Journal of Memory and Language 65 (2011) 406–430

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds

Hongbo Ji, Christina L. Gagné *, Thomas L. Spalding

University of Alberta, Edmonton, Alberta, Canada

ARTICLE INFO

Article history:
Received 22 November 2010
revision received 30 June 2011
Available online 2 August 2011

Keywords:
Compound
Decomposition
Complex structure
Semantic transparency
Lexical processing
Morphological structure

ABSTRACT

Six lexical decision experiments were conducted to examine the influence of complex structure on the processing speed of English compounds. All experiments revealed that semantically transparent compounds (e.g., *rosebud*) were processed more quickly than matched monomorphemic words (e.g., *giraffe*). Opaque compounds (e.g., *hogwash*) were also processed more quickly than monomorphemic words. However, when the experimental materials and/or procedure encouraged decomposition/integration, this advantage disappeared. This research suggests that morphological decomposition initiated by the existence of complex structure results in the availability of both the lexical and semantic representations of compound constituents, regardless of whether the compounds are transparent or opaque, and that meaning composition is attempted. This meaning composition further speeds up transparent compound processing beyond lexical facilitation but slows down opaque compound processing because the computed meaning for opaque compounds conflicts with the retrieved meaning.

Introduction

How do people recognize morphologically complex words such as *snowball* and *successor*? One possibility is that complex words are represented in the mental lexicon as whole word units, as are monomorphemic words such as *giraffe* (e.g., Butterworth, 1983). However, in the past three decades, ample evidence has shown that complex words can also be accessed via their constituents (Andrews, 1986; Andrews, Miller, & Rayner, 2004; Hyönä & Pollatsek, 1998; Juhasz, Starr, Inhoff, & Placke, 2003; Libben, 1998; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Pollatsek & Hyönä, 2005; Pollatsek, Hyönä, & Bertram, 2000; Taft, 1979, 1994; Taft & Forster, 1975, 1976; Zwitserlood, 1994). A fundamental underlying assumption behind most theories of complex word processing is that there is a processing cost associated with using a decomposition-based route; consequently, direct access is assumed to be the faster and more efficient processing option. The current research examines the validity of this assumption in the context of compound words. Our particular aim is to evaluate the relative costs and benefits associated with the decomposition of a compound into its constituents and with the subsequent integration of those constituents.

A guiding principle behind the architectures proposed for the processing of complex words has been the balance of lexical storage vs. morphological computation. Many papers in the literature on complex word processing have been explicitly concerned with establishing how the balance of storage and computation affects lexical processing (e.g., Baayen, 2007; Bertram, Laine, Baayen, Schreuder, & Hyönä, 2000; Kuperman, Bertram, & Baayen, 2010; Kuperman, Schrueder, Bertram, & Baayen, 2009; Libben, 2005). A wide range of theories of complex word processing have emerged as the various theories have sought to balance the demands associated with storage vs. computation, and these theories differ substantially in terms of the extent

* Corresponding author. Address: Department of Psychology, University of Alberta, P-217 Biological Sciences Building, Edmonton, Alberta, Canada T6G 2E9. Fax: +1 780 492 1768.

E-mail address: cgagne@ualberta.ca (C.L. Gagné).

0749-596X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2011.07.003

to which the morphological structure plays a role in the processing (i.e., the comprehension or production) of complex words. Some theories take a full-listing stance, in which all words are stored, and posit that morphological structure plays no (or a minimal) role (Lukatela, Carello, & Turvey, 1987; Manelis & Tharp, 1977). Other theories posit that morphological structure is involved (Chialant & Caramazza, 1995; Dell, 1986; Frauenfelder & Schreuder, 1991; Laudanna & Burani, 1985, 1995; Schreuder & Baayen, 1995; Taft & Forster, 1975, 1976). The latter theories vary in terms of the point at which the constituents' representations become available (see Kuperman et al. (2010) for an excellent overview).

The literature on the processing of complex words contains an interesting paradox. Two common assumptions underlying the discussion of the balance of storage and computation is that morphological parsing is time-consuming and that direct access of full-form representations is a way of reducing computational load (see, e.g., Bertram, Laine, & Karvinen, 1999). Therefore, of the two routes (direct access and decomposition), the direct access route is generally thought to be the preferable route (e.g., "As decomposition is assumed to be slower and more error-prone than full-form processing, it is economical to represent words that are used often as full forms since that guarantees effective access to them", Lehtonen et al., 2007, p. 124). Yet, despite the cost associated with decomposition, there are numerous findings which suggest that morphological decomposition routinely occurs during the processing of complex words (Andrews, 1986; Andrews et al., 2004; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Libben, 1998; Marslen-Wilson et al., 1994; Pollatsek & Hyönä, 2005; Pollatsek et al., 2000; Taft, 1979, 1994; Taft & Forster, 1975, 1976; Zwitserlood, 1994).

