsen.

6 Publication IV

Accepted for publication in *The Mental Lexicon*

Morphological Generalization of Hebrew Verb Classes: An Elicited Production Study in Native and Non-Native Speakers

Yael Farhy
Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam

Abstract

The present work investigated how morphological generalization, namely the way speakers extend their knowledge to novel complex words, is influenced by sources of variability in language and speaker properties. For this purpose, the study focused on a Semitic language (Hebrew), characterized by unique non-concatenative morphology, and native (L1) as well as non-native (L2) speakers. Two elicited production tasks tested what information sources speakers employ in verbal inflectional class generalization, i.e., in forming complex novel verbs. Phonological similarity was tested in Experiment 1 and argument structure in Experiment 2. The analysis focused on the two most common Hebrew inflectional classes, *Paal* and *Piel*, which also constituted the vast majority of responses in the two tasks. Unlike the commonly found outcomes in Romance inflectional class generalization, the results yielded, solely for Piel, a graded phonological similarity effect and a robust argument structure effect, i.e., more Piel responses in a direct object context than without. The L2 pattern partially differed from the L1: (i) argument structure effect for L2 speakers was weaker, and (ii) L2 speakers produced more Paal than Piel responses. The results are discussed within the framework of rule-based and input-based accounts.

How are speakers able to understand and produce words and sentences they have never read or heard before? In the domain of morphology, for example, speakers routinely extend their knowledge to novel forms in order to produce and understand complex novel words, in a process that is referred to as *morphological generalization*. Thus, when encountering a novel verb like *ploamph,* English speakers can effortlessly produce a past tense form, *ploamphed*. Several approaches have attempted to explain the process of morphological generalization. Connectionist and analogical single-system models have argued that morphological generalization can be solely explained by similarity between a novel item and learned items stored in memory (e.g., Rumelhart & McClelland, 1986; Skousen, Lonsdale & Parkinson, 2002). Essentially, the higher the similarity between a novel form and a set of stored forms, the higher the probability that the same type of form is generated. In contrast, other accounts have argued for a dual-route system, in which an application of productive symbolic rules is additionally required (Prasada & Pinker, 1993; Pinker, 1999). Despite the disagreement about its exclusiveness, both approaches agree similarity plays a major role in morphological generalization.

To reach a deeper understanding of similarity-based generalization, different sources of similarity should be taken into account. Research on Indo-European languages has provided evidence that native speakers generate complex novel forms based on phonological and semantic similarity. When presented with a novel word like *spling*, English speakers often produce the past tense *splang (*alongside with *splinged)* due to the phonological similarity of *spling* to other irregular verbs such as *sing* and *ring*. Robust evidence for the role of phonological similarity was found in various languages and morphological phenomena, such as the English past tense (e.g., Albright & Hayes, 2003; Eddington, 2000; Rumelhart & McClelland, 1986), French gender classes (e.g., Matthews, 2010) and German plural inflection (e.g., Hahn & Nakisa, 2000). Similarity-based generalization in morphology has also been demonstrated with semantic sources, although to a smaller extent. When presenting a novel verb like *frink* in a context that is similar to the irregular verb *drink* (such as in a sentence that includes drinkable objects), an irregular inflection is produced more often than a regular one, and when the semantic context is similar to a regular verb (*wink*), the pattern reversed (Ramscar, 2002).

So far, the majority of morphological generalization studies have examined native speakers of Indo-European languages, not taking into account that the great variability in morphological properties across languages and within characteristics of language users can possibly affect the type of information speakers use in forming complex novel words.

Addressing these two sources of variability will help to better explain and predict the type of information used by the speaker to generate novel complex forms. Therefore, the present study examined morphological generalization of Hebrew *verb classes*, characterized by a relatively unique non-concatenative morphological structure, and tested non-native (L2) speakers, a less explored population, along with native (L1) speakers. In two elicited production experiments, the present work sought to reveal what type of information, phonological similarity (Experiment 1) and/or argument structure (Experiment 2), are utilized in the formation of novel Hebrew verbs by L1 and L2 speakers. To better understand the rationale and design of two experiments, the following sections introduce (i) relevant theoretical and empirical background related to Hebrew verb classes and their generalization, and (ii) an overview of the current literature about morphological generalization in L2 speakers, explaining how certain properties of L2 users can potentially influence morphological generalization.

Generalization of Hebrew verb classes

Verbal inflectional classes, common in Romance and Slavic languages, are morphological classes that determine the form of a verb stem and its inflection paradigm. The system is obligatory, in a sense that every verb must belong to a class (Aronoff, 1994). Generalization of verbal inflectional classes has been suggested to rely strongly on phonological information. According to similarity-based single-mechanism approaches, phonological properties alone can explain verb generalization (Albright, 2002; Eddington, 2002). In addition, it is a central factor in dual-mechanism approaches, alongside with symbolic rule-based operations (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014). Studies on generalization of inflectional classes of Romance languages have manipulated phonological similarity based on rhyming of novel words to existing words (Say & Clahsen, 2002) or simulation of computational models (e.g., analogical models; Eddington, 2002), and overall yielded a robust evidence that individuals make use of phonological similarity when generalizing inflectional class to novel verb forms.

Hebrew also displays verbal inflectional classes that determine the inflection paradigm of the verb. Every Hebrew verb belongs to one of seven classes called *binyanim* (singular: *binyan;* see Table 1 for an overview). Every verb lexeme (or lexical entry) is formed by a non-concatenative combination of a consonantal root and one of seven verb-specific vowel patterns. The root (usually contains three consonants, such as *L-M-D*) is inserted non-linearly into the designated positions of the vowel pattern (such as *CaCaC*, resulting in the form *lamad*). This non-linear combination creates a verb lexeme with a certain meaning and subcategorization properties, and at the same time determines the verb class of the lexeme (e.g., *lamad* ‘learn’ is a lexical entry which is assigned to class *Paal*; Aronoff, 1994). The same root is sometimes combined with more than one vowel pattern, creating different lexical entries with the same root. For instance, one can combine the root *L-M-D* also with the pattern *CiCeC* and create the lexical entry *limed* 'teach' in class *Piel*. Each verb has a paradigm of forms, where verbs are inflected to express morpho-syntactic features such as tense, person, number and gender (e.g., within the lexical entry *lamad* the form *lamadti* is inflected for past tense first person singular).

Two main points differentiate the Hebrew classes from inflectional classes of Romance and Slavic languages. First, stem formation of verbs (and most nouns) is non-concatenative. This type of morphological system has been previously proposed to have a lexical space that is organized around roots as units of morphological combination, with a weaker dependence on semantic and orthographic properties compared to other languages (Boudelaa & Marslen-Wilson, 2015; Deutsch, 2016; Velan & Frost, 2011; but see Farhy, Veríssimo, & Clahsen, 2018a). The constituents (roots and patterns) in the non-concatenative structure cannot be pronounced on their own and therefore they cannot rhyme with other roots and patterns (a root like *L-M-D* does not rhyme with *X-M-D*, despite sharing two consonants). This point may be crucial for the reliance on phonological similarity in Hebrew verb class generalization, since rhyming between words tend to overlap with phonological similarity (e.g., words like *sing* and *ring* are considered highly similar phonologically: Bybee & Moder, 1983; Prasada & Pinker, 1993). Therefore, Hebrew verb class generalization may involve a reduced reliance on phonological sources[11] (but they are still central for other Semitic phenomena, such as Hebrew plural suffix inflection: Berent, Pinker, & Shimron, 1999; Levy, 1983; and broken plurals in Arabic: Dawdy-Hesterberg & Pierrehumbert, 2014).

The second unique property of the Hebrew classes compared to other inflectional classes is that they have certain semantic and syntactic tendencies. Their main tendency refers to their argument structure (can also be addressed in terms of subcategorization frame), that is, certain classes have a tendency to take a direct object (also called transitive verbs), while others do not take a direct object (Table 1). Since argument structure is strongly linked to semantic properties (via linking rules which map semantics onto syntactic functions, for details see, for example, Pinker, 2013), semantics is also closely associated with certain classes.

[11] Nonetheless, phonological restrictions do exist in verb class generalization; certain classes can take only three consonant roots, while others can take more.

Table 1

Overview of the properties of the Hebrew inflectional classes (binyanim)

Class name	Vowel pattern (3p.sg.m.past)	Argument structure (transitivity)
Paal	CaCaC	no tendency
Piel	CiCeC	usually contain a direct object (transitive)
Hitpael	hitCaCeC	never contain a direct object (not transitive)
Hifil	hiCCiC	usually contain a direct object (usually transitive)
Nifal	niCCaC	never contain a direct object (not transitive)
Pual	CuCaC	never contain a direct object (not transitive)
Hufal	huCCaC	never contain a direct object (not transitive)

Psycholinguistic evidence suggests that syntactic and semantic tendencies of the Hebrew classes affect their generalization. In an elicited production study (Berman, 1993) adults and children were instructed to produce a novel verb that describes the action in a picture, such as a picture where a man is causing the child to swim. In 90% and 77% of the cases adults and eight-year-old children produced a novel verb by altering the class of an existing intransitive verb to a class that is typically transitive, and added a direct object (e.g., alternating the verb *soxe* 'swim' [Paal, 3p.sing.present] to the novel Hifil verb *masxe* and the novel Piel verb *mesaxe*). Children at age two and three performed at chance level. These findings showed that at the age of eight Hebrew speakers are already influenced by the classes' argument structure tendencies. In another elicited production study (Bolozky, 1999a) adult native Hebrew speakers were presented with sentences, and were instructed to fill in the blank with novel verbs from loaned words (existing nouns, such as *sheriff*). The results showed that Piel verbs were the most frequent response when the focus was on the agent (e.g., *shirfu* in Example 1), and *Hitpael* – when the focus was on theme (e.g., *hishtaref* in Example 2). Again, Hebrew speakers were influenced by argument structure tendencies of verb classes in generating novel verbs.

(1) *geri kuper lo ratsa lihiyot sherif, toshvei ha'ir ____ oto bexo'ax. (asu oto lesherif)*
'Gary Cooper did not want to be a sheriff. The town's people _____ him with force. (Made him a sheriff)'

(2) *geri kuper lo ratsa lihiyot sherif. hu _____ lamrot zot. (hafax lesherif)*
'Gary Cooper did not want to be a sheriff. He ______ despite that. (Became a sheriff)'

However, these studies were not able to examine the use of argument structure and semantic information separately, because, as noted earlier, those properties are strongly linked. The present study attempts to fill this gap, examining the role of argument structure alone in generalization of the Hebrew classes, minimizing any influence of semantics by asking the participants to form a verb out of novel words and not out of existing words. To the author's knowledge, the relation between argument structure and Hebrew inflectional classes is unique and is not present non-Semitic inflectional classes. Nonetheless, a relation between argument structure and morphological form can be seen in non-Semitic languages. For example, the derivational German prefix *be-* often changes the argument structure of a verb, turning an intransitive verb to transitive one (*sprechen* 'to talk' [intransitive] → *besprechen* 'to discuss' [transitive], Maylor, 2002). Language users are known to be sensitive to argument structure information, guiding their interpretation of semantic meaning of the novel verb (e.g., Fisher, Hall, Rakowitz, & Gleitman, 1994; Naigels, 1996; Naigles & Kako, 1993). Furthermore, in the case of familiar verbs, the information about their argument structure guides the initial processing of the sentence, enabling the speaker to predict the argument structure after reading the verb (Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Traxler, 2005).

Morphological generalization in non-native (L2) speakers

Morphological generalization in L2 speakers can be examined from different directions. First, observing L2 morphological errors in production of complex words shows that speakers often extend their morphological knowledge in ways that lead to errors. Such errors are common in L2 speakers, even in highly proficient ones (e.g., Dewaele & Veronique, 2001; McCarthy, 2008; McDonald & Roussel, 2010; Montrul, 2011; Parodi, Schwartz, & Clahsen, 2004; White, Valenzuela, Kozlowska–Macgregor, & Leung, 2004). The most common types of morphological errors in L2 speech are production of forms less specified than required, default forms, or over-applying bare word forms (e.g., McCarthy, 2008; White et al., 2004), as well as *overregularization*s, in which speakers extend regular morphological patterns to irregular words, such as producing *singed* as the past tense of *sing*, over-applying the *-ed* suffix.

Another direction to look at L2 morphological generalization is via novel word production tasks. In forming novel complex words, speakers can rely on numerous cues; they can employ grammatical cues, such as argument structure of the verb, or non-grammatical cues, such as the phonology or orthography of the word. So far, studies have focused on the use of phonological similarity, showing that L2 speakers make use of phonological similarity like L1 speakers (English past tense: Cuskley, Colaiori, Castellano, Loreto, Pugliese, & Tria, 2015; Plag, 2000; German participles: Hahne, Mueller, & Clahsen, 2006) or even to a larger extent than L1 speakers (Greek verb inflection for active perfective aspect: Agathopoulou & Papadopoulou, 2009; Stavrakaki & Clahsen, 2009). For example, Cuskley et al. (2015) conducted an elicited production study on English novel verbs, and reported that for both L1 and L2 speakers past form formation (regular/irregular) was dependent on the phonological distance between the novel word and existing words (regular and irregular). The main difference between the groups was that the L2 speakers produced more irregular past tense forms in overall, which was attributed to the relative limited input they have been exposed to, resulting in larger weight to the very frequent irregular verbs in generating novel forms.

It is still not clear how L2 speakers apply more abstract cues, like argument structure information, in morphological generalization. Evidence has shown that argument structure guides L2 initial processing, enabling the speaker to predict the verb’s argument structure (Jiang, 2004, 2007), yet sometimes to a smaller extent than L1 speakers (Dussias & Schaltz, 2008). In self-paced reading task, Dussias & Scaltz (2008) reported slower reading times of L2 speakers when the argument structure did not match with the verb structural bias (e.g., when the verb *admit* appeared with a direct object, and not with a complement phrase). However, the effect was weaker than for the L1 speakers. In an offline task (Dussias, Marful, Gerfen, & Molina, 2010) English L2 speakers were reported to possess the knowledge about argument structure (or subcategorization frame) biases associated with English verbs, although less successfully than L1 speakers.

Another method of investigating morphological generalization in L2 is via Artificial Language (AL) experiments, in which the participants acquire a novel language throughout the experiment session (Morgan-Short, Sanz, Steinhauer, & Ullman, 2010; Morgan-Short, Steinhauer, Sanz, & Ullman, 2012; Steinhauer, White, & Drury, 2009). Such experiments have reported that adults were able to acquire and generalize subclasses based on phonological markers (Brooks, Braine, Catalano, Brody, & Sudhalter, 1993; Culbertson, Gagliardi, & Smith, 2017; Frigo & McDonald, 1998), semantic properties (e.g., Culbertson & Wilson, 2013) and syntactic context (also can be referred to as distributional cues; Reeder, Newport, & Aslin, 2017). However, evidence from AL studies also yielded that age of acquisition may influence the strategy employed in morphological generalization. While children performed overregularizations, generating the regular forms in almost 100% of the cases, adults employed a probability matching strategy, generating regular forms based on their token frequency (Kam & Newport, 2005, 2009; Schuler, Reeder, Newport, & Aslin, 2017). This finding may be generalized to non-artificial languages, suggesting that L1 and L2 speakers make use of different information sources in morphological generalization.

The present study

The present work examines Hebrew inflectional class generalization, focusing on (i) whether Hebrew speakers employ phonological associations (which are dominant for morphological generalization cross-linguistically) and argument structure information (considered to be Semitic-specific, at least in inflectional class generalization), and (ii) whether L2 speakers employ such information (if at all) to a different degree than L1 speakers. For this purpose, two elicited production experiments were conducted (as in much previous research testing morphological generalization, for example, Berent et al. 1999, Prasada & Pinker, 1993; Say & Clahsen, 2002; Veríssimo & Clahsen, 2014), manipulating phonological similarity of novel roots to classes (Experiment 1) and argument structure (specifically the presence of a direct object; Experiment 2). The two experiments directly compared L1 and L2 Hebrew speakers (highly proficient late-learners) on their production of novel verbs from novel nouns. Production of novel verbs requires combining a novel root (extracted from a novel noun) with a vowel pattern. For example, the novel root *S-L-Z* is extracted from the novel noun *selez* and when chosen to be combined with the Paal vowel pattern *CaCaC*, it yields the novel verb *salaz*. Of course, seven vowel patterns are always possible for forming a novel verb, reflecting the seven inflectional classes. However, both Experiment 1 and Experiment 2 focused on production of the two most common classes, Paal and Piel.

Experiment 1 tested whether Hebrew speakers use phonological analogy in generalizing inflectional classes from novel roots via an implementation of an analogical model to Hebrew, which simulated generalization of the classes based on analogical similarity of roots (Eddington, 2002). If phonological analogy is used in generalization, production of Paal verbs will increase when the novel roots are similar to Paal (compared to novel roots that are not similar to any class based on the model). Similarly, production of Piel verbs will increase when the novel roots are similar to Piel (compare to roots of no similarity to any class). Given previous evidence (e.g., Agathopoulou & Papadopoulou, 2009; Cuskley et al., 2015), L2 speakers are predicted to perform similarly to L1 speakers, or even show a stronger reliance on phonological similarity.

Experiment 2 was similar to Experiment 1, but manipulated argument structure (while minimizing semantic/thematic features), a property that was expected to affect generalization of Hebrew inflectional classes, unlike in non-Semitic languages. Do Hebrew speakers use information about the argument structure in generalizing Paal and Piel classes? Verbs in Piel are often transitive (79% of Piel verbs[12]), followed by a direct object (see Piel example (3a)), while Paal verbs do not have any tendency in this regard (46% transitive vs. 54% intransitive; see Paal examples (3b,c); Berman, 1997).

(3) a. hayeled nishek et hayalda
the-boy kissed ACC the-girl
'The boy kissed the girl.'
b. hayeled daxaf et hayalda
the-boy pushed ACC the-girl
'The boy pushed the girl.'
c. hayeled halax im hayalda
the-boy walked with the-girl
'The boy walked with the girl.'

Given previous findings showing that both adults and children are sensitive to the relationship between classes and argument structure (Berman, 1993), only Piel verb production would be affected by the argument structure manipulation; an increase in Piel responses is expected when the sentence contains a direct object (compared to only a prepositional phrase without a direct object). Regarding the L2 group, if their reduced sensitivity to argument structure during processing (reported in Dussias & Schaltz, 2008) applies also to morphological generalization, they will show a smaller Piel increase (if any) in sentences with a direct object.

[12] This number is based on manual coding of all verbs from a large Hebrew corpus (Itai & Wintner, 2008).

Experiment 1

Method

Participants

The experiment included two groups: native (L1) and non-native (L2) Hebrew speakers. The L1 group included 28 native speakers of Hebrew (24 females) between the ages of 21 and 33 (mean: 24.29 years), born in Israel.

