

Introduction

Complex words have been the target of experimental research on word processing for the past 30–40 years. Studying polymorphemic words can contribute to a better understanding of the structure of the mental lexicon. There have been numerous studies on how morphologically complex and compound words are stored and accessed in different languages (Longtin et al., 2003 for French; Rastle et al., 2004 for English; Kazanina et al., 2008 for Russian; among others). However, very little has been done on less studied languages like Persian. The goal of the present study is to shed some light on the nature of the mental lexicon, and more specifically, on how compound verbs are represented and accessed. The study aims to assess the current theories of compound word processing in a less studied language (i.e., Persian) with a very productive compounding system.

This chapter is organized as follows. In Section 1.1, the mental lexicon will be defined followed by a description of the models and theories on lexical processing. In Section 1.2, the goal of the study will be explained, followed by the book outline in Section 1.3.

1.1 – The Mental Lexicon

The mental lexicon encompasses ‘words,’ which are employed by the speakers of a language to represent the mappings between phonological forms that are used to refer to objects in the world and their corresponding meanings. Research on the mental lexicon has mainly concentrated on lexical processing and lexical representation.

To this end, scholars, who have tried to provide a definition of the mental lexicon, have found it quite challenging, since the boundaries of the subject as well as its contents are rather vague. Jarema and Libben (2007:2) have defined the mental lexicon as, “the cognitive system that constitutes the capacity for conscious and unconscious lexical activity.”

The mental lexicon has been examined and explained through studies on the structure of the lexicon and on the lexical processing and representation for several decades. Studies on the structure of the lexicon have attracted theoretical linguists, while studies on lexical processing and representation have enticed psycholinguists.

Models of the lexicon in theoretical linguistics can be considered as lying on a continuum. At one end are theories that consider the lexicon quite rich (Lexicalist Theories) and at the other, theories that consider the lexicon quite poor (e.g., Distributed Morphology). Lexicalist Theories hold that the lexicon encompasses words, roots, affixes and word formation rules, as well as more detailed information, such as information about the argument structures of various verbs. It has been argued that, in Lexicalist Theories, where the argument structures of verbs are linked via universal principles to particular syntactic positions, the argument structure alternations are accomplished through a separate generative process that occurs within the lexicon and prior to projection (Levin and Rappaport, 1995). Distributed Morphology, on the other hand, holds that the mental lexicon only includes atomic roots, such as sound-meaning pairings (Marantz, 1997). Between these two ends and yet closer to the Lexicalist end are theories such as other Generative theories, which are not as radical as the strong Lexicalist theories and which posit a robust computational model where both the new forms and new meanings are derived through computations using morphological rules.

It must be mentioned that there are two principal versions of Lexicalist theories, i.e., strong and weak. Strong Lexicalist theories consider the lexicon quite rich and hold that syntactic transformations cannot be used in either derivational or inflectional morphology. In these theories, the mental lexicon is so rich that it includes words, roots, affixes, word-formation rules, and subcategorization frames (information about the verb arguments).

On the other hand, weak Lexicalist theories (such as Distributed Morphology) consider the lexicon quite poor and hold that syntactic transformations cannot be used in derivational morphology. In these theories, the mental lexicon only includes atomic roots, such as sound-meaning pairings.

Psycholinguists have long been interested in the processing of complex words because, through the study of polymorphemic words, they can understand the structure of the mental lexicon, while at the same time discover how words are stored and accessed from the mental lexicon. Models of the

lexicon in psycholinguistics have yielded different theories of complex word processing, ranging from decompositional to non-decompositional theories.

Taft and Forster (1975) initially put forth the idea of the decompositionality of polymorphemic words when being processed. They proposed that morphologically complex words are accessed via their constituent morphemes, for example (teacher $\rightarrow$ teach + er). This view was later challenged by Butterworth's (1983) non-decompositional full-listing approach to morphologically complex word processing, which held that polymorphemic words are stored and processed as whole words. Between the two extreme decompositional and non-decompositional approaches to morphologically complex word processing, there are mediating dual access theories that support morphological decomposition for some polymorphemic words, but non-decompositional status for others (Stanners et al., 1979; Baayen, 1992; Schreuder and Baayen, 1995; Baayen et al., 1997; among others). According to dual access theories, both decompositional and non-decompositional representations may exist simultaneously when processing complex words, while different factors, such as semantic transparency, cumulative frequency, productivity of the affix, and surface frequency would determine which lexical representation should be activated.