Given that morphological decomposition is both costly and yet commonly used, it is essential to consider both the costs and benefits associated with decomposition. To do so, it is helpful to examine the extent to which morphological decomposition influences the processing time of complex words because it is not obvious *a priori* whether decomposition helps or hinders processing overall. Some theories predict that the decomposition route races with a direct access route and that response time on any given trial is determined by the route with the faster completion time (Baayen, Dijkstra, & Schreuder, 1997; Bertram et al., 1999; Frauenfelder & Schreuder, 1991; Schreuder & Baayen, 1995). Consequently, response times should be faster for complex words than for frequency-matched monomorphemic words. Other theories suggest that the decomposition and direct-access routes compete for cognitive resources which increases the time required to process complex words (Laine, Vainio, & Hyönä, 1999).

For the most part, research addressing this issue has used either inflected or derived words. Some studies have found that complex words exhibit processing difficulties relative to comparable monomorphemic words (Bertram et al., 1999; Hyönä, Vainio, & Laine, 2002; Laine et al., 1999; Lehtonen et al., 2007). Other studies have found a processing advantage for complex words relative to monomorphemic words (Bertram et al., 1999; Burani & Thornton, 2003; Hudson & Buijs, 1995). Thus, it does not appear to be the case that complex words (which can be accessed via either route) are necessarily more difficult to process relative to monomorphemic words (which can only be accessed via the presumably faster direct-access route).

The focus of our investigation will be on compound words because they provide an especially useful test case for evaluating questions about the costs and benefits associated with decomposition. First, compound words are especially likely to be decomposed because they tend to be low frequency (e.g., 58% of noun–noun compounds in the English CELEX have written frequencies of less than one in a million, Libben, 2005). Indeed, there are several studies showing that the constituents of a compound are available, at least at the lexical level (e.g., Andrews, 1986; Andrews et al., 2004; Fiorentino & Poeppel, 2007; Giraudo & Grainger, 2001; Hyönä & Pollatsek, 1998; Juhasz et al., 2003; Pollatsek & Hyönä, 2005; Pollatsek et al., 2000; Sandra, 1990; Zwitserlood, 1994). Second, relative to inflected and derived words, compound words are likely to have increased costs associated with decomposition. Inflections and derivations consist of one independent meaning unit plus an affix. Because affixes are a closed-class set, parsing can be accomplished by stripping affixes from stems (Taft & Forster, 1975). In contrast, the constituents of compounds are an open-class set and, consequently, there is no reliable heuristic for parsing. Moreover, all possible parses of a compound appear to be attempted; for example, *clamprod* is parsed as both *clam-prod* and *clamp-rod* (Libben, 1994; Libben, Derwing, & Almeida, 1999). Thus, "compound words present a much greater challenge to on-line morphological parsing than do affixed words" (Libben, 2005, p. 270).

Studies comparing the processing speed of compound and monomorphemic words have found mixed results. Inhoff, Briihl, and Schwartz (1996) presented compound words (e.g., *blueberry*), suffixed words (e.g., *ceaseless*), and monomorphemic controls (e.g., *arthritis*) in neutral sentence contexts and found that compound words received longer first fixation durations than suffixed and monomorphemic words. In an online naming task, however, naming latencies were shorter for compounds than for suffixed and monomorphemic words. Using a sentence-reading paradigm, Juhasz (2006) found that compound words with a high frequency first constituent had shorter first fixation times and gaze durations than did frequency and length matched monomorphemic words. However, compound words with a low frequency first constituent did not differ from monomorphemic words. In addition, whereas Juhasz, White, Liversedge, and Rayner (2008) found no processing differences between compound and monomorphemic words, Drieghe, Pollatsek, Juhasz, and Rayner (2010) found an advantage for compound words. Finally, Fiorentino and Poeppel (2007) found an advantage for compound words in both lexical decision times and in their ERP data.

Overall, the existing data suggest that despite the additional cost associated with morphological decomposition, the availability of the constituents can, in some cases, yield processing advantages. These advantages have been attributed to the lexical access of the constituent representations facilitating access of the whole word representation