The L2 group included 23 non-native speakers of Hebrew (9 females) between the ages of 22 and 37 (mean: 28.00 years). The L2 participants followed two important criteria: (1) they were late learners, acquiring Hebrew after the age of 7 (mean age of onset: 12.7 years, SD: 3.65) and (2) spoke Spanish as L1, so that all participants were familiar with the phenomenon of verbal inflectional classes from their L1. They were living in Israel and had emigrated from South America. Nonetheless, they all had at least 11 years of Hebrew exposure (mean: 15.35 years, SD: 2.82), and very high proficiency – they achieved a mean score of 98.45% in a section of the YAEL proficiency test for university candidates (including sentence completion, sentence rephrasing, and reading comprehension), with every participant scoring above 80%. Furthermore, the L2 participants estimated (in percentages) their relative use of Hebrew overall and in four specific domains (speaking, hearing, writing, and reading). Overall reported Hebrew usage had a mean of 63.70% (SD: 17.27). Relatively high ratings were reported also for the four specific domains (speaking: 67.17%; hearing: 63.48%; writing: 78.26%; reading: 66.96%). Participants used their native languages (Spanish) less often, with a mean overall usage of 25.00% (SD: 11.68). All participants had normal or corrected-to-normal vision and none had been diagnosed with any language disorders.

Simulation

The materials were constructed based on an implementation of the *Analogical Modelling of Language* (AML; Eddington, 2002)[13]. The AML is an exemplar-based model, predicting linguistic behavior based on analogy to stored linguistic experiences. The aim of the implementation was to predict a class for novel Hebrew roots based on analogy to a large database of existing roots. For this purpose, a large Hebrew lexical database of over 130 million

[13] The simulation was completed with the great assistance of Mr. Jonathan Engel, a student of Stanford University, during his internship at Potsdam Research Institute for Multilingualism.

tokens (Itai & Wintner, 2008) was used for the model's input. The database contained 4,694 verbs in the seven verb classes. Verbs with roots of more than three consonants were removed (563 verbs), because these roots have phonological constraints that allow them to only be inserted into certain verb classes (Piel, Pual and Hitpael; Bolozky, 1999a). Since verbs from the classes Pual and Hufal are fully predictable as the passive forms of Piel and Hifil, they were removed from the model's input. The verbs from the remaining classes (3,247) were encoded in a phonetic transcription. In the end, the database included 3,247 entries, each entry was a verb represented by a root and a class name (e.g., *lamad* '(he) learned' was represented as *L-M-D*, Paal). Afterwards, a list of novel roots (1,534) was computed, consisting all possible three consonant combinations that did not appear in the database (avoiding roots that have the same consonant in first and second position, as such roots only exist in few very rare cases and therefore do not sound natural in Hebrew, see Berent & Shimron, 1997).

In order to produce a certain output to a given input, in the current case – a certain class to a given novel root – the AML algorithm searches throughout the lexical memory for entries with roots that are most similar to the given input form, i.e., the novel root. It then creates groups of entries from the database called sub-contexts, each shares phonological similarities with the input form (in the case of Hebrew, it is reflected by root consonant overlap). Entries that share more features with the given context appear in more sub-contexts, and carry more weight in predicting the class of the given input. In the present case the output of each given input was a probability of belonging to each one of the five classes. Yet in many cases only one class had a probability larger than zero to be assigned; for example, the simulation assigned the class Paal to the non-existing root *SH-TS-R*, as sub-contexts like _-*TS-R*, appear more in Paal, such as Y-TS-R (*yatsar* 'created') and ʕ-TS-R (*atsar* 'stopped'). On the other hand, the novel root *SH-X-S* was assigned to Piel, based on similarity to roots such as *Y-X-S*, appearing in Piel (*yixes* 'referred').

Materials

Based on the list of novel roots and the implementation of the AML model, 60 novel roots were selected; 20 of them were phonologically similar to Paal (and not to any other class), 20 – to Piel (and not to any other class) and 20 were not similar to any verb class[14]. All roots contained

[14] The strength of similarity produced by the model was taken into account in selecting the items. I attempted to select the items so the range of similarity strength will be as similar as possible between the classes. Overall, the strength of the Piel roots tended to be higher (range in Paal-similar roots: 27–2282, and range in Piel similar roots: 38–15698), therefore a complete match of the similarity strength between the root groups was not possible.

three letters, in line with the typical Hebrew root structure and were selected to sound as natural as possible in Hebrew. Furthermore, only regular roots were selected, to allow simple production of verbs of all classes. For example, roots which contained the letters 'Y' (representing the phonemes /j/ or /i/) or 'V' (/v/ /o/ or /u/) were not included because these letters are often silent or missing in certain classes and inflections (e.g., the Paal verb *shar* '(he) sang' contained the root *SH-Y-R*). The next step was to form nouns from these roots. The nominal pattern *CeCeC* was selected, since it consists of only the root letters without affixes and the vowels are visually represented by (optional) diacritic marks under the letters. Since the purpose was for the verb to be elicited only from the root consonants (and not from affixes), the pattern *CeCeC* was a suitable choice. All novel roots were inserted into the vowel pattern (including the diacritic marks), forming novel nouns like *SeLeZ* סֶלֶז.

Procedure

Participants performed the task on the web via *SoSci Survey* (Leiner, 2016). The experiment included 60 experimental items. Every experimental item was presented only once and was embedded within a two-sentence frame. The first sentence presented the novel noun in bold and with diacritic marks, so the vowels of the word were explicitly represented in the written form. The second sentence contained a temporal adverbial referring to the past (such as 'yesterday'), a first person singular subject pronoun (*ani*) and blank space (see Table 2).

The participants were instructed to fill in the blank space with a verb using the bolded word and write 'the first verb that comes to mind and sounds most suitable'. It was emphasized that in most cases the bolded word does not exist in Hebrew and the verb to be formed is also non-existing. Due to the context of the sentence, a verb inflected in the first person singular past tense (represented by the suffix *-ti*) was expected, in any of the seven vowel patterns of the classes, for example appropriate responses for the Paal similarity example in Table 2 can be *SaLaZTI* (Paal), *SILaZTI* (Piel), *HiSLaZTI* (Hifil) and so on. Two examples were presented including possible answers, stressing that more than one answer is possible.

In addition to the 60 experimental items, 6 control items were included (intermixed within the experimental items), presented in the same procedure as the experimental items, but containing existing nouns instead of novel nouns. For example, 'I have a glue (*devek*). With it yesterday I ____'. For these items, there was only one possible answer, namely a real verb (*hidbakti* 'I glued'). The aim of the control items was to prevent participants from using the same verb class for all items and to assure that they have read the sentences presented to them.

Table 2

Example set of Experiment 1

Condition	Sentence 1	Sentence 2
Piel similarity	יש לי **שֶׁדֶס**.	
	*yesh li **shedes**.*	
	there to-me **shedes**	
	‘I have a **SheDeS**.’	
Paal similarity	יש לי **סֶלֶז**	בעזרתו אתמול אני ______
	*yesh li **selez***	*beezrato etmol ani* ______
	there to-me **selez**	With-it yesterday I ___
	‘I have a **SeLeZ**.’	‘Yesterday I ______ with it’
No similarity	יש לי **זֶמֶד**.	
	*yesh li **zemed**.*	
	there to-me **zemed**	
	‘I have a **ZeMeD**.’	

The 66 items (60 experimental and 6 control) were presented in a way that the same phonological similarity condition (Paal, Piel or no similarity) did not appear more than twice in a row. A second version was created, with the reversed order of items, in order to control for a possible influence of the order of presentation. The participants were randomly assigned with one of the two versions.

Data analysis

Responses for each item were coded according to their class (Paal, Piel or Other). Answers that did not belong to any class, contained existing verbs, or were not written in the first person singular past were discarded from further analysis (2.16% of the responses).

The 6 control items were analyzed and it was shown that in all cases participants produced an existing verb. This finding indicates that they have read the sentences and were able to identify existing roots and produce existing verbs out of them.

The data were analyzed with generalized mixed-effects regression (binomial family, logit link function; Jaeger, 2008) with crossed random effects for subjects and items (e.g., Baayen, Davidson, & Bates, 2008). Three logistic mixed-effects models were fit, each with Group (L1, L2), Phonological Similarity (No Similarity, Paal Similarity, Piel Similarity) and their interaction as fixed effects. In addition, Trial order (standardized) was included as a continuous predictor, as well as its interaction with Group. This was done to control for trial-level task effects that can reflect fatigue, boredom, or learning (Baayen & Milin, 2010). The interaction of Trial with Group was included to assess whether Trial has a different influence on the two groups. Possible difference can be obtained due to a general slower processing of L2 speakers and limited cognitive capacity (e.g., McDonald, 2006), which can lead to a stronger fatigue, or on the other hand, could lead to a smaller boredom effect than in L1 speakers. The three models differed in their dependent variable. The first model examined effects of similarity for Piel verbs, and was therefore fit with responses of Piel compared to the rest of the responses (Paal and Other) as the dependent variable. The second model examined effects of similarity for Paal verbs and was therefore fit with responses of Paal compared to the rest of the responses (Piel and Other) as the dependent variable. The third model aimed to compare Paal and Piel responses and was fit with responses of Piel compared to Paal responses (Other responses were removed). Following the recommendation of Matuschek, Kliegl, Vasishth, Baayen, and Bates (2017) random slopes were included if they improved model fit (as measured by AIC). All possible random structures were assessed. The best fit for all models (i.e., the model with the lowest AIC) was achieved by employing an intercept-only random structure.

Results

Table 3 displays proportion rates of responses by class in each condition, for the L1 and L2 speakers. For a summary of the output of the statistical models used in the experiment, see Appendix A.

In order to examine a similarity effect for Piel, that is, whether the proportions of Piel responses will increase when the roots are similar to Piel, an analysis of Piel responses versus the rest of the responses was performed. This analysis did not yield an interaction between similarity (No similarity vs. Piel similarity) and group ($b = 0.1039$, $z = 0.44$, $p = .659$), indicating the groups were not influenced by the similarity in different ways. In order to examine effects across groups, the group factor was centered. The analysis yielded an effect of similarity (see Figure 1); a greater proportion of Piel responses was produced for roots in the Piel similarity condition compared to the no-similarity condition ($b = -0.4010$, z = 2.46, $p = .014$), which contained roots that did not resemble any class. When comparing proportions of Piel responses in the Piel and Paal similarity conditions, no significant difference was found ($b = 0.2410$, z = 1.48, $p = .137$).

Table 3

Percentages of responses by similarity condition and group in Experiment 1

Group	Response	Similarity condition		
		No similarity	Paal similarity	Piel similarity
L1	Paal	48.75	45.00	42.86
	Piel	36.96	37.68	43.39
	Other	14.29	17.32	13.75
L2	Paal	58.04	59.35	53.70
	Piel	27.39	30.87	32.61
	Other	14.57	9.78	13.70

A similar procedure was performed to examine a similarity effect for Paal. The analysis of the Paal responses versus the rest of the responses did not yield an interaction between similarity and group (b = 0.2813, z = 1.30, p = .193), showing no evidence for different group patterns. In contrast to Piel verbs, an effect of similarity for Paal verbs was not found (group factor was centered; Paal similarity vs. no similarity: b = -0.0624, z = -0.42, p = .677); participants did not show an increase in Paal responses when presented with roots that are similar to Paal compared to roots that are not similar to any class. Also when comparing proportions of Paal responses in the Piel and Paal similarity conditions, no significant difference was found (b = 0.2204, z = 1.47, p = .142).

In order to directly compare the proportions of Paal and Piel responses, only Paal and Piel responses were included and similarity was centered. After removing responses of other classes, in the L1 group 53.10% were Paal verbs, and 46.90% were Piel verbs. In the L2 group, 63.37% of the responses were Paal verbs and 36.63% were Piel verbs. A main effect of Group was obtained (b = -1.1735, z = -2.07, p = .038), indicating more Paal responses in L2 than L1 across all conditions, whereas L2 speakers produced significantly more Paal than Piel responses (b = -1.3770, z = -3.22, p = .001). In the L1 group, similar proportions of Paal and Piel responses were obtained (b = -0.2036, z = -0.54, p = .590). The analysis also yielded a significant interaction of Trial and Group (b = 0.0217, z = 4.01, p < .001). While the L2 group produced a trial effect (b = -0.0170, z = -3.96, p < .001), specifically producing more Paal than Piel responses with time, the L1 group did not show any trial effect (b = 0.0048, z = 1.43, p = .152).

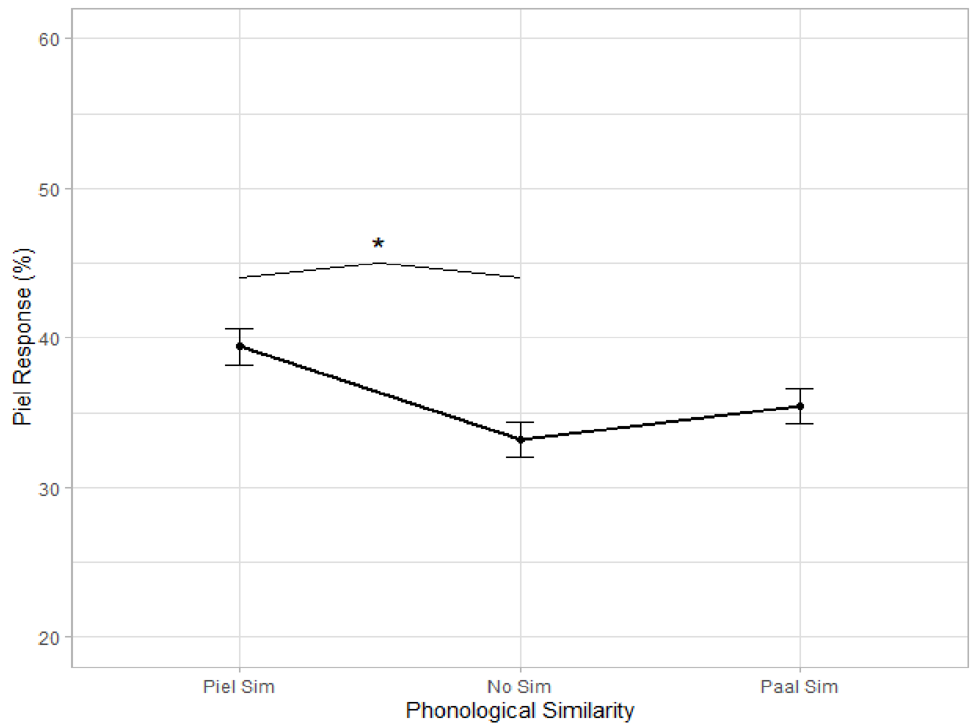

Figure 1. Percentages of Piel responses (out of overall responses) in each similarity condition of Experiment 1. Responses of L1 and L2 speakers are presented together. Piel Sim: Piel Similarity, No Sim: No Similarity, Paal Sim: Paal Similarity.

Overall, the findings show very high production of Paal and Piel novel verbs. The analysis suggests that phonological similarity has a graded effect in generalization of Piel verbs for L1 and L2 speakers (as Piel responses in the Piel similarity condition were larger only in comparison to one out of two conditions), but not in production of Paal verbs. An interesting contrast between L1 and L2 speakers was found in their overall responses of Paal and Piel; unlike L1 speakers, L2 speakers showed a preference to produce Paal over Piel verbs. The next experiment, in a similar design, tested whether argument structure – a source less central cross-linguistically in morphological generalization – is employed in inflectional class generalization in Hebrew.

Experiment 2

Method

Participants

The experiment included a group of 40 L1 speakers of Hebrew (31 females) between the ages of 22 and 35 (mean: 27.00 years), born in Israel and a group of 30 L2 speakers of Hebrew (17 females) between the ages of 22 and 38 (mean: 28.90 years). The L2 group was carefully selected to follow the same criteria as the L2 group in Experiment 1: late learners (mean age of onset: 12.53, SD: 3.50) with L1 Spanish. A subset of the L2 participants has taken part in Experiment 1 (n = 17). However, Experiment 2 took place six months after Experiment 1. The L2 speakers have been long exposed to Hebrew (mean: 16.37 years, SD: 3.36) and achieved a mean score of 97.19% in the YAEL proficiency test, with every participant achieving 80% or higher. As in Experiment 1, the L2 participants reported to use Hebrew on a daily basis (speaking: 64.33%; hearing: 60.83%; writing: 72.83%; reading: 66.16%; total: 62.83%) and to use their L1 Spanish less often, with a mean overall usage of 27.37% (SD: 15.63). All participants had normal or corrected-to-normal vision and none had been diagnosed with any language disorders.

Materials

For the present experiment, 72 novel roots were selected from the same list constructed for Experiment 1. All roots were selected based on identical criteria to the roots in Experiment 1. For consistency purposes, and since phonological similarity was not manipulated in the current experiment, all selected roots had similarity only to the Piel class based on the AML algorithm in Experiment 1. Identical to Experiment 1, all novel roots were inserted into the vowel pattern *CeCeC*, forming novel nouns like *GeDeS*.

Two experimental conditions were created in order to test the effect of argument structure (AS), that is, whether the number and type of arguments in the sentence influences the type of class used to form the verb. The exact AS manipulation focused on the presence of a direct object; either a direct object (+DO) followed the blank space or no direct object (–DO) was displayed to constitute a baseline for comparison (instead, a prepositional phrase was added, to maintain length similarity). In addition, a condition was included to assure that the participants followed the instructions. In this condition a by-phrase followed the blank space. Since only passive verbs can precede a by-phrase, participants were expected to form only verbs of classes that contain passive verbs (Nifal, Pual and Hufal). The nouns in all phrases which followed the expected verb were novel nouns (see Table 4 for an example).

Procedure

Participants performed the task on the web via *SoSci Survey* (Leiner, 2016). There were 72 experimental trials presented for every participant. Every experimental item was presented only once and embedded within a two sentence frame. Identically to Experiment 1, the first sentence presented the novel noun in bold and with diacritic marks and the second sentence had a temporal adverbial that indicated a past event (such as ‘yesterday’), a first person singular subject, and blank space. Unlike the first experiment, the blank space was followed by a direct object (+DO), a prepositional phrase (–DO) or a by-phrase.

Table 4

Experimental design of Experiment 2

Condition	Sentence – part 1	Sentence – part 2
No direct object (–DO)		ליד הגנדסון *leyad hagandason* ‘near the gandason’
Direct object (+DO)	יש לי **גֶדֶס**. בעזרתו אתמול אני ______ *yesh li* ***gedes****. beezrato etmol ani* _____ there to-me **gedes**. With-it yesterday I ___ ‘I have a **GeDeS**. Yesterday I ______ with it’	את הגנדסון *et hagandason* ‘ACC the gandason’
By-phrase		על-ידי הגנדסון *al-yedey hagandason* ‘by the gandason’

The participants were instructed to fill in the blank space with a verb using the bolded word and to type ‘the first verb that comes to mind and sounds most suitable’. It was emphasized that the bolded word does not exist in Hebrew and the verb to be formed is also non-existing. Due to the context of the sentence, a novel verb in the first person singular past tense was expected, but could have been produced in any one of the seven classes. Three examples were presented including possible answers, stressing that more than one answer is possible.

The items were spread in the task so that the same condition (+DO, –DO and a By-phrase) did not appear more than twice in a row. In addition, three versions of the task were created, so that every novel noun appeared only once in every version, each time in a different condition. The three versions were distributed randomly between the participants.

Data analysis

Responses for each item were coded according to their class (Paal, Piel and Other). Answers that did not belong to any class, or were not written in the first person singular past were discarded (0.8%) from further analysis. As the aim for the by-phrase condition was to control for attention of participants, only percentages of the by-phrase condition were computed but the analysis was performed on the +DO and –DO conditions.