Experimental research has provided the field with valuable insights into how a word and its relevant information, such as its root, affixes and word formation rules, are stored in and accessed from the mind. These experiments have examined the effects of morphological relatedness (writer-write), semantic relatedness (doctor-nurse), syntactic relatedness (easily-surely), and orthographic or phonological relatedness (tape-ape) or a combination of these effects on the processing of polymorphemic words. Many studies have supported a decompositional approach to polymorphemic word processing (Longtin et al., 2003 for French; Rastle et al., 2004 for English; Fiorentino & Poeppel, 2007 for English; Kazanina et al., 2008 for Russian; among others). On the other hand, the non-decompositional approach has been claimed for some kinds of polymorphemic words, including words that are formed by morphological processes such as regular past-tense formation (Baayen et al., 1997; Sereno & Jongman, 1997; among others).

Most models of complex word processing consider abstract representations and algebraic rules to combine those representations for complex words, such as the models described above. However, there are some other models of lexical processing that challenge a computational theory of mind (e.g., Sidenberge and Gonnerman, 2000; Hay and Baayen, 2005; among

others). For example, the distributional-connectionist model of lexical processing considers morphology to be a graded, inter-level representation that reflects orthographic, phonological and semantic relationships. Another model that emphasizes orthographic, phonological or semantic representation is the symbolic dual-route model (Coltheart et al., 1993; Coltheart et al., 2001).

The next section will discuss the goals of the present study, followed by the research questions posed, and an overview of the experimental studies.

1.2 – Goals of the Present Study

The present study investigates the processing of polymorphemic words through examining compound verbs in Persian, which is a less studied language. It investigates the role of semantic transparency as well as morphological structure during lexical retrieval of noun-verb compound verbs. Semantic transparency refers to the compositionality of the compound, which differentiates between transparent (*ghazaa-xordan* ‘food-to eat’ – *GHAZAA* ‘food’ *to eat*), and opaque (e.g., *qasam-xordan* ‘oath-to eat’ – *QASAM* ‘oath’ *to swear / oath*) compounds. Morphological relatedness is broader and encompasses both semantically transparent and semantically opaque compounds. Moreover, it can differentiate between morphologically related prime-target pairs (*ghazaa-xordan* ‘food-to eat’ – *GHAZAA* ‘food’ *to eat* or *qasam-xordan* ‘oath-to eat’ – *QASAM* ‘oath’ *to swear / oath*) and orthographically related prime-target pairs (e.g., *shenaaxtan* ‘to recognize’ – *SHENAA* ‘swimming’).

Studies on morphologically complex nouns, such as *teach-er* (Longtin et al., 2003; Kazanina et al, 2008; among others) and compound nouns, such as *quarter-back* (Fiorentino, 2006; among others) abound, while studies on morphologically complex verbs are not as many (e.g., Smolka et al, 2009; Smolka et al., 2010 for German; Feldman et al., 2002 for Serbian; among others). In addition, there is a glaring lack of studies on processing compound verbs, perhaps because such structures are not common in many languages. In Persian however, compounding is a common verb formation process, which is even applied to produce novel verbs with loan nonverbal elements and Persian base verbs.

Persian compound verbs are a combination of two constituents: the non-verbal and the verbal, respectively. The nonverbal part can be a noun (e.g., *karye kardan* ‘tear-to do’ ‘noun-verb’ *to cry*), an adjective (e.g., *daagh kar-*

dan ‘hot-to do’ ‘adjective-verb’ *to get mad*), a prepositional phrase (e.g., *az bar kardan* ‘of-on-to do’ ‘prep-prep-verb’ *to memorize*), an adverbial phrase (e.g., *pish bordan* ‘forward-to take’ ‘prep-verb’ *to succeed*), or a complex nominal phrase (e.g., *sar-be-sar-gozaashtan* ‘head-to-head-to put’ *to tease*). The specific compound verbs that are examined in the present study are noun-verb compound verbs.