The data analysis was kept as similar as possible to the analysis in Experiment 1. The data were analyzed with generalized mixed-effects regression (binomial family, logit link function; Jaeger, 2008) with crossed random effects for subjects and items (e.g., Baayen et al., 2008). Three logistic mixed-effects models were fit, each with Group (L1, L2) and AS (–DO, +DO) as fixed effects. In addition, Trial (standardized) was included as a continuous predictor, as well as its interaction with Group. The first model aimed to examine effects of AS on Piel verbs and was therefore fit with responses of Piel compared to the rest of the responses (Paal and Other) as the dependent variable. The second model aimed to examine effects of AS on Paal verbs and was fit with responses of Paal compared to the rest of the responses (Piel and Other). The third model aimed to compare proportions of Paal and Piel responses and was therefore fit with responses of Piel compared to Paal responses (Other responses were removed). Like in Experiment 1, random slopes were included if they improved model fit (as measured by AIC). All possible random structures were assessed. The best fit for all models was achieved by employing an intercept-only random structure.

Results

Table 5 displays proportion rates of responses by verb class (Paal, Piel and Other) in each condition, for the L2 and L1 groups. For a summary of the output of the statistical models used in the experiment, see Appendix A.

To examine AS effects on Piel verbs, the analysis compared Piel responses versus the rest of responses and yielded a main effect of AS (Group was centered; $b = -1.1158$, $z = -11.97$, $p < .001$); L1 and L2 speakers showed an increase in Piel in the +DO condition compared to the

–DO condition. In addition, a significant interaction between AS and Group was obtained (b = 0.5084, z = 2.67, p = .008), showing that the increase in Piel responses in the +DO condition was larger in L1 (b = 1.3337, z = 11.43, p < .001) than in L2 (b = 0.8253, z = 5.46, p < .001), indicating a stronger AS effect in L1 (Figure 2).

Table 5

Percentages of responses by Argument Structure condition and Group in Experiment 2

Group	Response	- Direct object	+ Direct object	By-phrase
L1	Paal	39.31	36.29	5.77
	Piel	28.88	51.30	2.52
	Other	31.80	12.41	91.72
L2	Paal	49.24	50.00	19.33
	Piel	28.37	38.61	6.95
	Other	22.39	11.39	73.71

To examine AS effects on Paal verbs, the analysis compared Paal responses versus the rest of responses and did not yield a main effect of AS (b = 0.0779, z = 0.93, p = .350); L1 and L2 speakers did not show a difference in Paal responses between the +DO and –DO conditions. An interaction between AS and Group was also not obtained (b = -0.2239, z = -1.33, p = .184), indicating that Paal responses were not influenced by AS.

To compare Paal and Piel responses, an analysis of only Paal and Piel responses yielded a marginal main effect of Group (b = 1.0765, z = 1.82, p = .068); the L2 group produced more Paal responses than Piel responses across conditions (b = -1.0452, z = -2.31, p = .021), but not the L1 group (b = 0.0313, z = 0.08, p = .935), a pattern similar to the one found in Experiment 1. The analysis also yielded a significant interaction of Trial and Group (b = 0.0156, z = 3.03, p = .002). While the L1 group produced a trial effect (b = -0.0164, z = -4.35, p < .001), specifically producing more Paal than Piel verbs over time, the L2 group did not show any trial effect (b = -0.0009, z = -0.19, p = .846).

Generally, the findings suggest a robust effect of AS on inflectional class generalization, with a strong increase in Piel responses when a direct object followed the verb. This effect was stronger in the L1 than the L2 group. In addition, the pattern of responses from Experiment 1 repeated; L2 speakers produced more Paal than Piel responses, a preference that did not occur in the L1 group.

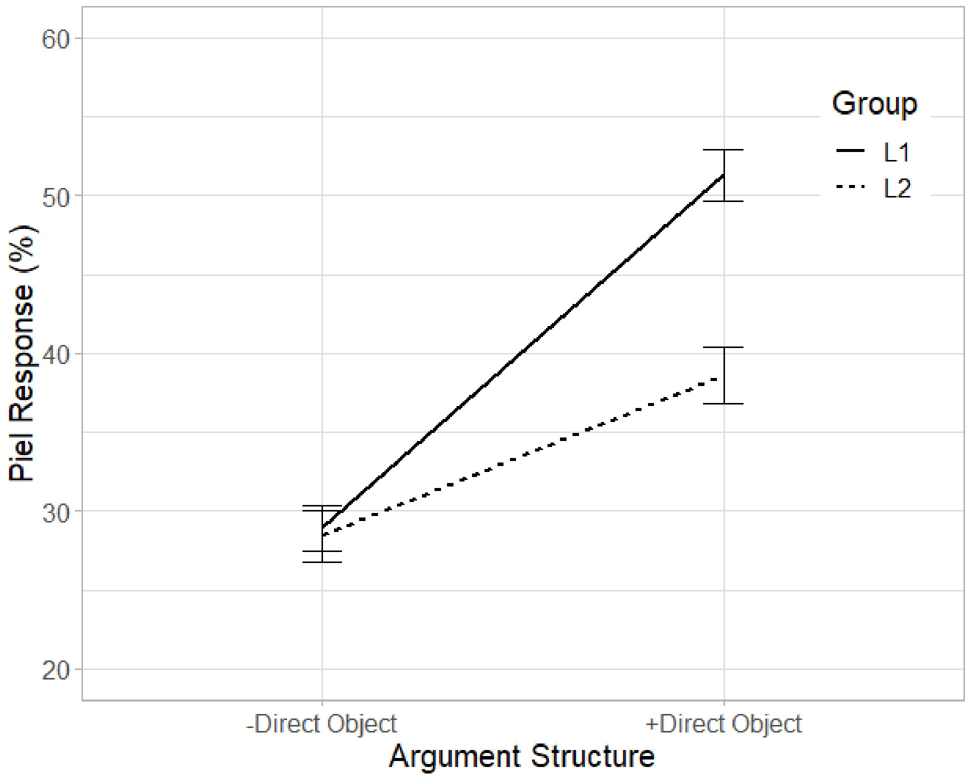

Figure 2. Percentages of Piel responses (out of overall responses) in each condition of Experiment 2 in L1 and L2 speakers.

Discussion

The present work examined what sources of information are employed in generalization of Hebrew inflectional classes in two elicited production experiments. Experiment 1 examined the use of phonological similarity of novel roots to existing roots, a source of information which is considered central in morphological generalization across languages. Experiment 2 examined the use of argument structure, an information source less explored in morphological generalization. Responses in both experiments displayed a very high rate of Paal and Piel verbs. In Experiment 1 Paal responses were not affected by phonological similarity, while Piel responses were, with more Piel responses for roots that were similar to Piel (by the AML) compared to roots that were not similar to any class. Experiment 2 showed a clear effect of argument structure; the presence of a direct object strongly increased the production of novel Piel verbs (but not Paal verbs). In addition, L1 and L2 speakers were tested to examine whether during morphological generalization L2 speakers make use of phonological associations and argument structure information to different extent than L1 speakers. Overall, L2 speakers showed a similar pattern to L1 speakers, with the exception that the argument structure effect for Piel responses was weaker in the L2 group, and that production of Paal was larger than Piel in the L2 group but not the L1 group (significant in Experiment 1 and marginal in Experiment 2). I suggest these findings reflect effects of variability: (i) language-specific effects and (ii) speaker-specific effects.

Language-specific effects on morphological generalization

The present findings showed a graded phonological effect in generalization of Hebrew verb classes. In Experiment 1 Paal responses were similarly high in all similarity conditions. For example, the root *S-L-Z*, analogical to verbs of Paal by the AML, was produced as Paal, *salazti*, to the same extent that the Piel-similar root *Sh-D-S* was produced as Paal, *shadasti*. In contrast to Paal, Piel responses showed a graded effect of phonological similarity; more Piel responses were produced when the roots were similar to Piel (Piel similarity condition), but only in comparison to one out of the two conditions that were not similar to Piel (significant compared to roots from the no-similarity condition, but not the Paal similarity condition). A full phonological effect would have yielded a contrast also between Piel similarity and Paal similarity conditions.

Overall, the role of phonological similarity found here is relatively minor, since it yielded a partial effect for the Piel class and no effect for the Paal class. This can be related to the non-concatenative morphological structure of Semitic languages, in which most words, and all verbs, are formed by a non-linear combination of a consonantal root and a vowel pattern (e.g., root *L-M-D* + pattern *taCCiC* = *talmid* ‘a pupil’). Phonological analogy of concatenative stems overlaps with rhyme between two pronounceable stems (*sing* and *ring* have high similarity and also rhyme; Bybee & Moder, 1983). Analogy of non-concatenative roots, however, cannot be rhyme-based as roots are not pronounceable. Instead, analogy of roots is dependent on matching root consonants, for example, *L-M-D* has high analogy to *X-M-D*, sharing two consonants in the same position. Such rhyme-less analogy is probably less effective in morphological generalization. While previous literature has argued that the non-concatenative structure underlies the processing of complex words, and the way they are organized and accessed in the lexicon (Deutsch, 2016; Velan & Frost, 2011), I propose that the non-concatenative structure of Hebrew can possibly also influence the way they are formed, relying less on phonological information in complex novel word formation.

Furthermore, in comparison to previous findings of morphological generalization in Indo-European languages, specifically inflectional classes (e.g., Eddington, 2002; Say & Clahsen, 2002; Veríssimo & Clahsen, 2014), it appears that the phonological effect here is less robust. For comparison, in an elicited production of Italian inflectional classes (Say & Clahsen, 2002) the findings showed that when a novel verb had no similarity to any inflectional class, a response with second or third class verbs was 4%, but when it had a high similarity to the second or third class, a response with second or third class verbs was between 35% and 38% out of all responses. In comparison to this 30% difference between conditions, the present study yielded a 6% increase at most (for Piel verbs). Nonetheless, the experiments from the different studies might not be directly comparable, thus the observation about the relatively small phonological effect in Hebrew should be carefully considered and be subject for further investigation.

What was shown to be very central in generalization of the Hebrew classes is the argument structure of the verb. The findings showed that participants clearly preferred to produce Piel verbs when the frame included a direct object compared to when it did not (although still relatively frequently produced), but the presence of a direct object did not influence Paal production. This indicates that when forming novel verbs Hebrew speakers are sensitive to the argument structure tendencies of the classes; the majority of existing Piel verbs (around 80%) include a direct object as part of their argument structure, but Paal verbs do not have any tendency in this regard. The present findings are consistent with previous findings about Hebrew verb formation, which found that Hebrew speakers are sensitive to semantic and argument structure information during novel verb formation and class assignment (Berman, 1993; Bolozky, 1999a). While in those studies the contribution of semantic and argument structure information could not have been examined separately, the present work sharpens these previous findings further, showing that argument structure alone (with minimum semantic information) can be employed for generalization of Hebrew verb classes.

The sensitivity of Hebrew speakers to argument structure tendencies of verb classes can be explained by several theoretical frameworks, yet none is completely convincing. One way is within a frequency-based approach, which proposes that frequency and input-based information guides language performance. For argument structure information, it was suggested that language users are able to extract input-based information about the statistical frequencies of argument structure of individual verbs, and use them for prediction during sentence processing (e.g., Traxler, 2005; Wilson & Garnsey, 2009). Along these lines, a frequency-based approach would argue that after receiving ongoing input, Hebrew speakers have acquired knowledge about the statistical frequencies of argument structure of Paal (46%

are transitive) and Piel verbs (79% are transitive, as noted earlier in the introduction), and use them in generalization. Therefore, a frequency-based approach can explain the increase in production of Piel verbs in the context of a direct object, and why Paal responses were not affected by the argument structure manipulation. However, this approach would encounter difficulties explaining the full pattern of the results. If we consider the statistical frequency of Piel verbs from all intransitive verbs in Hebrew, it yields that only 8% of all intransitive verbs in Hebrew are Piel verbs (as opposed to 40% of all transitive verbs). This seems incompatible with the Piel response rates found in Experiment 2; Even in sentences without a direct object Piel verbs were produced in about 30% of the cases in both groups. A frequency-based approach would struggle to explain the large number of Piel responses in contexts without a direct object by terms of statistical frequency alone.

A second way to explain the argument structure sensitivity is via a rule-based approach. Within this framework, Piel stems are formed by a symbolic rule of [root + pattern] (Farhy & Veríssimo, 2019) and are therefore easily extended to novel items, which explains the high productivity of Piel in Modern Hebrew in general (many new verbs are created in the Piel class; Bolozky, 1999a) and the high Piel response rates in this study particularly. This description is similar to the productive first class in Romance languages. However, this explanation alone cannot account for the argument structure effect found for Piel. A rule-based account would probably add that the Piel rule is context-specific as it is syntactically constrained with a specific argument structure specified to Piel, and not to Paal. Still, this approach does not completely align with the present results, and would struggle, similarly to the frequency-based approach, to explain the relatively high rate of Piel responses in contexts without a direct object. This finding indicates a more graded sensitivity to argument structure in the case of Piel, rather than a clear-cut constraint.

Unlike Piel, Paal responses were not affected by the manipulations of both experiments. However, the Paal pattern was relatively surprising in that it presented a high response rate overall (ranged between 36% and 59% in the different conditions and experiments). The Paal class has an unusual combination of characteristics; it is highly frequent (both in type and token) – 19.4% of all verbs (with three consonantal roots) belong to Paal (Piel is second with 17.1%). Paal also constitutes the most phonologically basic form, *CaCaC* (without prefixes or gemination). Paal's vowel /a/ is considered to be a prominent vowel in Hebrew (Bolozky, 1999b). It is the default choice in Hebrew acronym formation (Bolozky, 1999b) and overused by early readers of Hebrew, a script which mostly lacks vowel representations (Leikin, Schwarz, & Share, 2010). Given the high frequency of Paal and its prominent phonological form, one would expect it to be a productive class, however it is non-productive in Modern Hebrew in terms of accepting new members into the class (Aronoff, 1994). Paal's low productivity has been proposed to explain the contrast between Paal and Piel verbs in experimental effects of previous studies (elicited production: Bolozky, 1999a; masked priming: Farhy et al., 2018a; cross-modal priming: Farhy & Veríssimo, 2019). In an elicited production task, Bolozky (1999a) showed that Paal responses were highly uncommon in generating novel verbs, in contrast to the present findings. One possibility is that this discrepancy can be settled if the nature of the experimental task is considered along with certain properties of Paal. In contrast to Bolozky's work, semantic information in the current study was minimized; (i) novel verbs were produced from novel nouns (and not existing ones), and (ii) the objects of the novel verbs were novel nouns. Given that Paal encompasses a large variety of verbs with different meanings without a clear semantic tendency, production of a Paal verb would be appropriate in a minimal semantic context. Furthermore, one reason for the low productivity of Paal is that more than three consonant roots cannot be inserted to its vowel pattern, yet in the present study only three letter roots were selected, eliminating the effect of Paal's phonological constraint.

Speaker-specific effects on morphological generalization

The L2 results had several similarities to the L1 results, along with some differences. Similarities were yielded in the effect of both (i) phonological similarity and (ii) argument structure. First, no evidence was found for a difference in phonological similarity effects between the L1 and L2 groups, which were, as described earlier, rather graded. Second, the L2 group showed increased Piel responses when the sentences contained a direct object (Experiment 2), similarly to the L1 group. This indicates that L2 speakers were sensitive to specific properties of the Hebrew classes, namely argument structure tendencies (which complements findings on L2 sentence processing; Dussias & Scaltz, 2008; Jiang 2004, 2007). It is worth noting that the L2 group in the study had an L1 Spanish, a Romance language with verbal inflectional classes. Previous research has consistently shown that generalization of inflectional classes of Romance languages in L1 speakers strongly relies (at least for certain classes) on phonological associations (Spanish: Brovetto & Ullman, 2005; Italian: Eddington, 2002; Say & Clahsen, 2002; Portuguese: Veríssimo & Clahsen, 2014), and taken together with the current findings, I propose that L2 speakers do not globally rely on the same cues in morphological generalization for their L1 and L2. Rather, L2 speakers rely on information sources which are relevant for the L2 phenomenon. For the present case it means that L2 Hebrew speakers with L1 Spanish were able to generalize Hebrew inflectional classes based on argument structure information, despite the fact that this is not a source that is applied in generalization of Romance inflectional classes. Nonetheless, this proposal should be carefully examined in future research, for example by testing participants in both their L1 and L2.

The analysis detected two differences between the responses of L1 and L2 speakers. First, the effect of argument structure for Piel responses was larger in L1 than L2 speakers. This pattern is similar to Dussias & Scaltz (2008), in which L2 speakers were shown to employ argument structure cues in sentence processing, but to a lesser extent than L1 speakers. One possibility to explain the reduced argument structure effect for L2 in the present study is to assume that properties of L1 can influence the way L2 speakers use information sources at least to some extent, in line with a large body of literature which proposes that some aspects of L2 acquisition are influenced by L1 (known as L1 transfer; e.g., Ionin & Montrul, 2010). Argument structure information is not thought to play a role in generalization of inflectional classes in Spanish (the group's L1). Since it is not directly available from their L1 grammar, perhaps it was not fully acquired yet in all the individuals of the L2 group. Thus, the relatively limited use of argument structure information can be attributed to L1 properties of the L2 group.

Alternatively, the reduced effect of argument structure in L2 can reflect a more general property of L2 generalization. Generally, the finding aligns with sentence processing studies which argue that L2 speakers, even at very high proficiency, show less sensitivity to morpho-syntactic and morphological features than L1 speakers (e.g., Chen, Shu, Liu, Zhao, & Li, 2007; Farhy, Veríssimo, & Clahsen, 2018b; Hopp, 2013; Jiang, 2004, 2007; Krause, Bosch, & Clahsen, 2015). Based on these studies, *the Shallow Structure Hypothesis* (SSH) has argued that in parsing L2 speakers rely less on abstract information and more on surface information, like orthography and phonology (Clahsen & Felser, 2006, 2018). Although the SSH was originally described for L2 processing, the claim of the SSH can be relevant for morphological generalization. In the present study the reduced effect of argument structure in the L2 group can be interpreted as reflecting a reduced ability to make use of abstract syntactic information, specifically argument structure.

The second L1/L2 contrast found in the study was that L2 speakers produced more Paal than Piel responses overall in both experiments, but L1 speakers did not. As noted earlier, both groups produced more Paal verbs than expected given the low productivity of Paal in extending to new verbs in Modern Hebrew, and yet L2 speakers produced even more Paal forms than L1 speakers. I propose this was the result of a combination of two factors. The first is that producing Paal forms simplified the task demands for the L2 participants. It allowed them to rely less on abstract syntactic information – which they have been argued to process less efficiently than L1 speakers (e.g., Krause et al., 2015) – as the Paal class, unlike Piel, does not have any tendency for a specific argument structure. In addition, the Paal stem is the simplest in terms of phonology (as mentioned earlier). Thus, phonological simplicity, along with syntactic simplicity, may explain the particularly high response rate of Paal among the L2 speakers.

The second account for the over-application of Paal is frequency-based. Adult language learners have been proposed to rely more closely on token frequencies in their input to generalize morphological forms (Kam & Newport, 2005, 2009; Schuler et al., 2017). In addition, late L2 learners usually have a relatively small vocabulary size (probably a result of limited input and exposure – the L2 participants reported to have a Hebrew daily usage of 63% and to speak mostly Spanish with their family). Consequentially, words with higher token frequency have more weight in L2 generalization. Since the most frequent verbs in Hebrew belong to the Paal class (in the Hebrew database [Itai &Wintner, 2008] 39 of the 100 most frequent verbs are Paal, followed by Piel verbs with 23 verbs), it is possible that L2 speakers generate more Paal verbs due to its high token-frequency. Previous evidence from English has shown over-application of irregular past tense verbs (Cuskley et al., 2015), which was also proposed to occur due to the high token-frequency of the irregular verbs along with lack of knowledge of less frequent words, often assigned with the regular inflection *-ed*.

Nonetheless, the Paal/Piel contrast between the groups should be considered cautiously, as the analysis yielded that this contrast was dependent on the trial order in both experiments, but with contrasting patterns; in Experiment 1 the L2 group produced more Paal (compared to Piel) with time, while in Experiment 2 it was the L1 group that showed this pattern. The trial effects can indicate a fatigue or boredom of the participants, considering that Paal verbs are phonologically the simplest to produce, as discussed earlier. The question why the L2 group produced a trail effect in Experiment 1 and the L1 produced the same effect in Experiment 2 is beyond the scope of the current study, and can be addressed in future research. It is recommended that future studies will carefully consider web-based experiments' sensitivity to trial and the implications of such sensitivity on the interpretation of the results.

Conclusions

By examining generalization of Hebrew inflectional classes in L1 and L2 speakers, the present study demonstrated how variability in (i) properties of languages and (ii) characteristics of speakers influences the the type of information individuals employ when extending their knowledge to novel forms, specifically phonological similarity and argument structure. With regard to the first source of variability (language properties), Hebrew inflectional class generalization yielded both similarities and differences in comparison to corresponding studies on Romance languages (Say & Clahsen, 2002; Veríssimo & Clahsen, 2014). The results relate to the ongoing debate about whether the specific properties of Semitic morphology lead to substantial differences between Semitic and other languages in the organization of the mental lexicon (e.g., Velan & Frost, 2011), or whether they make use of the same universal architecture like other languages (e.g., Farhy et al., 2018a). The present study suggests that there is no clear-cut answer to this debate, as demonstrated in the complex picture of the current results, involving both similar and different effects between Hebrew and Romance inflectional classes. For better precision, future psycholinguistic theories should take into account both similarities and differences found between Semitic and Indo-European languages in psycholinguistic research. The second source of variability (characteristics of speakers) yielded subtle differences between L1 and L2 generalization, which could reflect different generalization strategies of the two groups, as well as reduced sensitivity to abstract properties, which is often reported in L2 processing studies. Whereas extensive L2 research has focused so far on processing, the present findings support the need for further research on L2 generalization, in order to reach a deeper understanding of how the specific characteristics of L2 speakers (such as age of acquisition and lower exposure) influence the formation of complex words. The findings of the present work emphasize the importance of testing a variety of languages and populations in psycholinguistics, aiming to develop theoretical accounts of language performance that are more precise, but at the same time able to encompass a wide variety of languages and populations.

Acknowledgements

This work was supported by a Ph.D. scholarship awarded by the Minerva Fellowship Program. I thank Harald Clahsen and João Veríssimo for the constant feedback on earlier versions of this paper, and grateful to two anonymous reviewers for helpful comments on an earlier version of this paper. I would like to thank Jonathan Engel for his wonderful work on this paper’s computational modeling and Olivia Taylor for the final proofreading of the manuscript.

Appendices

Appendix A

Logistic regression models employed in Experiment 1 and Experiment 2

EXPERIMENT 1

Fixed effect	Estimate (*b*)	SE	*z*-value	*p*-value
Model 1: Piel vs. Rest				
Reference for Similarity: No Sim				
Intercept	-1.2180	0.2814	-4.33	<.001***
Similarity: Paal (vs. No Sim)	0.1599	0.1631	0.98	.327
Similarity: Piel (vs. No Sim)	0.4010	0.1628	2.46	.014*
Similarity: Paal (vs. No Sim) X Group	0.2214	0.2359	0.94	.348
Similarity: Piel (vs. No Sim) X Group	0.1039	0.2351	0.44	.659
Reference for Similarity: Piel Sim				
Intercept	-0.8171	0.2801	-2.92	.003**
Similarity: Paal (vs. Piel Sim)	-0.2410	0.1623	-1.48	.137
Similarity: Paal (vs. Piel Sim) X Group	0.1175	0.2332	0.50	.614
Model 2: Paal vs. Rest				
Reference for Similarity: No Sim				
Intercept	0.2764	0.2493	1.11	.268
Similarity: Paal (vs. No Sim)	-0.0624	0.1500	-0.42	.677
Similarity: Paal (vs. No Sim) X Group	0.2813	0.2161	1.30	.193
Reference for Similarity: Piel Sim				
Intercept	-0.0064	0.2493	-0.03	.979
Similarity: Paal (vs. Piel Sim)	0.2204	0.1502	1.47	.142
Similarity: Paal (vs. Piel Sim) X Group	0.2173	0.2163	1.00	.315

	Model 3: Paal vs. Piel			
Reference for Group: L1				
Intercept	-0.2036	0.3774	-0.54	.590
Group	-1.1734	0.5655	-2.07	.038*
Trial (in L1)	0.0048	0.0033	1.43	.152
Trial X Group	-0.0217	0.0054	-4.01	<.001***
Reference for Group: L2				
Intercept	-1.3770	0.4272	-3.22	.001**
Trial (in L2)	-0.0170	0.0043	-3.96	<.001***

**p*<.05, ***p*<.01, ****p*<.001

EXPERIMENT 2

Fixed effect	Estimate (*b*)	SE	*z*-value	*p*-value
	Model 1: Piel vs. Rest			
Reference for Group: L1, AS:-DO				
Intercept	-1.3613	0.3337	-4.08	<.001***
AS in L1	1.3337	0.1167	11.43	<.001***
AS X Group	-0.5084	0.1906	-2.67	.008**
Reference for Group: L2, AS:-DO				
Intercept	-1.8760	0.3965	-4.73	<.001***
AS in L2	0.8253	0.1513	5.46	<.001***
	Model 2: Paal vs. Rest			
Group: centered, AS: centered				
Intercept	-0.4488	0.1996	-2.25	.024*
AS	0.0779	0.0834	0.93	.350
AS X Group	-0.2239	0.1687	-1.33	.184
	Model 3: Paal vs. Piel			
Reference for Group: L1, AS: centered				
Intercept	0.0313	0.3840	0.08	.935

Group	-1.0765	0.5900	-1.82	.068
Trial (in L1)	-0.0164	0.0038	-4.35	<.001***
Trial X Group	0.0156	0.0051	3.03	.002**
Reference for Group: L2				
Intercept	-1.0452	0.4517	-2.31	.021*
Trial L2	-0.0009	0.0044	-0.19	.846

**p*<.05, ***p*<.01, ****p*<.001

7 General Discussion

7.1 Summary of Results

The present work addresses the question how language-specific morphological properties and speaker-specific characteristics shape processing and generalization processes of complex forms. We focused on investigating the role of (i) Hebrew morphological properties and (ii) language 'nativeness' (i.e., being an L2 speaker as compared to an L1 speaker) in processing and generalization of productive and unproductive verbal inflectional classes. Hebrew verb lexemes (or lexical entries) are formed by a non-concatenative combination of a consonantal root and a vowel pattern, one out of seven vowel patterns possible for verbs. Each pattern is unique for a certain inflectional class. For example, the Paal verb lexeme *lamad* 'learn' is formed by the root *L-M-D* and the pattern *CaCaC*, while the Piel verb lexeme *limed* 'teach' is formed by the root *L-M-D* and the pattern *CiCeC*. Despite the seemingly identical structure and similar type frequency, the classes contrast in productivity; whereas new verbs are often created in Piel, they are hardly ever created in Paal.

The current theoretical accounts of morphology and lexical access are strongly guided by the knowledge accumulated about Indo-European languages, especially English. Indeed, the detailed research conducted on the English past tense, for example, yielded important insights about the processing of regular and irregular inflectional morphology, arguably reflecting universal properties of the human language system. However, it raises the question of whether universal theories of morphological processing and organization can be developed based on English morphology, which is considered to be relatively poor (Blevins, 2006), 'an extreme case' with idiosyncrasies that perhaps should not be taken as a template for the basic properties of a universal theoretical framework (Orsolini & Marslen-Wilson, 1997). Instead, if we wish to form a theory of language processing that captures the universal properties of a language, and at the same time acknowledges the idiosyncratic characteristics of different languages, we need to examine typologically different languages (Boudelaa & Marslen-Wilson, 2015). Via cross-linguistic testing, we can reach a better understanding of how the specific constraints of an individual language shape the types of solutions that are adopted to optimize the mapping between form and meaning (Frost et al., 2008). The present thesis aims to address this gap, in order to contribute to the development of theories about the mental lexicon.

In addition, current theories should be able to account for speakers with different individual characteristics, such as late age of acquisition. Late acquisition of a second language usually goes hand in hand with differences in the patterns of language input, exposure and means of acquisition, compared to L1 speakers. Sampling mainly L1 speakers might yield precise results with less variance but captures a narrow subset of the speakers of today's global world, who are often multilingual, proficient in more than one language.

The data presented in this dissertation extend existing psycholinguistic research on morphological processing and generalization by capturing a vaster degree of variance in morphological properties of different languages and variance in characteristics of speakers by testing a less explored language and a speaker group:

1. Hebrew
2. L2 speakers (compared with L1 speakers)

Table 3 summarizes the findings obtained in all four publications of the dissertation. The table shows the main findings in each experimental technique (masked priming, cross-modal priming and elicited production) for each group (L1 and L2 speakers) and for each verbal class (Paal and Piel). In summary, the findings of this dissertation revealed several dissociations in morphological effects between Paal and Piel classes and L1 and L2 speakers, both in word recognition and generalization of classes to novel verbs. L1 speakers show (i) a lack of root-priming effect in masked priming for Paal but a significant root-priming effect for Piel, (ii) a semantic transparency effect for Paal verbs in cross-modal priming, but not for Piel verbs, (iii) a robust effect of argument structure in generalization of Piel verbs to novel items, but not Paal verbs. L2 speakers showed a partly different pattern than the L1 speakers in both word recognition and generalization: (i) a similar root-priming effect for both Paal and Piel but only in the infinitive form in masked priming, and (ii) an effect of argument structure in generalization of Piel verbs to novel items, but a less robust effect than the L1 speakers demonstrated.

The present chapter discusses the results in two parts. First, the results of the L1 speakers will be discussed in comparison to results from the current literature about Indo-European languages, as part of the ongoing debate as to what extent the specific morphological properties of Hebrew modify the way complex words are accessed and represented in the mental lexicon. Second, the results of the L2 group will be discussed in a direct comparison to the L1 pattern, within current theoretical L2 approaches in an attempt to understand the nature of the difference in processing and generalization properties between L2 and L1 speakers.

Table 3

Summary of results across the four publications

	Native speakers		Non-native speakers	
	Paal	Piel	Paal	Piel
Processing				
Masked priming: root priming effect	-	+	+/- (only in infinitive forms)	+/- (only in infinitive forms)
Cross-modal priming: semantic transparency effect	+	-	not tested	not tested
Generalization				
Elicited production: phonological similarity effect	-	+/-	-	+/-
Elicited production: argument structure effect	-	++ (larger than in L2 speakers)	-	+

Note: the sign '+' denotes an effect was detected, and '-' denotes an effect was not detected. The sign +/- shows the effect was partial.

7.2 The Role of Hebrew-Specific Morphology in Processing and Generalization

7.2.1 Morphological processing

Morphological processing was examined in the present work by using two experimental paradigms: masked priming and cross-modal priming. We presuppose that these different priming techniques tap into distinct levels of word recognition processes, allowing us to look at different angles of morphological representations. This notion has been advanced in priming studies in the last 20 years (Marslen-Wilson, 2007, but was also criticized; see, for example, Feldman, Milin, Cho, Moscoso del Prado Martín, & O'Connor, 2015). As described in the General Introduction, masked priming has been claimed to tap into an initial step of accessing a word (the access level), in which words are decomposed into constituents, blind to semantic compositionality. The cross-modal priming technique is thought to tap into a later stage of recognition processes, a level including the central representations of the lexical entries, consisting of morpho-semantic representations, where both form and meaning are processed (Marslen-Wilson, 2007). One of the central arguments for this hypothesis is that morphologically related prime-target pairs with opaque semantic relation between them tend to display a significant priming effect in masked priming but not in overt or cross-modal priming (e.g., Longtin et al., 2003; Rastle et al., 2000). Nonetheless, in Semitic morphological studies, such contrast has not been shown (Boudelaa & Marslen-Wilson, 2005, 2015; Frost et al., 1997, 2000). The predominant claim in the Semitic literature (described in Boudelaa & Marslen-Wilson, 2015, for example) has been that due to their rich non-concatenative nature of morphology, complex words in Semitic languages are always processed via extraction of their consonantal roots, regardless of semantic compositionality (also in cross-modal priming).

However, our results seem to question this Semitic-specific claim. The results in Publications I and III indicate that despite the relatively unique non-concatenative structure of Hebrew (and Semitic languages in general), Hebrew recognition of complex words shows similar patterns to those that have been reported in Indo-European languages: (i) decomposition of constituents in the case of structured representations, while unstructured representations are accessed as wholes, and (ii) central morphological representations are related to semantic properties. More specifically, under masked priming conditions (Publication I), L1 speakers extracted the root from productive but not unproductive verbal classes. This was detected regardless of the type of the inflected form of the prime (infinitive or first-person singular past [1sg Past]). Under cross-modal priming conditions (Publication III), it was found that the degree of semantic relatedness between primes and targets affected RTs of targets preceded by Paal primes (but not Piel).

Both results are only partially consistent with previous priming studies on morphology in Semitic languages. Priming studies on Hebrew and Arabic have constantly shown that primes and targets that share a regular root (i.e., three root consonants which are all displayed in the primes and targets) yield a clear root-priming effect. The present work, however, suggests that root priming is more constrained than previously considered. Paal and Piel primes were directly compared by preceding the same Hitpael target (Paal condition: *LiLMOD–HiTLaMeD*; Piel condition: *LeLaMeD–HiTLaMeD*). Paal primes failed to display a significant root-priming effect in both form conditions (infinitives and 1sg Past), in sharp contrast to Piel primes, showing a significant root-priming effect in both forms. Importantly, target recognition facilitation from Paal primes was significantly reduced compared to target facilitation of Piel primes.

In addition, the semantic transparency effect found for Paal primes is inconsistent with a large body of cross-modal priming studies, mostly in Arabic, that failed to find an effect of semantic transparency in prime-target pairs that share a root. All cross-modal findings from Arabic and Hebrew reported that both transparent and opaque prime-target pairs show a significant morphological priming effect (e.g., Boudelaa & Marslen-Wilson, 2015; Frost et al., 2000), unlike the typical findings from Indo-European languages, in which the opaque pairs fail to show a significant priming effect. The Hebrew cross-modal results (Frost et al., 2000) differed slightly from the Arabic results, in that a larger priming effect was detected for transparent than opaque pairs. Those studies examined semantic transparency as a categorical variable, and therefore to deepen the understanding of this topic this work tested Semitic semantic transparency as a continuous variable, which is a more genuine reflection and less artificial measure of semantic overlap.

How can the unusual pattern in Publications I and III be explained by the current accounts of the complex word processing in Semitic language? In the following paragraphs, I will discuss how a root-based approach, a stem-based approach and a distributed approach would explain the current findings, and argue that they are able to do so only partially.

First, consider the most common approach for Semitic morphology, the root-based approach, which postulates that the Semitic lexicon is organized by different principles compared to the Indo-European ones, relying more closely on morphological principles than semantic and orthographic ones (e.g., Deutsch, 2016). The root-based approach claims that the Semitic mental lexicon is organized by consonantal roots. During lexical recognition, *all* root-based derived words are immediately decomposed into roots and patterns in what is referred to as a 'morphological unit' (Deutsch & Meir, 2011), 'morphological level' (Deutsch et al., 1998), 'word-form level' (Deutsch, 2016) or 'access level' (Boudelaa & Marslen-wilson, 2004), where roots and patterns are represented separately. This level is claimed to be purely based on morphological form, independent of semantic and syntactic properties and not restricted to orthographic properties. Therefore, this level is not a morpho-orthographic level nor a morpho-semantic level, terms that have been proposed in word recognition accounts based on Indo-European languages (Crepaldi, Rastle, Coltheart, & Nickels, 2010; Diependaele, Sandra, & Grainger, 2009; Rastle et al., 2004; Xu & Taft, 2015).

The root-based approach, however, can explain the present findings only partially. The pattern of results of the productive class (Piel) fits the root-based approach: (i) the root of Piel was accessed from Piel primes, as revealed by the root-priming effect from Piel primes (Publication I) and (ii) the access to the morphological structure was independent from semantic properties (Publication III). The pattern of Paal, however, challenges the root-based approach, as it showed an opposite pattern to Piel: (i) the root of Paal was not accessed from Paal primes under masked priming conditions (Publication I) and (ii) the access to the morphological structure of Paal, shown by a significant root-priming effect, was dependent on semantic properties (Publication III). If, as proposed by the root-based approach, a purely morphological level exists where all root-based forms are decomposed into roots and patterns, Paal verbs had been expected to show a similar root-priming effect to Piel verbs, without showing a semantic overlap effect. To sum, the root-based approach, which is the predominant view on Semitic lexical access, would struggle to account for the Paal pattern within its current theoretical assumptions.

An alternative account of the organization of the Hebrew mental lexicon is the stem-based approach, which emphasizes the similarities between the Semitic and Indo-European mental lexicon. The stem-based approach has emerged from a theoretical framework for Semitic morphology, claiming that unique constituents that are specific for Hebrew and Arabic are not required, and that the Semitic morphology with its relatively unusual patterns can be explained within current universal frameworks which encode stems (Bat-El, 1994; Ussishkin, 2005). There is only limited psycholinguistic evidence that supports this approach (Berent et al., 2007), yet in fact previous findings that were claimed to support the root-based approach can be explained by a stem-based approach as well (as discussed in the General Introduction).

However, also the stem-based approach can only partially account for the pattern of the current results. According to this view, the root is not extracted during lexical recognition, which is in line with the absence of the root-priming effect for Paal. If we assume that the prime *lilmod* ‘to learn’ and the target *hitlamed* ‘to intern’ do not share a mental representation of a root *L-M-D*, but are accessed via whole stems, *lmod* (allomorph of *lamad*) and *hitlamed*, then a priming effect would not occur, as they have two distinct stems. The root-priming effect for Paal in the cross-modal study can have more than one source. Since semantic transparency in the Paal condition was found to have a significant effect on RTs, the root-priming effect can be the result of a graded overlap of form and meaning between primes and targets (e.g., Gonnerman et al., 2007). Alternatively, the stem-based approach can explain the effect as the result of an activation of morphologically related stem-based representations following the presentation of the prime (Berent et al., 2007).

What the stem-based approach would struggle to explain is the pattern of Piel primes and particularly the contrast between Paal and Piel. If we assume that stems are the minimal constituent in Hebrew, then Piel and Hitpael stems do not share a morphological constituent. If so, then what is the source of the priming found for Piel under masked priming and cross-modal conditions? One can refer to an overlap in meaning and form between prime and target as the source that has elicited the effect, potentially driven by activation of associative links between words related in form and meaning. This explanation is unlikely in both studies. In the masked priming, the same meaning and form overlap existed also between Paal primes and Hitpael targets, but a priming effect was not found, and was significantly reduced compared to the Piel facilitation. In the cross-modal priming, semantic transparency did not affect the Piel condition; thus, it appears that the source of the Piel priming effect is not related to semantic properties.