There are two main approaches to compound verb formation in Persian, mainly the syntactic and the lexical approaches. The syntactic approach to Persian compound verb formation entails that the syntactic category of a lexical element is determined only by the syntactic context in which it appears.

Persian compound verbs are formed by combination and incorporation, the two lexical processes involved in compound verb formation in Persian (Dabir-Moghaddam, 1997). These two processes ordinarily result in a non-compositional compound in the first one (i.e., combination, e.g., *qasam-xordan* ‘oath-to eat’ *to swear*), but a compositional compound in the latter (i.e., incorporation, e.g., *ghazaa-xordan* ‘food-to eat’ *to eat*).

In fact, Persian is an interesting language for investigating lexical representations because it belongs to the same language family as English, French and German and thus possesses the concatenative morphology of Indo-European languages. Yet, it is different from many Indo-European languages in its morphological richness and productivity as demonstrated by the productivity of Persian compound verbs in later chapters in this book.

The present study will help to compare the observed results with the experimental studies on the processing of complex and compound words in other concatenative languages. In fact, due to the morphological richness of Persian, I anticipate that the results of the current study will be similar to the data derived from other priming experiments in morphologically rich languages like German (Smolka et al., 2010; among others) but less similar to those derived from priming experiments in morphologically less rich languages such as English or French (Longtin et al., 2003; Fiorentino, 2006; among others).

The research question addressed in the present study is twofold: first, whether or not compound verb constituents show significant priming in the masked priming paradigm; and second, whether or not priming effects are constrained by semantic transparency and morphological relatedness. To answer these questions, three masked priming experiments were carried out to investigate the processing of compound verbs in Persian. Experiments 1, 2a and 2b all investigated the effect of semantic and morphological

relatedness on priming. However, the conditions in each experiment were different. In Experiment 1, the priming effect of transparent and opaque compounds on their nominal and verbal constituents was investigated. In Experiment 2a, the priming effect of transparent and opaque compounds as well as pseudo-compounds on their nominal constituents or pseudo-constituents was examined. Finally, in Experiment 2b, the priming effect of transparent and opaque compounds as well as pseudo-compounds on their verbal constituents or pseudo-constituents was investigated.

1.3 – Book Outline

Following on this introductory chapter, Chapter 2 discusses previous experimental research on complex and compound word processing. It outlines theories of polymorphemic word processing and its repercussions on theories of word processing. This is followed by relevant studies on aphasic patients, which have focused on their processing of complex words. The effects of different linguistic and nonlinguistic factors on the processing of polymorphemic words will be analyzed further. Some of these factors include prime duration and frequency effects.

Chapter 3 reports the findings of the three experiments conducted in this study, which examined how Persian compound verbs are represented and accessed while processing. The first study included one masked priming experiment, while the second encompassed two masked priming experiments; one of which concentrated on the priming effect of the first constituent, the nominal constituent or the non-head, and the other concentrated on the priming effect of the second constituent, the verbal constituent or the head.

The first study did not separate the nominal and the verbal constituents into two experiments, but rather encompassed all the conditions in one single experiment. Although the transparent and the opaque conditions were analyzed separately, it was difficult to form a definitive conclusion since several conditions were included in one single experiment. Therefore, in the second study, the experiments were organized in such a way as to distinguish the priming effects in the verbal and the nominal constituents more clearly. The experiments investigated the role of semantic transparency as well as morphological structure in on-line visual recognition of Persian noun-verb compound verbs using a constituent priming paradigm in native speakers of Persian.

Chapter 4 discusses the theoretical background of compound verbs in Persian.¹ This includes syntactic and semantic accounts of noun-verb compound verbs. This section is followed by a discussion on noun incorporation. The results of the present study will help to clarify the controversy existing in the field about whether the nominal element is incorporated into the verb or is simply an internal argument of the verb.

Chapter 5 discusses the results in light of the existing literature and presents the implications of the results for the psycholinguistic and linguistic fields. In the end, it summarizes the findings of the study and concludes with suggestions for further research.

Finally, Appendix A lists the stimuli for Experiment 1, and Appendix B lists the stimuli used in Experiment 2. Appendix C includes the sample results of minF' Tests, which will be discussed briefly in Chapter 3.