Taking the two approaches together, we can conclude that both are able to provide a rather partial explanation to the current findings. While the root-based approach explains the Piel pattern but not the Paal pattern, the stem-based approach is able to explain the Paal pattern but not the Piel one. It is possible that other frameworks, that do not postulate a specific morphological level of representation, nor decomposition of constituents, can explain the present results better. Consider, for example, the distributed approach, which postulates that morphology reflects a learned sensitivity to the systematic relationship between the word form and its meaning, and thus morphological priming effects in general, and root-priming effect in particular, arise from a graded overlap of form and meaning between the prime and target (Gonnerman et al., 2007). Plaut and Gonnerman (2000) argued that distributed connectionist models can account also for the patterns found in complex word recognition of Semitic languages, showing that in morphologically rich languages (defined by them as languages with a structure that almost always enables a decomposition of forms to separate constituents, which participate in the formation of many words) morphological priming can be simulated in the absence of semantic overlap and at the same time still be affected by semantic overlap. Thus, their model is consistent with the semantic effect found for Paal. Yet, this account would still struggle to explain the full pattern of results, especially the contrast found between Paal and Piel in Publications I and III. Paal and Piel conditions were matched in form overlap and semantic relatedness to the targets. Nonetheless, they displayed a clear asymmetry in priming effects in Publication I and an asymmetry in semantic transparency effects in Publication III.

What all the above approaches do not take into account is the *contrasting degree of productivity* between Paal and Piel classes. Previous psycholinguistic evidence has shown a dissociation in patterns of productive and unproductive morphology, a significant part of them in priming studies (e.g., Sonnenstuhl et al., 1999). Some evidence for productivity effects was found in Hebrew (e.g., Velan & Frost, 2011, who briefly noted the possibility of whole-stem access to forms with unproductive roots), but so far, the role of productivity in Hebrew morphology has not been thoroughly examined and discussed. The central role of productivity in Arabic morphological processing was recently demonstrated in a lexical decision study of Arabic verbal classes (Wray, 2016). While the RTs for recognition of verbs from a productive class were affected by both the base frequency and the surface frequency of the verb, the RTs for recognition of verbs from unproductive classes were affected only by the surface frequency of the verbs. This pattern was interpreted as reflecting a contrast in the way processing occurs in productive and unproductive morphophonological patterns; a full decomposition pathway (root and pattern in the case of Semitic languages) to productive forms and a direct access as whole stems for the unproductive forms. This interpretation can easily be implemented for the present findings, which include two types of contrasts between productive (Piel) and unproductive (Paal) classes: one in root priming in Publication I and the other in semantic transparency in Publication III. Our results strengthen the need for a deeper understanding of the role productivity plays in recognition processes of Hebrew complex words.

Addressing the role of productivity in complex word recognition, we propose a dual-route access to Hebrew stems and a balance between decomposable and undecomposable stems, in a similar architecture that has been proposed in the dual-route approach for Indo-European languages, which also includes two levels of representation: an access level and a central level. The results can be best explained if we postulate that Paal stems and Piel stems have different types of mental representations, structured for Piel and unstructured for Paal, both at the access level (tapped by the masked priming technique) and central level (tapped by the cross-modal technique). For a visual illustration of the proposed model, see Figure 2. According to our proposed model, at the access level Piel stems are decomposed into roots and patterns (e.g., *limed* ‘(he) taught’ in Figure 2), regardless of semantic properties, in a similar manner to decomposition in root-based accounts. After the decomposition, the consonantal root is activated. The recognition of the Hitpael target also follows a decomposition at the access level, and reactivates the same root displayed at the Piel prime. Since the same constituent (the root) was already activated following the prime, activation of the root in Hitpael is faster, leading to a faster recognition of the Hitpael targets.

We propose that Paal stems, however, are represented as whole stems at the access level (e.g., *lamad* in Figure 2). It is important to note that we presuppose that the access level in this model does not only include morpho-orthographic properties as sometimes described in the Indo-European literature (Crepaldi et al., 2010; Diependaele et al., 2009; Rastle et al., 2004; Xu & Taft, 2015), but phonological properties as well, as commonly assumed in Semitic models of lexical processing (described earlier). This is crucial due to the very opaque orthographic script of Hebrew, which includes very partial representation of vowels. Therefore, our assumption is that all access representations include the vowels that are not necessarily represented orthographically (e.g., the word *sevel* 'suffering' is written without vowels: *SVL*, but its access representation would include the vowels: *sevel*). Thus, despite the fact that the Paal stem is orthographically identical to the root (e.g., both are orthographically represented as *LMD*), we assume that their access representations are different (Paal stem: *lamad* and root: *L-M-D*). Root priming is not yielded for Paal primes because Paal stems are accessed directly as whole stems; thus, the whole stem is activated (*lamad*) and not the separate constituents (root *L-M-D* and pattern *CaCaC*). When extraction and activation of the consonantal root does not occur, the root in the Hitpael target is activated for the first time, and thus target facilitation based on shared root activation is not possible.

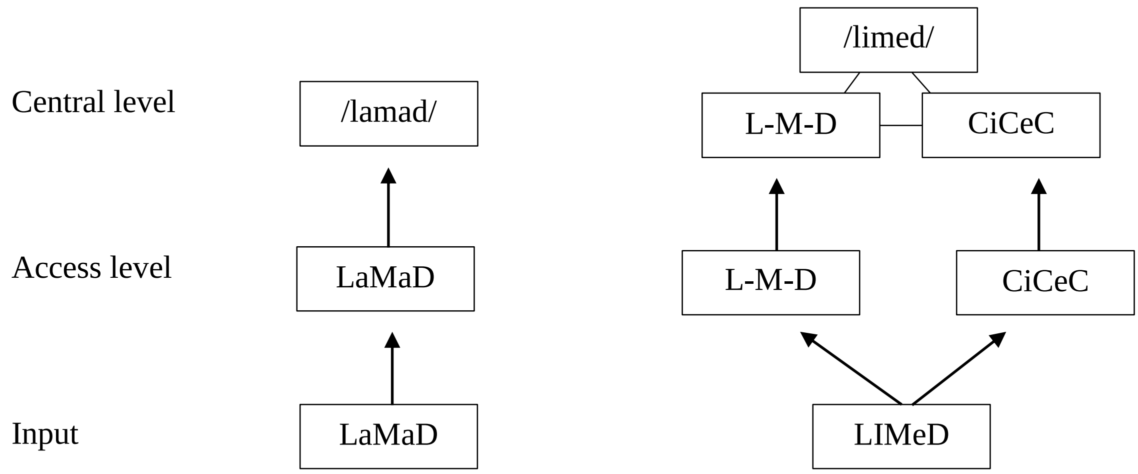

Figure 2. Our proposed model of the representation and processing of verbal stems in Hebrew. On the left is the example of how Paal stems are represented and accessed (demonstrating the representation of the lexical entry *lamad* '(he) learned'); on the right is an example for the representation of Piel stems, which is based on a combinatorial representation of a root and a pattern (demonstrating the representation of the lexical entry *limed* '(he) taught').

Further support for our model and the representation of Paal stems as unstructured arises from a recent masked priming study with MEG (Kastner et al., 2018), which tested root priming of Paal and Hifil (another verbal class) as targets, and Paal and Hifil vowel pattern priming. Interestingly, the behavioral results failed to yield any facilitation effect, perhaps related to the very short SOA (33 ms), which has not been applied in Hebrew priming studies so far. The MEG analysis also failed to find a root-priming effect with Paal verbs but found an effect with Hifil. This finding is in line with the finding of Publication I and may imply that roots are not extracted from Paal at the access level, arguably due to its unstructured representation. Interestingly, however, the MEG results showed that the vowel pattern priming was significant (e.g., *SHaTaF–DaXaF*), which indicates L1 Hebrew speakers were able to extract the vowel pattern *CaCaC* of a visual presentation of Paal, even though the vowel /a/ was not visually represented at all. How and why a pattern-priming effect but not a root-priming effect was found for Paal remained an open question in this study, which demonstrates that despite the relatively large body of research on Hebrew morphology, many questions are still open.

At the central level, our model is less detailed. This is in part because morphological processing evidence in Hebrew has strongly focused on the masked priming paradigm, and therefore the findings are often dependent on visual processes and are constrained to the initial access-level processes. Research on Arabic contributed more findings of cross-modal studies that are not confined to the visual domain. Yet, a detailed model of lexical recognition in Arabic also has not been presented yet, arguably due to insufficient data from online fine-grained methods (Boudelaa & Marslen-Wilson, 2015). Most Hebrew accounts of lexical access have postulated only one level of morphological representations, organized around networks of roots and patterns, predicting the same patterns for various experimental paradigms. Perhaps this is partly why other paradigms, which tap into later stages of processing and non-visual domains, were not given much attention. However, in recent years the picture-word interference task with auditory and visual distractors has been implemented (Deutsch, 2016; Deutsch & Meir, 2011; Kolan et al., 2011), providing more depth to the research of morphological processing in Hebrew by tapping into the conceptual level, central level and production processes. So far, the results are relatively parallel to previous masked priming studies. Nonetheless, in a picture-word interference task, Kolan et al. (2011) found a facilitation effect for a vowel pattern distractor, but only when it had a semantic contribution, unlike the root which always had a facilitatory effect. This finding has led the authors to propose a word retrieval model, where morphology is employed in two levels of word recognition, which are essentially parallel to the access level and the central level described in the present dissertation (in their model, they make use of the terms ‘lexeme level’ and ‘lemma level’, based on the model of word retrieval of Levelt, Roelofs, & Meyer, 1999). According to the proposal of Kolan et al., the root is represented at the access level and the vowel pattern at the central level. Although the authors did not postulate a root representation at the central level, their model provides the first step towards establishing that a morpho-semantic level is essential also in the Semitic mental lexicon.

The semantic transparency effect found for Paal in Publication III, along with the semantic transparency effect for roots found for Hebrew in the cross-modal study of Frost et al. (2000; where an opaque semantic relation elicited a priming effect but to a lesser extent than transparent pairs) and the semantic influence in the vowel pattern effect found in the picture-word interference study of Kolan et al. (2011) require us to reconsider whether semantics and morphology are truly independent from each other in the Semitic lexicon, or whether, at least for a subset of the lexicon, they are connected to each other. I suggest that a morpho-semantic level of representation is required also for Semitic languages (see central level in Figure 2), where each Paal and Piel verb constitutes a lexical entry in which information about the morphology, subcategorization, semantics, form and inflectional paradigm of the entry is specified. I posit that also at that level the morphological representations of Piel stems are structured. Therefore, at the cross-modal priming, after the auditory prime had been heard, it was decomposed (*L-M-D* + *CiCeC*) and its root was activated at the central level. The recognition of the target also involved decomposition and its root was reactivated at the central level, leading to a target recognition facilitation. An alternative explanation of the effect as stemming from an overlap of semantic and form is relatively implausible, since the root-priming effect was elicited regardless of semantic transparency, which strongly suggests that the source of the effect was morphological and cannot be fully explained by an overlap of form and meaning.

The exact type of morpho-semantic representation of Paal stems (structured/unstructured) is less clear. What is apparent, though, is that the semantic transparency effect found for Paal is typical to findings from overt priming and cross-modal priming in Indo-European languages (Feldman & Soltano, 1999; Longtin et al., 2003; Marslen-Wilson et al., 1994; Rastle et al., 2000; Reid & Marslen-Wilson, 2003; Rueckl et al., 2008). This effect is often interpreted as reflecting a dependency of morphological representation on semantic properties. Therefore, it appears that the central level of the Semitic lexicon is more similar to Indo-European languages than previously considered. The semantic transparency effect in Indo-European languages was interpreted as an indication that at the central level only complex words with transparent semantic relation share a constituent (*sad–sadness*), but words like *department–depart* do not share the same lexical stem representation – *depart* (*department* is stored as a whole stem and not a structured stem); presumably, this is why they do not show a priming effect under cross-modal conditions (Marslen-Wilson, 2007). If such rational is applied for the present findings, that would imply that some Paal stems are structured at the central level and some are not, depending on the semantic compositionality of the root and vowel pattern. However, implementing the same explanation from Indo-European to Hebrew lexical entries is not straightforward. Derivation processes in Hebrew and a language like English are highly different. In English, a base stem like *depart* also has its own lexical entry with specified meaning and syntactic properties, and therefore an evaluation of semantic relatedness between *depart* and *department* is possible, while a root like *L-M-D* is not a lexical entry, and is assumed, at least by most theories of Modern Hebrew, to not carry an invariant meaning (for review: Aronoff, 2013). Thus, one cannot assess the semantic compositionality of *L-M-D* and *lamad*, and definitely not make a dichotomous distinction based on this term. Under this framework, it is not clear which Paal stems are expected to be unstructured and which – structured.

Alternatively, a distributed approach would claim that Paal stems have distributed representations rather than discrete constituents and thus do not follow the clear-cut distinction of being structured or unstructured. Within this approach, Paal verbs would be represented via patterns of activity across different processing units that reflect meaning, sound and spelling and the connections between them. Morphological knowledge is stored in the weights on those connections rather than via explicit levels of representations (Gonnerman et al., 2007). This type of framework also proposed that priming effects reflect graded overlap of form and meaning between words, which seems to be compatible with the effects for Paal: the priming in publication III can be explained by graded effects of semantic and form overlap. This might fit Paal stems, but as discussed before, it is less suitable for the findings from the Piel condition. It is not likely that a distributed mechanism will be applied to a language only partially, and therefore also this explanation is not satisfying.

Perhaps the most probable possibility is that all Paal stems are unstructured at the central level, parallel to the Paal representation at the access level. The source of root-priming effect of Paal under cross-modal conditions will then be attributed to semantic properties and not decomposition to morphological constituents. The proposal that Paal stems have unstructured representations is also in line with some of Paal's traits. Other than low productivity, which has been thoroughly discussed throughout the dissertation, another trait of Paal, which we did not discuss so far, is its vowel pattern ambiguity/homonymy; that is, the vowel pattern serves more than one function. The Paal pattern *CaCaC* does not only denote a verb but can also denote adjectives (e.g., *katan* 'small') and nouns (*xalav* 'milk'; Kastner et al., 2018). In that sense, it is similar to affixes like *-er* in English that appear both in a derived word like *teacher* and a comparative form like *smaller*. It has been suggested that homonymous and ambiguous affixes are less likely to be decomposed and structured than words with unambiguous affixes (based on base and surface frequency effects in lexical decision tasks, Bertram et al., 1999, 2000;

Sereno & Jongman, 1997). It was explained that such affixes make the recognition process more complex, since they require the individual to consider two sets of grammatically different possibilities, and thus, they trigger whole-form storage. Our proposal that Paal stems are unstructured and therefore do not decompose fits quite well with this account. The recognition process of forms with the vowel pattern *CaCaC* is more complex and requires considering the various grammatical possibilities of *CaCaC*, and thus it gives rise to the storage of Paal verbs as whole stems rather than structured by a root and the ambiguous vowel pattern *CaCaC*.

7.2.2 Morphological generalization

The second morphological aspect that was examined regarding Hebrew-specific effects is generalization of novel verbs in an elicited production task (Publication IV). The main findings here were (i) a weak effect of phonological analogy (which will be referred to here also as phonological similarity), (ii) a significant effect of argument structure and (iii) a high response frequency for Paal and Piel (i.e., the majority of responses across conditions in both experiments of Publication IV were the Paal and Piel verbs). The present section discusses how these results contribute to understanding the role of universality and variability in the mental process of language. I will argue that the pattern of results in Publication IV generally suggests that the generalization properties of Hebrew inflectional classes are more Hebrew-specific, since they rely strongly on argument structure information rather than phonological analogy, in contrast to inflectional classes of languages such as Italian and Portuguese.

First, we consider the effect of phonological analogy, which was not a strong predictor of class generalization, present to some extent in Piel but not at all for Paal (the unproductive class). This pattern differs from the typical patterns in other languages, which consistently show analogy effects in the productive morphological patterns (e.g., English past tense: Prasada & Pinker, 1993). Novel roots that according to the AML (the analogy-based model; Eddington, 2000, 2002) were similar to Paal (Paal similarity condition) did not yield more Paal responses compared to other roots that were not similar to Paal (Piel similarity condition and no-similarity condition). In the present work, an indication for the influence of phonological similarity effect was found to some extent for the productive Piel class, but this effect was rather weak. Roots that were similar to Piel (by the AML) yielded more Piel responses compared to roots that were not similar to any class. Yet, Piel responses in the Piel similarity condition were not different compared to Piel responses in the Paal similarity condition, although they are analogically not similar to Piel. If Piel was generalized by similarity, we would have expected to also find a response contrast between the Piel similarity condition and the Paal similarity condition. In addition, the Piel effect was basically a 6% increase in responses between the Piel condition and the no-similarity condition. When observing parallel findings from Indo-European languages, such as Italian inflectional classes, we detect a larger phonological effect – a 30% increase in responses (Say & Clahsen, 2002). In summary, we find some indication of a phonological similarity effect, but a relatively weak one and not in the unproductive class that was predicted to show an effect of analogy, but in the productive class which in previous literature has shown mixed results.

However, this conclusion should be addressed with caution, since the present findings may strongly relate to the specific algorithm used and the way the AML model (Eddington, 2000, 2002) defines similarity, and not necessarily reflect the overall use of phonological context in general. The AML is an exemplar-based model; namely, it makes use of a database of exemplars, which assumes to represent the contents stored in the mental lexicon, and uses it to predict linguistic behavior. In the AML, similarity depends on how subcontexts from the database are determined, and how the analogs are selected from the subcontexts. A possible limitation of the AML framework is that a lexical item might not be selected as a possible analog in the case of disagreements between members of the same subcontext. This indicates that when phonological segments occur in more than one binyan, their influence will be eliminated. Since in Hebrew the same consonantal root can appear in more than one binyan, this poses a serious problem for the algorithm, potentially losing a large amount of information. Other models, despite similarities to the AML, operate in a different way and define similarity differently. Thus, although models such as the connectionist networks also assumes that generalization relies on analogy between novel and stored forms and not on context-free rules, their algorithm is different; it relies on the connection weights between distributed units of form and meaning, and defines similarity as the overlap between them. Such models could possibly lead to a different output and align better with the human data found in Publication IV.

Still, if we assume the model reflects analogy-based generalization, one can speculate about the reasons for the weakness of the phonological effects in generalization of Hebrew classes, which can be attributed to two sources related to specific properties of Hebrew: the non-concatenative structure of Hebrew morphology and the derivational properties of the binyanim. First, phonological analogy of non-concatenative consonants is not comparable to analogy of concatenative clusters of vowels and consonants. Rhyme-based associations are a central part of similarity-based generalization (Bybee & Moder, 1983), but such associations require pronounceable phonological clusters, which is not the case when dealing with non-concatenative constituents. The AML model (Eddington, 2000, 2002), employed in Publication

IV, simulated subcontexts of similarity, expressed by overlap of root consonants without vowels. The model revealed that the subcontext *_-TS-R* appears in Paal and therefore predicts that novel roots that overlap with this subcontext (e.g., *D-TS-R*) would also appear in Paal (*datsar*). This kind of overlap is based on 'consonant matching' and cannot be pronounced; thus, of course, the overlapping parts cannot rhyme. This is conceptually very different from analogy-based generalizations that have been examined so far in Indo-European languages, and perhaps can explain to some extent the weakness of the analogical effect we found. The second source for the reduced phonological effect is the derivational properties that are also involved in formation of the binyanim. Recall that the root and pattern combination is a derivational process, whose output creates a lexical entry. The same root can derive many lexical entries when combining with different vowel patterns, including more than one verbal pattern. Consequentially, classes are not built around typical phonological clusters, as opposed to the English irregular past tense (for example, the clusters -ing/-ang). This is generally in line with previous literature, which has mostly focused on phonological-based generalization in inflected forms rather than derived forms (e.g., Prasada & Pinker, 1993). Alegre & Gordon (1999a) did report phonological-based generalization in derived English forms, but this was restricted to forms which go through a stem change. The novel roots in our design were all regular (irregular roots exist in Hebrew as well; see Velan et al., 2005). Thus, it seems unlikely that phonological characteristics of the root itself would be sufficient in explaining how novel verbs are created and how verbal classes are assigned.

If phonological similarity is not a sufficient source of Hebrew binyan generalization, what are the sources that Hebrew speakers employ when extending binyanim to novel verbs? One source examined in the present work is argument structure, since all binyanim have certain constraints or biases in this domain, except Paal. However, are Hebrew speakers sensitive to biases of argument structure in novel verb formation? According to the present findings, the answer is positive. Significantly more Piel verbs were produced in the context of a direct object. The present findings provide clear evidence that abstract structure can be employed to assign an inflectional class to a novel verb. We managed to isolate argument structure from semantic properties by minimizing the semantic information in the experimental design, unlike previous studies (Berman, 1993; Bolozky, 1999a), in which the semantic meaning of the novel verb was provided, and therefore it was not clear whether speakers were sensitive to argument structure alone or, at least partly, to the semantic information provided to them. To summarize, Hebrew speakers are sensitive to the specific properties and biases of Hebrew classes and make use of cues that are relevant to the Hebrew classes when producing novel verbs.

The third pattern of results that differs from the typical findings of morphological generalization of inflectional classes is that two classes were clearly the preferred response across groups, conditions and experiments, without very strong evidence of the response being affected by phonological analogy. Typically, at least in generalization of Romance inflectional classes, one morphological pattern is the most dominant, highly preferred compare to others. In Romance inflectional classes, for example, that would be class I (e.g., Say & Clahsen, 2002). In contrast, in the present findings, Paal and Piel responses comprised more than 80% of the total answers in both experiments. The other five binyanim were produced rarely, with Hifil following Paal and Piel with 8% overall.

To understand why the two classes were clearly preferred, we will first have a closer look at Paal to unravel the mystery of Paal's high response frequency, which stands in sharp contrast to the fact that the Paal class hardly extends to novel verbs in Modern Hebrew (for a review, see Bolozky, 1999a). I suggest that this pattern demonstrates influence from the current design that stems from the lack of any semantic information provided in the task about the meaning of the verb. While this type of design allows a precise examination of a 'purer' morphological process, the generalization under the conditions of the task did not resemble a natural generalization process of inflectional classes, outside of lab conditions. Inflectional class generalization outside of the lab conditions involves the formation of a new lexical entry with certain semantic properties, which are often correlated with the class of the verb in Hebrew. For example, verbs with causative meaning are often Piel or Hifil (e.g., *heevid* 'caused someone to work' belongs to Hifil). The Paal class is the only class that equally accommodates all types of verbs (active and stative, transitive and intransitive), irrespective of their semantic properties. Aronoff (1994) therefore referred to Paal as 'derivationally marginal' and 'not semantically robust in the way that we expect of productive patterns' (p.146), pointing out the differences between Paal and class I of Romance languages, which is a default that is both frequent and productive. Yet, despite the low productivity, he defines Paal as the default, since its function is 'to sweep up what the more powerful classes have left in their wake' (p.146). This might be the reason why so many Paal verbs were formed in Publication IV, much more than Paal's actual productivity in Modern Hebrew, and much more than Paal responses in previous studies in which the meaning of the novel verb was provided (Berman, 1993; Bolozky, 1999a). Once semantic information was unspecified, the strong derivational binyanim were not employed, leaving the Paal class to 'sweep up' those verbs.

Piel response frequency, similarly to Paal, was very high, but unlike Paal, Piel responses had a clear effect of argument structure context and a relatively weak effect of phonological similarity. Yet, this does not mean Piel responses were restricted to contexts with a direct object or with high similarity to existing Piel roots. On the contrary, relatively a high number of Piel responses were produced in the non-direct object condition, both in experiment 1 (which did not include a direct object) and experiment 2 (39% in experiment 1, 29% in experiment 2), second only to Paal (45% and 39%, accordingly). Very high preference for a Piel response was found in conditions that were not phonologically similar to Piel in experiment 1 (37% Piel responses in the no-similarity condition and 38% in the Paal similarity condition). In summary, Piel class is preferred over other classes (besides Paal) also in less 'ideal' phonological, semantic or syntactic contexts, resulting in a very high response preference overall.

Taking together, binyan generalization demonstrates different properties from the Romance inflectional classes: (i) argument structure context has a strong influence on generalization, (ii) phonological similarity is relatively weak for the productive class and failed to be found at all for the unproductive class, and (iii) the system displays two preferred classes which are generalized beyond phonological similarity and a specific argument structure. But how do these findings fit with our proposal following the priming results that Paal and Piel have different types of representations? The idea that Paal verbs have an unstructured, whole-stem central representation seems to be problematic when considering the generalization results. If we assume Paal verbs are not represented as structured by root + pattern, this implies they are not formed by rules. In addition, a phonological similarity effect was not found, which suggests Paal verbs are not extended by phonological analogy. If not by symbolic rules and not by analogy, how are they extended so frequently to novel forms in our elicited production experiment?

There is not a clear answer for this complex issue. Two possibilities are discussed here which may settle this discrepancy. One possibility is that Paal is not extended by symbolic rules but by analogy to stored forms in the lexicon. Admittedly, we failed to find evidence for generalization based on phonological analogy in our implementation of the AML (experiment 1 of Publication IV), yet this does not determine that all types of analogy-based generalization are absent in binyan generalization. It merely demonstrates that the specific algorithm we implemented did not explain the pattern of responses of our participants. As discussed above, other models and algorithms based on phonological similarity, such as connectionist networks, might be a better fit for Hebrew binyan generalization. Generally, in analogy-based models, patterns with high type frequency (such as the English regular past tense and class I of the Romance inflectional classes, which have been claimed to be rule-based by the dual-route approach) are more likely to be similar to an arbitrary novel form because their phonological space is more evenly and broadly distributed. Given Paal's high frequency and broad phonological space, such models may, in principle, be able to predict Paal's high response preference.

In addition, to better predict binyan generalization, similarity-based models may benefit from employing non-phonological information. So far, similarity-based models mostly employed phonological analogy to predict linguistic behavior, but other types of non-phonological analogies can also in principle drive morphological generalization (e.g., orthographic analogy in Dutch plural inflection; see Keuleers et al., 2007). This is particularly relevant for binyan generalization, which already showed in the present work and previous studies to be influenced by argument structure and semantic information. An analogy-based model for binyan generalization, whether a connectionist model or an exemplar-based model, would probably yield more precise results when it incorporates semantic and argument structure information to predict the Hebrew class in novel verb formation. To conclude, it is possible that Paal representations are stored as whole stems and not rule-based, and generalized by analogy to stored forms, although we have not found the exact algorithm that explains the results.

A second possibility is that Paal (like Piel) is extended by derivational rules of root and pattern, but at the same time Paal stems are not accessed via decomposition of their constituents. The process of how forms are generalized and how their stored representations are accessed are not always aligned, since they present two different questions, addressing separate issues (Alegre & Gordon, 1999b). Regarding accessing stored representations, several dual-route frameworks (e.g., the *augmented addressed morphology* [Caramazza, Laudanna, & Romani, 1988] and the *morphological race model* [Frauenfelder & Schreuder, 1992]) have proposed that access to complex forms can be operated via two routes that run in parallel and compete with each other: a decomposition route and a whole-form route. In most cases, the whole-form route wins, yet in some cases the decompositional route will win. Factors that determine the route are complex and include base and surface frequency, as well as phonological and semantic transparency. For example, it has been proposed that forms of the regular English past tense are accessed as whole forms when the form is highly frequent, and via decomposed constituents when the frequency of the form is low, and at the same time regular patterns are generalized by rules and not analogy (Alegre & Gordon, 1999b). In the case of novel words, access is proposed to be performed via the decompositional route (Caramazza et al., 1988). We can adopt this type of framework to the present findings, postulating that both routes are in principle possible, but Paal stems tend to be accessed as wholes, with the exception of perhaps Paal stems with very low frequency and novel Paal stems, which are accessed by decomposition. Therefore, it is possible that Paal verbs are stored also in a structured manner and acquired by a rule, but in most cases the whole-stem route wins in the case of Paal. This raises the question why Paal stems would be more susceptible to run via the whole-form route while Piel stems run via the decomposition route. One possibility that we discussed earlier is that due to the ambiguous nature of the vowel pattern of Paal, it is more efficient to access it as a whole. Clearly, this line of thought still requires a more detailed discussion and examination in future research.

To conclude, the results of the generalization study revealed complex patterns that do not perfectly align with the results of the priming studies (Publications I and III) or with previous generalization studies of Hebrew verbs (e.g., Bolozky, 1999a). However, such discrepancies should encourage us to adapt and develop our theoretical assumptions. They provide us the opportunity to achieve a deeper understanding of the Hebrew mental lexicon, and the properties that are specific to the Semitic family. One such specific property is the highly significant role that we found for argument structure information in generalization of Hebrew inflectional classes, a property that, to the author's knowledge, does not play a role in generalization of inflectional classes of non-Semitic languages.

7.3 Morphological Processing and Generalization in L2 Speakers

The present experimental work directly compared L2 and L1 speakers in two dimensions of morphology: complex word processing and generalization of novel complex forms. The results revealed significant L1/L2 differences in sensitivity to morphological and morpho-syntactic properties in recognition of existing forms and formation of novel forms. Table 3 illustrates that compared to L1 speakers, highly proficient L2 speakers of Hebrew showed reduced sensitivity to abstract morphological differences between two inflectional classes that are semantically and phonologically similar. This was reflected in (i) similar root priming effects in Paal and Piel, which were (ii) restricted to non-finite forms (Publication II), and (iii) a smaller effect of argument structure for Piel responses compared to L1 speakers (Publication IV).

Recall that theoretical frameworks for L2 have generally debated the question of whether the source of L1/L2 performance differences lies within general cognitive properties, like working memory capacity (e.g., McDonald, 2006), or whether specific linguistic properties are involved (e.g., Clahsen & Felser, 2018). We argue that the contrasts found in the present thesis are more compatible with the proposal that linguistic-specific properties, rather than solely general cognitive ones, are the source of the L1/L2 performance differences, and that the results are best to be viewed as evidence that abstract structural, morphological and morpho-syntactic properties are challenging for late bilinguals, even at high proficiency levels, and that L2 morphological processing and generalization are less affected by structural information sources (Clahsen & Felser, 2018).

First, we consider the cognitive-general approach (Cunnings, 2017; McDonald, 2006; Mcdonald & Roussel, 2010) and explain why it is not able to explain the full L2 pattern found in Publications II and IV. As previously described, this view claims that the processing architecture of L2 is the same as L1, and attributes the L1/L2 performance differences at high levels of proficiency to cognitive sources, arguing that L2 processing requires increased demands and capacities of working memory (McDonald, 2006), or that it is more susceptible to interference in memory retrieval (Cunnings, 2017). Although these accounts mostly refer to L2 parsing of full sentences, some parts of the present results can be regarded as stemming from a general cognitive overload. For instance, the L2 group showed very slow RTs compared to the L1 group across conditions (Publication II), which can be thought to reflect a cognitive overload for the L2 group. It is reasonable to assume a cognitive overload when considering the overall procedure and design, namely the limited display of 336 target words (2000ms) in a complex orthographic script, which includes only partial vowel representation and is different from the L1 script. The similarities in L1/L2 patterns in the generalization study (Publication IV) can also indicate that when no time pressure and processing load are given, L2 speakers generally make use of the same information sources as L1 speakers: both in employing phonological similarity information and argument structure information.

However, some fine-grained group differences are difficult to explain solely by the general cognitive sources. First, L2 speakers showed selective root-priming effects: priming effects for non-finite but not for finite forms. Within the framework of the cognitive-general approach, it is expected that the brief presentation of the prime poses a cognitive overload on the L2 speakers that would result in either lack of priming effects in all conditions or priming effects overall that are generally weaker than in L1 (as was found in effect powers for Arabic L2 speakers in Freynik, Gor, & O'Rourke, 2017). When the prime presentation is not visually brief, but auditory, similar morphological priming effects (whether derived or inflected forms) in L2 and L1 speakers usually occur (Basnight, Chen, Hua, Kostić, & Feldman, 2007; Feldman et al., 2010; Gor & Jackson, 2013), arguably by the cognitive-general approach, due to the longer and conscious presentation of the prime, which creates less cognitive overload. Within this framework, it is challenging to explain why a root-priming effect was not found for finite forms but was found for non-finite forms. It is more likely that the specific linguistic properties of the finite and non-finite forms are the source of the empirical contrast that was found between them. The second group effect is the root-priming effect for Paal in the non-finite form which stands in contrast to the lack of such effect in the L1 group. The fact that priming was found for L2 but not L1 despite any cognitive overload that might be present in the L2 group raises doubts about general cognitive properties as the sole source of differences between L1 and L2. In summary, cognitive sources may account for some of the patterns of L2 morphological processing, but to fully understand the picture, linguistic representational sources should be taken into account as well.

Another explanation for the masked priming results can be that the effects are orthographic-driven. It is true that in general orthographic effects have not been found in masked priming in Arabic and Hebrew when testing L1 speakers (Boudelaa & Marslen-Wilson, 2005; Frost et al., 2005) nor transposed letter effects (TLEs; in root-based words: Perea et al. 2010 [Arabic], Velan & Frost, 2009 [Hebrew]; although see Oganyan, 2017 for TLEs in Hebrew), strengthening the claim that the source of root-priming effects is morphological in nature and not orthographic. On the other hand, for L2 speakers, this claim is not as strong. A recent masked priming study on L2 Hebrew speakers (Oganyan, 2017) showed TLEs for L2 both in nouns and verbs, suggesting that L2 Hebrew speakers rely more on orthographic information than L1 speakers. This finding complements several recent findings, reporting orthographic priming effects in L2 speakers (Diependaele et al., 2011; Feldman et al., 2010; Heyer & Clahsen, 2015; Li et al., 2017; Qiao & Forster, 2017). Thus, an alternative explanation to the present L2 findings might be orthographic-driven. The priming effect yielded for the non-finite forms in L2 could be driven by orthographic overlap and not necessarily by morphological overlap. Explaining the full pattern of results based on only orthographic overlap, however, is not feasible, since the finite primes had similar orthographic overlap to the target compared to the non-finite primes and yet they did not yield a priming effect. Still, increasing evidence for orthographic effects in L2 is to be considered in the design of future studies on morphology, either by closely controlling for orthographic overlap or by putting more effort into testing the auditory domain.

Our proposal is that the most compatible explanation for the full pattern of results is one that takes into account linguistic-specific properties. The proposal for lexical access in L2 on the basis of the findings in Publication II will be discussed first, followed by a discussion about morphological generalization in L2 with regard to the results of Publication IV.

Regarding the group contrasts found in Publication II (masked priming), I propose that they can best be interpreted as resulting from (i) an L2 reduced sensitivity to the morphological difference between inflectional classes and (ii) L2 processing costs in the recognition of finite forms. Both claims will be further discussed in the following paragraphs. As described in the previous section, L1 speakers were able to access the class information at the initial access level. Consequentially, they processed Paal and Piel in different ways; Paal stems were accessed as wholes and Piel stems were accessed via decomposition to their constituents (root and pattern). Failing to find a root-priming asymmetry between Paal and Piel in both forms in L2, in sharp contrast to the L1 results, suggests that the L2 group has accessed both classes in a similar way. Since the main difference between Paal and Piel is morphological, i.e., membership in different inflectional classes, accessing information of Paal and Piel, like the L1 pattern suggests, requires access to abstract morphology during initial recognition; that is, it is not based on concatenative affixes but requires an extraction of the abstract vowel pattern from the verbal form. I suggest that the L2 group did not access the vowel pattern constituent which carries the class information during the presentation of the prime. This is consistent with the *shallow structure hypothesis* (SSH; Clahsen & Felser, 2006a, 2006b, 2018), which posits that L2 speakers are less efficient with computation and manipulation of abstract grammatical representations in real-time processing and claims that they tend to rely on processes which do not involve decomposition. Another possible explanation is the one by Gor et al. (2017), proposing that L2 lexical access focuses mainly on accessing the lexical meaning, therefore extracting the root and/or the stem, while underusing recombination of the constituents and checking of the inflected form, which provide access to more abstract grammatical information.

I propose that the finding that root priming was restricted to non-finite forms can be attributed to larger processing costs for finite forms. The two types of forms differ greatly in the complexity of the information they carry, although on the surface they both include inflectional affixes. Unlike affixes of non-finite forms, affixes of finite forms specify complex morpho-syntactic information, like gender, person and number. Accessing such complex and abstract morpho-syntactic features can yield processing costs to the L2 processor. L2 speakers have been reported to show reduced sensitivity to morpho-syntactic properties, mainly in sentence processing (e.g., Coughlin & Tremblay, 2013; Hopp, 2010, 2013; Jiang, Hu, Chrabaszcz, & Ye, 2015; Tokowicz & MacWhinney, 2005), but also in word-level morphological processing (e.g., Bosch & Clahsen, 2016; Gor et al., 2017). Given our findings and previous literature, the core problem of L2 complex word recognition seems to be the morpho-syntactic features encoded in the affixes. I suggest that, in principle, L2 speakers are able to strip off morphological affixes, but struggle to do so when the affix carries complex morpho-syntactic information (see also Jacob et al., 2017).

While it is relatively agreed upon that finite forms carry more complex morpho-syntactic information, it is not clear whether the lack of root-priming effect in finite forms in Publication II reflects direct access as whole form or simply a lack of full access to the form. On the one hand, several accounts, like the SSH (Clahsen & Felser, 2018) and the *declarative/procedural model* (Ullman, 2005), have interpreted the lack of morphological priming from inflected forms in masked priming studies (*walked–walk)* as an indication for a direct whole-form access of inflected finite forms to the mental lexicon without initial decomposition (Kirkici & Clahsen, 2013; Silva & Clahsen, 2008). This claim has been further supported by frequency studies, for instance, Bowden et al. (2010), showing that L2 production of Spanish inflected forms is affected by surface frequency even when L1 production is not (since surface frequency effect is often interpreted as whole-form access). Such accounts would therefore claim that, based on the results of Publication II, L2 Hebrew speakers did not decompose the finite inflected forms of Paal and Piel but accessed them directly; the Hebrew suffix *-ti* (specifying first person, singular, past form) was not stripped from the stem and the root was not extracted. Thus, the L2 access representation of a word like *limdati* would also be /limadti/.

A second possibility is that the L2 speakers were not able to fully access the finite prime at all (Gor, Chrabaszcz, & Cook, 2018), not in a decomposed manner nor via a whole form. It is possible that forms that carry complex morpho-syntactic information require more time to be fully accessed in L2 speakers, and this does not tend to occur under masked priming conditions. It may be that when the prime is presented for longer, or auditorily, highly proficient L2 speakers would be able to decompose an inflected form. This is a particularly relevant argument for our L2 group, when taking into account the challenge they are facing with regard to reading an L2 script that is different from their L1 script, written in a different direction (right to left) and presenting only partial vowel information. Indeed, the L2 group in Publication II showed very slow RTs overall (mean raw RTs of over 900ms in all conditions). This finding is consistent with the numbers of the cross-modal priming study of Freynik et al. (2017), who tested L2 Arabic speakers (with L1 English) and reported over 1000ms RTs in all conditions. Typically, priming studies of highly proficient L2 speakers do not yield such slow RTs, but reading in an L2 script that differs from the L1 script tends to increase the RTs (e.g., slower RTs of L2 English for L1 speakers of Chinese than L1 speakers of German; Silva & Clahsen, 2008). Future L2 studies should be attentive to the special challenge the Semitic script presents to L2 learners, which is not directly related to their morphological processing abilities.

Regarding the effects from the non-finite primes, from a decompositional perspective, it is possible that the root effect reflects that the non-finite verbal forms (of Paal and Piel), unlike the finite forms, were accessed through a decompositional route. I propose that two types of decomposition were carried out: inflectional and derivational. The first decomposition involved inflectional affixes. The affix *le-* is an inflectional affix but it does not carry any morpho-syntactic features like gender, person or tense. It simply marks the infinitive form. It is therefore less demanding for L2 speakers to access, in contrast to the finite suffix *-ti*, which encodes the features of person, number and tense (first person singular past). Thus, the L2 speakers were able to decompose the form into a prefix and a stem (*lelamd* 'to teach' → *le+lamed*), but they were not able to decompose the more demanding finite form into a stem and a suffix (*limadti '(I) taught'* → *limad+ti*). I postulate that the decomposition of the inflectional affixes is a crucial step, and only if this step is executed will the second step of lexical recognition be possible. Therefore, since the finite suffix *-ti* had not been decomposed, the second phase of decomposition, namely root extraction, was not carried out.

The root priming found for the infinitive forms is consistent with previous findings testing highly proficient L2 speakers. Root-priming effects were already reported in a cross-modal priming study testing L2 speakers of Arabic (Freynik et al., 2017), and priming effects in derived forms have been consistently found in Indo-European languages (e.g., Diependaele et al., 2011; Heyer & Clahsen, 2015; Jacob et al., 2017; Kirkici & Clahsen, 2013; Li et al., 2017; Silva & Clahsen, 2008; Voga et al., 2014). Jacob et al. (2017) directly compared priming from inflected and derived forms on the same target (in a masked priming paradigm) and found a priming effect for derived but not inflected forms in highly proficient L2 German speakers. There has been a debate regarding whether priming effects from derived forms necessarily imply that a decomposition of constituents occurred. One can interpret the root-priming effects found for the non-finite forms in more than one way. The first possibility is that the root priming occurred due to a derivation-based decomposition where the root was extracted from the stem. For the non-finite forms, which were first decomposed into a stem and an inflectional affix (*lelamd* 'to teach' → *le+lamed)*, the extraction of the root was possible (*lamed* → *L-M-D*). An alternative explanation for the root priming in the non-finite condition would be based on the SSH (Clahsen & Felser, 2018; Kirkici & Clahsen, 2013), arguing that derivational effects in L2 reflect lexically mediated activation between two partially overlapping lexical entries. In general, these two interpretations are difficult to tease apart.

Publication IV focused on the properties of generalization of Paal and Piel to novel words. Here, the results are also consistent with the claim regarding a reduced L2 sensitivity to structural information, reflected by a reduced argument structure effect; L2 speakers produced more Piel forms in the context of a direct object than without, but not to the same extent as L1 speakers. An argument structure bias that refers to a specific lexical entry (e.g., the verb *admit* is usually followed by a complement phrase and less often by a direct object) generally requires linking a specific lexical entry and an abstract syntactic structure (e.g., Bernolet & Hartsuiker, 2010; Peter, Chang, Pine, Blything, & Rowland, 2015), so the process can be evaluated as consisting of both lexical and syntactic properties. The bias tested in the present work is more abstract, as it is an argument structure bias of a whole abstract morphological class, rather than a specific lexical entry, and requires linking an argument structure and a morphological class. Despite the more abstract nature of the argument structure bias, the L2 speakers in our study were sensitive to this bias and able to extend their knowledge about the tendency of Piel verbs to include a direct object to novel forms. However, the generalization of the bias for Piel was not as strong as the one the L1 group displayed. While it is true that statistical effects of L2 speakers tend to have smaller power than those of L1 speakers (e.g., Gor et al., 2017), the fact that L2 speakers showed similar effects to L1 speakers in employing analogical information (partial similarity effect for Piel) indicates that their reduced sensitivity is restricted to structural and morphological properties rather than to surface processes like phonological analogy, and suggests that they did not rely on argument structure information consistently and efficiently. Taken together, the results of Publication IV present further evidence for L2 reduced sensitivity to structural information.

Additional L1/L2 contrast in responses was detected; the L2 group displayed excessive production of Paal forms, more than in L1, in both experiments of Publication IV. This may imply that the L2 group implemented a simpler strategy of generalization in addition to the use of structural cues. The Paal class fits equally well in any syntactic context, as it does not have an argument structure bias. Under the experimental conditions of Publication IV, generating Paal verbs allowed to easily avoid the processing of the direct object in the sentence. L2 speakers have been previously reported to tend to avoid grammatical information if it is not essential for the task and process this information only when the design indirectly draws their attention to it (Gor et al., 2017). The fact that the overuse of Paal responses appeared also in experiment 1, which did not include direct objects, can hint that this strategy, although not necessarily intentional, is rather general, allowing the reader to skip the processing of potential structural information. This is naturally only a post-hoc speculation, but this pattern fits well with the claim that L2 speakers underuse structural information and generally prefer the simplest analysis that is compatible with the input (e.g., Rah & Adone, 2010; Roberts & Felser, 2011).

8 Conclusions and Future Perspectives

The present dissertation investigated sources of variability within the morphological domain. To what extent are morphological processes and representations universal? To what extent are they shaped by specific properties of languages and speakers? To address these questions, the work focused on (i) a less explored language with unique morphological properties, Hebrew, and (ii) a less investigated population, late learners of a second language.

Admittedly, while those questions are very general and carry wide implications, the scope of the present investigation is somewhat limited. The dissertation examined only the linguistic phenomenon of inflectional classes, emphasizing two verbal Hebrew classes (Paal and Piel) throughout the whole experimental path of the dissertation. While such focus provided us the opportunity to reach a precise and detailed picture of the Hebrew binyanim, it may be questionable, based on this narrow investigation, to provide general insights about universals and particulars in the mental lexicon. Expanding this type of research to other phenomena would be beneficial to validate the generalizability of our conclusions and assure us they are not specific to the phenomenon of binyanim. In addition, while they constitute a profound basis for the role of productivity in Hebrew morphology in L1 and L2 speakers, the methods used in the present dissertation are relatively limited, producing behavioral data. The implementation of online measures, such as ERP, would have enabled us to achieve a deeper understanding of Hebrew morphological processing in L1 and L2 speakers.

Still, our results provide significant insights about universals and particulars in the language system. With regard to possible effects of the *specific properties of Hebrew morphology*, we found that, on the one hand, the empirical patterns were more similar to patterns often reported in the Indo-European literature than claimed by the predominant root-based approach for Hebrew. We have found (i) evidence for structured and unstructured representations based on productivity (expressed by root-priming contrast between Paal and Piel; Publication I), and (ii) evidence for semantic influence on central-level morphological representations (reflected by a semantic transparency effect for Paal; Publication III). On the other hand, several unique patterns emerged: (i) the semantic influence appears to be selective, as it emerged only for Paal but not Piel (Publication III), and (ii) phonological-based generalization was rather weak and partial, while (iii) generalization based on argument structure information was very robust (Publication IV). With regard to possible effects of the *specific properties of L2 speakers*, a large part of the performance was very similar to the L1 speakers, mainly in the generalization of binyanim: a relatively weak reliance on phonological-based generalization and a robust influence of argument structure context. However, specific effects emerged for the L2 group as well, strongly displayed in the masked priming study (Publication II), with priming effects shown only from non-finite forms, without a difference in priming between productive and unproductive inflectional classes, arguably showing processing costs for complex morpho-syntactic information and reduced sensitivity to inflectional class information. To a smaller extent, this claim is also consistent with the generalization findings, yielding a reduced sensitivity to argument structure in generalization.

The different patterns have been explained here within a representational framework of language, postulating a discrete level for morphological structure. Within this framework, we were able to account for the differences in the empirical patterns in Hebrew compared to other languages and for differences in patterns of L2 compared to L1 speakers, while still maintaining a universal architecture of the language system. The results of the present work demonstrated the importance of addressing the sources of variability in language performance, whether they stem from characteristics of the speaker or the language under investigation. Testing one language typology and a homogenous group of L1 speakers in order to reach homogenous results allows to build theoretical accounts of the language system more easily, relying on clearer effects; however, it would be questionable to consider accounts that are built on such empirical basis as universal. Such accounts can lead to a misconception of how the mental architecture of language works. They disregard the flexible properties of the cognitive system which allow various speakers of different languages to successfully make use of the language system. On the other hand, variability in performance does not denote that the principles of the cognitive system should be defined differently for each language. Universal principles should be established and at the same time be flexible enough to accommodate the variance that stems from the different language and speaker properties. The goal of future psycholinguistic research should be to find an appropriate balance between universal and particular aspects of language performance, aiming to develop and adapt theoretical frameworks so they can account for certain language-specific and speaker-specific patterns, and at the same time address universal properties of morphology and language in general.

References

Adam, G. (2002). *From variable to optimal grammar: Evidence from language acquisition and language change* (doctoral dissertation). Tel-Aviv University, Israel.

Agathopoulou, E., & Papadopoulou, D. (2009). Regularity patterns of the Greek past perfective: a psycholinguistic study. In G.K. Giannakis, M. Baltazani, G.I. Xydopoulos & A. Tsangalidis (Eds.), *Proceedings of the 8th international conference on Greek linguistics, Ioannina* (pp. 1–13).

Albright, A. (2002). Islands of reliability for regular morphology: Evidence from Italian. *Language*, 684–709.

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition, 90*, 119–161.

Alegre, M., & Gordon, P. (1999a). Rule-based versus associative processes in derivational morphology. *Brain and Language*, *68*, 347–354.

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

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

Amenta, S., & Crepaldi, D. (2016). The variable mind? How apparently inconsistent effects might inform model building. *Frontiers in Psychology*, *7:*185.

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.

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

Arad, M. (2005). *Roots and patterns: Hebrew morpho-syntax*. Dordrecht: Springer.

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT press.

Aronoff, M. (2013). The roots of language. In S. Cruschina, M. Maiden, & J.C. Smith (Eds.), *The boundaries of pure morphology* (pp.161–180). Oxford: Oxford University Press.

Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1991* (pp. 109–149). Dordrecht: Kluwer.

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

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*, 12–28.

Balota, D. A., Aschenbrenner, A. J., & Yap, M. J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1563–1571.

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

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

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language & Linguistic Theory*, *12*, 571–596.

Bat-El, O. (2002). Semitic verb structure within a universal perspective. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based, morphology* (pp. 29–59). Amsterdam: John Benjamins.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Benmamoun, E. (1999). Arabic morphology: The central role of the imperfective. *Lingua*, *108*, 175–201.

Benmamoun, E. (2003). The role of the imperfective template in Arabic morphology. In J. Shimron (Ed.), *Language processing and language acquisition in a root-based morphology* (pp. 99–114). Amsterdam: John Benjamins.

Berent, I. (2002). Identity avoidance in the Hebrew lexicon: Implications for symbolic accounts of word formation. *Brain and language*, *81*, 326–341.

Berent, I., Everett, D. L., & Shimron, J. (2001). Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology*, *42*, 1–60.

Berent, I., Pinker, S., & Shimron, J. (1999). Default nominal inflection in Hebrew: Evidence for mental variables. *Cognition*, *72*, 1–44.

Berent, I., & Shimron, J. (1997). The representation of Hebrew words: Evidence from the obligatory contour principle. *Cognition, 64*, 39–72.

Berent, I., Vaknin, V., & Marcus, G. F. (2007). Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition*, *104*, 254–286.

Berman, R. A. (1993). Marking of verb transitivity by Hebrew-speaking children. *Journal of Child Language*, *20*, 641–669.

Berman, R. A. (1997). Modern Hebrew. In R. Hetzron (Ed.), *The Semitic languages* (pp. 312–333). New York, NY: Routledge.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming?. *Cognition*, *114*, 455–461.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research*, *28*, 213–226.

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 489–511.

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

Bick, A. S., Frost, R., & Goelman, G. (2010). Imaging implicit morphological processing: evidence from Hebrew. *Journal of Cognitive Neuroscience*, *22*, 1955–1969.

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

Bick, A. S., Goelman, G., & Frost, R. (2011). Hebrew brain vs. English brain: Language modulates the way it is processed. *Journal of Cognitive Neuroscience*, *23*, 2280–2290.

Blevins, J. P. (2006). English inflection and derivation. In B. Aarts & A. M. McMahon (Eds.), Handbook of english linguistics (pp. 507–536). London: Blackwell.

Blom, E., Polišenská, D., & Weerman, F. (2006). Effects of age on the acquisition of agreement inflection. *Morphology*, *16*, 313–336.

Bolozky, S. (1978). Word formation strategies in the Hebrew verb system: Denominative verbs. *Afroasiatic Linguistics*, *5*, 111–136.

Bolozky, S. (1982). Strategies of Modern Hebrew verb formation. *Hebrew Annual Review, 6*, 69–79.

Bolozky, S. (1999a). *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill.

Bolozky, S. (1999b). On the special status of the vowels a and e in Israeli Hebrew. *Hebrew Studies*, *40*, 233–250.

Bosch, S., & Clahsen, H. (2016). Accessing morphosyntax in L1 and L2 word recognition. *The Mental Lexicon*, *11*, 26–54.

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

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

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

Boudelaa, S., & Marslen-Wilson, W. D. (2015). Structure, form, and meaning in the mental lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience*, *30*, 955–992.

Boudelaa, S., Pulvermüller, F., Hauk, O., Shtyrov, Y., & Marslen-Wilson, W. (2010). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience*, *22*, 998–1010.

Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language Learning*, *60*, 44–87.

Brooks, P. J., Braine, M. D., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language*, *32*, 76–95.

Brovetto, C., & Ullman, M. T. (2005). The mental representation and processing of Spanish verbal morphology. In Eddington, D. (Ed.), *Selected proceedings of the 7th hispanic linguistics symposium* (pp. 98–105). Somerville, MA: Cascadilla Proceedings Press.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, *51*, 5–13.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*, 425–455.

Bybee, J. L., & Moder, C. L. (1983). Morphological classes as natural categories. *Language*, 251–270.

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

Carota, F., Bozic, M., & Marslen-Wilson, W. (2016). Decompositional representation of morphological complexity: Multivariate fMRI evidence from Italian. *Journal of Cognitive Neuroscience*, *28*, 1878–1896.

Chen, L., Shu, H. U. A., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition, 10*, 161–174.

Clahsen, H. (1997). The representation of participles in the German mental lexicon: evidence for the dual-mechanism model. In G. E. Booij & J. van Marle (Eds.), *Yearbook of morphology 1996* (pp. 73–95). Dordrecht: Kluwer.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, *22*, 991–1013.

Clahsen, H. (2006). Dual-mechanism morphology. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (Vol. 4, pp. 1–5). Oxford: Elsevier.

Clahsen, H., Balkhair, L., Schutter, J. S., & Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research*, *29*, 7–31.

Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics*, *23*, 201–250.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564–570.

Clahsen, H., & Felser, C. (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition, 40,* 693–706.

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

Clahsen, H., Gerth, S., Heyer, V., & Schott, E. (2015). Morphology constrains native and non-native word formation in different ways: Evidence from plurals inside compounds. *The Mental Lexicon*, *10*, 53–87.

Clahsen, H., Martzoukou, M., & Stavrakaki, S. (2010). The perfective past tense in Greek as a second language. *Second Language Research*, *26*, 501–525.

Clahsen, H., & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, *120*, 2627–2637.

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

Colombo, L., Stoianov, I., Pasini, M., & Zorzi, M. (2006). The role of phonology in the inflection of Italian verbs: A connectionist investigation. *The Mental Lexicon*, *1*, 147–181.

Coughlin, C. E., & Tremblay, A. (2013). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, *34*, 615–646.

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

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

Crystal, D. (2012). *English as a global language*. Cambridge university press.

Culbertson, J., Gagliardi, A., & Smith, K. (2017). Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language, 92*, 343–358.

Culbertson, J., & Wilson, C. (2013). Artificial grammar learning of shape-based noun classification. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the cognitive science society* (pp. 2118–2123). Austin, TX: Cognitive Science Society.

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition, 20*, 659–678.

Cuskley, C., Colaiori, F., Castellano, C., Loreto, V., Pugliese, M., & Tria, F. (2015). The adoption of linguistic rules in native and non-native speakers: Evidence from a Wug task. *Journal of Memory and Language, 84*, 205–223.

Dawdy-Hesterberg, L. G., & Pierrehumbert, J. B. (2014). Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience, 29*, 1268–1282.

Deutsch, A. (2016). The separability of morphological processes from semantic meaning and syntactic class in production of single words: Evidence from the Hebrew root morpheme. *Journal of Psycholinguistic Research, 45*, 1–28.

Deutsch, A., Frost, R., & Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1238–1255.

Deutsch, A., Frost, R., Pelleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review*, *10*, 415–422.

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes*, *20*, 341–371.

Deutsch, A., & Malinovitch, T. (2016). The role of the morpho-phonological word-pattern unit in single-word production in Hebrew. *Journal of Memory and Language*, *87*, 1–15.

Deutsch, A., & Meir, A. (2011). The role of the root morpheme in mediating word production in Hebrew. *Language and Cognitive Processes*, *26*, 716–744.

Deutsch, A., Velan, H., & Michaly, T. (2018). Decomposition in a non-concatenated morphological structure involves more than just the roots: Evidence from fast priming. *The Quarterly Journal of Experimental Psychology*, *71*, 85–92.

Dewaele, J. M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition, 4*, 275–297.

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

Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition*, *37*, 895–908.

Dussias, P. E., Marful, A., Gerfen, C., & Molina, M. T. B. (2010). Usage frequencies of complement-taking verbs in Spanish and English: Data from Spanish monolinguals and Spanish—English bilinguals. *Behavior Research Methods, 42,* 1004–1011.

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

Dussias, P. E., & Scaltz, T. R. C. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica, 128,* 501–513.

Eddington, D. (2000). Analogy and the dual-route model of morphology. *Lingua, 110,* 281–298.

Eddington, D. (2002). Dissociation in Italian conjugations: A single-route account. *Brain and Language, 81*, 291–302.

Farhy, Y. (under review). Morphological Generalization of Hebrew Verbal Classes: An Elicited Production Study in L1 and L2 Speakers, under review with *The Mental Lexicon*.

Farhy, Y., & Veríssimo, J. (2018). Semantic Effects in Morphological Priming: The Case of Hebrew Stems. *Language and Speech*, published online November 30 2018. doi: 10.1177/0023830918811863

Farhy, Y., Veríssimo, J., & Clahsen, H. (2018a). Universal and particular in morphological processing: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology, 71*, 1125–1133.

Farhy, Y., Veríssimo, J., & Clahsen, H. (2018b). Do late bilinguals access pure morphology during word recognition? A masked-priming study on Hebrew as a second language. *Bilingualism: Language and Cognition*, *21*, 945–951.

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

Feldman, L. B., Frost, R., & Pnini, T. (1995). Decomposing words into their constituent morphemes: Evidence from English and Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 947–960.

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

Feldman, L. B., Milin, P., Cho, K. W., Moscoso del Prado Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience*, *9:*111.

Feldman, L. B., O'Connor, P. A., & del Prado Martín, F. M. (2009). Early morphological processing is morphosemantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review*, *16*, 684–691.

Feldman, L. B., & Soltano, E. G. (1999). Morphological priming: The role of prime duration, semantic transparency, and affix position. *Brain and Language*, *68*, 33–39.

Feldman, L. B., Soltano, E. G., Pastizzo, M. J., & Francis, S. E. (2004). What do graded effects of semantic transparency reveal about morphological processing? *Brain and Language*, *90*, 17–30.

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, *92*, 333–375.

Foote, R. (2017). The storage and processing of morphologically complex words in L2 Spanish. *Studies in Second Language Acquisition*, *39*, 735–767.

Ford, M. A., Davis, M. H., & Marslen-Wilson, W. D. (2010). Derivational morphology and base morpheme frequency. *Journal of Memory and Language*, *63*, 117–130.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116–124.

Fraundorf, S. H., & Jaeger, T. F. (2016). Readers generalize adaptations to newly-encountered dialectal structures to other unfamiliar structures. *Journal of Memory and Language*, *91*, 28–58.

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

Freynik, S., Gor, K., & O’Rourke, P. (2017). L2 processing of Arabic derivational morphology. *The Mental Lexicon*, *12*, 21–50.

Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language, 39*, 218–245.

Frost, R. (2006). Becoming literate in Hebrew: The grain size hypothesis and Semitic orthographic systems. *Developmental Science*, *9*, 439–440.

Frost, R., Deutsch, A., Gilboa, O., Tannenbaum, M., & Marslen-Wilson, W. D. (2000). Morphological priming: Dissociation of phonological, semantic, and morphological factors. *Memory & Cognition*, *28*, 1277–1288.

Frost, R., Forster, K. I., & 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.

Frost, R., Grainger, J., & Carreiras, M. (2008). Advances in morphological processing: An introduction. *Language and Cognitive Processes*, *23*, 933–941.

Frost, R., Kugler, T., Deutsch, A., & Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*, 1293–1326.

Gafos, D. (1998). A-templatic reduplication. *Linguistic Inquiry*, *29*, 515–527.

Gafos, A. I. (2003). Greenberg's asymmetry in Arabic: A consequence of stems in paradigms. *Language*, *79*, 317–355.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language*, *37*, 58–93.

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

Gor, K., Chrabaszcz, A., & Cook, S. (2017). Processing of native and nonnative inflected words: Beyond affix stripping. *Journal of memory and Language*, *93*, 315–332.

Gor, K., Chrabaszcz, A., & Cook, S. (2018). Early and late learners decompose inflected nouns, but can they tell which ones are inflected correctly?. *Journal of Second Language Studies*, *1*, 106–140.

Gor, K., & Jackson, S. (2013). Morphological decomposition and lexical access in a native and second language: A nesting doll effect. *Language and cognitive processes*, *28*, 1065–1091.

Gwilliams, L., & Marantz, A. (2015). Non-linear processing of a linear speech stream: The influence of morphological structure on the recognition of spoken Arabic words. *Brain and Language*, *147*, 1–13.

Hagiwara, H., Sugioka, Y., Ito, T., Kawamura, M., & Shiota, J. I. (1999). Neurolinguistic evidence for rule-based nominal suffixation. *Language*, *75*, 739–763.

Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route?. *Cognitive Psychology*, *41*, 313–360.

Hahne, A., Mueller, J. L., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience*, *18*, 121–134.

Hare, M., Elman, J. L., & Daugherty, K. G. (1995). Default generalisation in connectionist networks. *Language and Cognitive Processes*, *10*, 601–630.

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

Heyer, V., & Kornishova, D. (2018). Semantic transparency affects morphological priming… eventually. *The Quarterly Journal of Experimental Psychology*, *71*, 1112–1124.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901–931.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, *29*, 33–56.

Hopp, H. (2016). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics*, *37*, 1253–1280.

Hopp H. (2018). *The Bilingual Mental Lexicon in L2 Sentence Processing*. Available at: http://site.uit.no/lava/files/2018/01/20171230_J_SLA_Yearbook_Hopp_4.pdf [accessed July 12 2018].

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing unpredictable variation: The roles of adult and child learners in language formation and change. *Language Learning and Development*, *1*, 151–195.

Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, *60*, 877–925.

Itai, A., & Wintner, S. (2008). Language resources for Hebrew. *Language Resources and Evaluation*, *42*, 75–98.

Jacob, G., Heyer, V., & Veríssimo, J. (2018). Aiming at the same target: A masked priming study directly comparing derivation and inflection in the second language. *International Journal of Bilingualism*, *22*, 619–637.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 422–450.

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

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

Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2017). The activation of grammaticalized meaning in L2 processing: Toward an explanation of the morphological congruency effect. *International Journal of Bilingualism*, *21*, 81–98.

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences USA*, *96*, 7592–7597.

Johnson, J. S., Shenkman, K. D., Newport, E. L., & Medin, D. L. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language*, *35*, 335–352.

Justus, T., Larsen, J., Yang, J., de Mornay Davies, P., Dronkers, N., & Swick, D. (2011). The role of Broca's area in regular past-tense morphology: An event-related potential study. *Neuropsychologia*, *49*, 1–18.

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*, 30–66.

Kastner, I., Pylkkänen, L., & Marantz, A. (2018). The form of morphemes: MEG evidence from masked priming of two Hebrew templates. *Frontiers in Psychology*, *9*:2163.

Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 1371–1390.

Kazanina, N., Dukova-Zheleva, G., Geber, D., Kharlamov, V., & Tonciulescu, K. (2008). Decomposition into multiple morphemes during lexical access: A masked priming study of Russian nouns. *Language and Cognitive Processes*, *23*, 800–823.

Keating, G. D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. *Research in Second Language Processing and Parsing*, *53*, 113–134.

Keuleers, E., Sandra, D., Daelemans, W., Gillis, S., Durieux, G., & Martens, E. (2007). Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology*, *54*, 283–318.

Kim, J. J., Marcus, G. F., Pinker, S., Hollander, M., & Coppola, M. (1994). Sensitivity of children's inflection to grammatical structure. *Journal of Child Language*, *21*, 173–209.

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

Kolan, L., Leikin, M., & Zwitserlood, P. (2011). Morphological processing and lexical access in speech production in Hebrew: Evidence from picture-word interference. *Journal of Memory and Language*, *65*, 286–298.

Krause, H., Bosch, S., & Clahsen, H. (2015). Morphosyntax in the bilingual mental lexicon: An experimental study of strong stems in German. *Studies in Second Language Acquisition, 37*, 597–621.

Laks, L. (2011). *Morpho-phonological and morpho-thematic relations in Hebrew and Arabic verb formation* (Doctoral dissertation). Tel-Aviv University, Tel-Aviv.

Lardiere, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second language research, 14*, 359–375.

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes*, *9*, 295–316.

Lavric, A., Clapp, A., & Rastle, K. (2007). ERP evidence of morphological analysis from orthography: A masked priming study. *Journal of Cognitive Neuroscience*, *19*, 866–877.

Leikin, M., Schwartz, M., & Share, D. L. (2010). General and specific benefits of bi-literate bilingualism: a Russian–Hebrew study of beginning literacy. *Reading and Writing*, *23*, 269–292.

Leiner, D. J. (2014). SoSci Survey (Version 2.5.00-i) [Computer Software]. Available from https://www.soscisurvey.de

Leiner, D. J. (2016). SoSci Survey (Version 2.6.00) [Computer software]. Available from https://www.soscisurvey.de

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

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Levy, Y. (1983). The acquisition of Hebrew plurals: The case of the missing gender category. *Journal of Child Language*, *10*, 107–121.

Li, J., Taft, M., & Xu, J. (2017). The Processing of English Derived Words by Chinese-English Bilinguals. *Language Learning*, *67*, 858–884.

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

Longtin, C. M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*, 313–334.

López-Villaseñor, M. L. (2012). The effects of base frequency and affix productivity in Spanish. *The Spanish Journal of Psychology*, *15*, 505–512.

Marcus, G. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*, 189–256.

Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 175–193). Oxford: Oxford University Press.

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

Marslen-Wilson, W. D., Hare, M., & Older, L. (1993). Inflectional morphology and phonological regularity in the English mental lexicon. In *Proceedings of the 15th annual conference of the cognitive science society* (pp. 693–698). Hillsdale, NJ: Erlbaum.

Marslen-Wilson, W. D., & Tyler, L. K. (1998). Rules, representations, and the English past tense. *Trends in Cognitive Sciences*, *2*, 428–435.

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

Masson, M. E., Rabe, M. M., & Kliegl, R. (2017). Modulation of additive and interactive effects by trial history revisited. *Memory & Cognition*, *45*, 480–492.

Matthews, C. (2010). On the nature of phonological cues in the acquisition of French gender categories: Evidence from instance-based learning models. *Lingua*, *120*, 879–900.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

Maylor, B. R. (2002). *Lexical Template Morphology: Change of state and the verbal prefixes in German* (Vol. 58). John Benjamins Publishing.

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, *24*, 459–486.

McCarthy, J. J. (1981), A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry*, *12*, 373–418.

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out?. *Trends in Cognitive Sciences*, *6*, 465–472.

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

McDonald, J. L., & Roussel, C. C. (2010). Past tense grammaticality judgment and production in non-native and stressed native English speakers. *Bilingualism: Language and Cognition, 13*, 429–448.

McElree, B. (2006). Accessing recent events. In Ross, B. (Ed.), *The Psychology of Learning and Motivation* (pp. 155–200). San Diego, CA: Academic Press.

Medeiros, J., & Duñabeitia, J. A. (2016). Not everybody sees the ness in the darkness: Individual differences in masked suffix priming. *Frontiers in Psychology*, *7:*1585.

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLOS ONE, 12*, e0171935.

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

Montrul, S. (2011). Morphological errors in Spanish second language learners and heritage speakers. *Studies in Second Language Acquisition, 33*, 163–192.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning, 60*, 154–193.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience, 24*, 933–947.

Morris, J., Frank, T., Grainger, J., & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, *44*, 506–521.

Morris, J., & Stockall, L. (2012). Early, equivalent ERP masked priming effects for regular and irregular morphology. *Brain and Language*, *123*, 81–93.

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

Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition, 58*, 221–251.

Naigles, L. G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child development, 64*, 1665–1687.

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

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., & Neville, H. J. (2007). An ERP study of regular and irregular English past tense inflection. *NeuroImage*, *34*, 435–445.

Norman, T., Degani, T., & Peleg, O. (2016). Transfer of L1 visual word recognition strategies during early stages of L2 learning: Evidence from Hebrew learners whose first language is either Semitic or Indo-European. *Second Language Research*, *32*, 109–122.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Oganyan, M. (2017). *The Role of Morphology in Word Recognition of Hebrew as a Templatic Language* (Doctoral dissertation). University of Washington, Seattle.

O'Malley, S., & Besner, D. (2013). Reading aloud: Does previous trial history modulate the joint effects of stimulus quality and word frequency? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*, 1321–1325.

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

Parodi, T., Schwartz, B. D., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, 669–706.

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

Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., & Carreiras, M. (2018). Does visual letter similarity modulate masked form priming in young readers of Arabic? *Journal of Experimental Child Psychology*, *169*, 110–117.

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15.

Pinker, S. (1999). *Words and rules*. New York, NY: Harper Perennial.

Pinker, S. (2013). *Learnability and cognition: The acquisition of argument structure*. MIT press.

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*, 456–463.

Plag, I. (2000). Irregular past tense formation in English interlanguage. *Language Use, Language Acquisition and Language History. (Mostly) Empirical Studies in Honour of Rüdiger Zimmermann*, 134–149.

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

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*, 21–69.

Plunkett, K., & Nakisa, R. C. (1997). A connectionist model of the Arabic plural system. *Language and Cognitive Processes*, *12*, 807–836.

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*, 1–56.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*, 103–133.

Prior, A., & Markus, E. (2014). Morphological activation in sentence context: when the root prevails over the meaning. *Language, Cognition and Neuroscience*, *29*, 1180–1188.

Qiao, X., & Forster, K. I. (2017). Is the L2 lexicon different from the L1 lexicon? Evidence from novel word lexicalization. *Cognition*, *158*, 147–152.

Rah, A., & Adone, D. (2010). Processing of the reduced relative clause versus main verb ambiguity in L2 learners at different proficiency levels. *Studies in Second Language Acquisition*, *32*, 79–109.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, *45*, 45–94.

Rastle, K. and Davis, M. H. (2003). Reading morphologically complex words: Some thoughts from masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 279–305). New York: Psychology Press.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, *23*, 942–971.

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

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review*, *11*, 1090–1098.

Rastle, K., Lavric, A., Elchlepp, H., & Crepaldi, D. (2015). Processing differences across regular and irregular inflections revealed through ERPs. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 747–760.

Reeder, P. A., Newport, E. L., & Aslin, R. N. (2017). Distributional learning of subcategories in an artificial grammar: Category generalization and subcategory restrictions. *Journal of Memory and Language*, *97*, 17–29.

Reid, A. A., & Marslen-Wilson, W. D. (2003). Lexical representation of morphologically complex words: Evidence from Polish. *Trends in Linguistics Studies and Monographs*, *151*, 287–336.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics*, *32*, 299–331.

Rodriguez-Fornells, A., Münte, T. F., & Clahsen, H. (2002). Morphological priming in Spanish verb forms: An ERP repetition priming study. *Journal of Cognitive Neuroscience*, *14*, 443–454.

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

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In J.L. McClelland, D. E. Rumelhart, & The PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol.2, pp. 216–271). Cambridge, MA: MIT Press.

Sabourin, L., & Stowe, L. A. (2008). Second language processing: When are first and second languages processed similarly?. *Second Language Research*, *24*, 397–430.

Say, T., & Clahsen, H. (2002). Words, rules and stems in the Italian mental lexicon. In S. Nooteboom, F. Weerman, & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 93–129). Dordrecht: Kluwer.

Schmidtke, D., Matsuki, K., & Kuperman, V. (2017). Surviving blind decomposition: A distributional analysis of the time-course of complex word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1793–1820.

Schuler, K. D., Reeder, P. A., Newport, E. L., & Aslin, R. N. (2017). The effect of Zipfian frequency variations on category formation in adult artificial language learning. *Language Learning and Development*, *13*, 357–374.

Schultink, H. (1961). Produktiviteit als morfologisch fenomeen. *Forum der letteren, 2,* 110–125.

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

Selkirk, E. O. (1982). *The syntax of words*. Cambridge, MA: MIT Press.

Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory & Cognition*, *25*, 425–437.

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

Skousen, R., Lonsdale, D., & Parkinson, D. B. (2002). *Analogical modeling: An exemplar-based approach to language* (Vol. 10). Amsterdam: John Benjamins Publishing.

Smolka, E., Komlosi, S., & Rösler, F. (2009). When semantics means less than morphology: The processing of German prefixed verbs. *Language and Cognitive Processes*, *24*, 337–375.

Smolka, E., Preller, K. H., & Eulitz, C. (2014). 'Verstehen'('understand') primes 'stehen'('stand'): Morphological structure overrides semantic compositionality in the lexical representation of German complex verbs. *Journal of Memory and Language*, *72,* 16–36.

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

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

Stavrakaki, S., & Clahsen, H. (2009). The perfective past tense in Greek child language. *Journal of Child Language*, *36*, 113–142.

Steinhauer, K., White, E. J., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research, 25,* 13–41.

Taraban, R., & Kempe, V. (1999). Gender processing in native and nonnative Russian speakers. *Applied Psycholinguistics, 20,* 119–148.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition*, *27,* 173–204.

Traxler, M. J. (2005). Plausibility and verb subcategorization in temporarily ambiguous sentences: Evidence from self-paced reading. *Journal of Psycholinguistic Research, 34*, 1–30.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*, 231–270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp.141–178). Washington, DC: Georgetown University Press.

Ussishkin, A. (1999). The inadequacy of the consonantal root: Modern Hebrew denominal verbs and output–output correspondence. *Phonology*, *16*, 401–442.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language & Linguistic Theory*, *23*, 169–218.

Vaknin-Nussbaum, V., & Shimron, J. (2011). Hebrew plural inflection: Linear processing in a Semitic language. *The Mental Lexicon*, *6*, 197–244.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*, 1176–1190.

Vannest, J., Polk, T. A., & Lewis, R. L. (2005). Dual-route processing of complex words: new fMRI evidence from derivational suffixation. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 67–76.

Velan, H., Deutsch, A., & Frost, R. (2013). The flexibility of letter-position flexibility: Evidence from eye movements in reading Hebrew. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1143–1152.

Velan, H., & Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review*, *14*, 913–918.

Velan, H., & Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language*, *61*, 285–302.

Velan, H., & Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition*, *118*, 141–156.

Velan, H., Frost, R., Deutsch, A., & Plaut, D. C. (2005). The processing of root morphemes in Hebrew: Contrasting localist and distributed accounts. *Language and Cognitive Processes*, *20*, 169–206.

Veríssimo, J., & Clahsen, H. (2009). Morphological priming by itself: A study of Portuguese conjugations. *Cognition, 112*, 187–194.

Veríssimo, J., & Clahsen, H. (2014). Variables and similarity in linguistic generalization: Evidence from inflectional classes in Portuguese. *Journal of Memory and Language, 76,* 61–79.

Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition, 25*, 315–326.

Voga, M., Anastassiadis-Symeonidis, A., & Giraudo, H. (2014). Does morphology play a role in L2 processing? Two masked priming experiments with Greek speakers of ESL. *Lingvisticæ Investigationes, 37*, 338–352.

Waltke, B. K., & O'Connor, M. P. (1990). *An introduction to Biblical Hebrew syntax*. Winona Lake, IN: Eisenbrauns.

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

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

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

Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language*, *60*, 368–392.

Wray, S. (2016). *Decomposability and the effects of morpheme frequency in lexical access* (Doctoral dissertation). The University of Arizona, Tucson.

Xu, J., & Taft, M. (2015). The effects of semantic transparency and base frequency on the recognition of English complex words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 904–910.

Yablonski, M., & Ben-Shachar, M. (2016). The morpheme interference effect in Hebrew. *The Mental Lexicon*, *11*, 277–307.