Using Conceptual Combination Research to Better Understand Novel Compound Words*

Christina L. Gagné & Thomas L. Spalding

Interpreting a novel modifier-noun phrase (e.g, licorice guitar) involves both the conceptual and lexical systems; one must access the concepts denoted by the words and select a relation (e.g., noun MADE OF modifier) to form a unified conceptual representation. We describe recent empirical work that demonstrates that the ease of interpreting a novel noun phrase is influenced by the availability of the required relation, and discuss the implications of this research for the processing of both novel and familiar compounds.

Key words: conceptual combination, interpretation, novel word, relation

How do people derive the meaning of novel words? Some novel words consist of a single unit (e.g., *dax*). The meaning of such novel words must be derived from the syntactic, semantic, and pragmatic context of their use, or by direct instruction (e.g., reading or being told a definition), or sometimes by analogy with similar words. Some novel words (e.g, *jazzercise*) can be understood through the known word and the known derivational or inflectional addition, as well as all the mechanisms available for single-unit words. Often, however, novel words are composed of two or more morphemes which are themselves words (e.g., *grasscord*). In English, nominal compounding is extremely productive and is a common way of introducing new words into the lexicon. This set of words, the compounds, is our focus in this paper. In particular, we examine the extent to which the processing of such words is guided by processes that are held in common with the interpretation of modifier-noun phrases.

Before continuing, it is necessary to clarify terminology because different terms are used in the psychological and linguistic literature. The creation of modifier–noun phrases such as *tofu bun* is often referred to as conceptual combination by cognitive psychologists and as nominal compounding by linguists. In both cases, two (or more) words are used to denote a single concept. The term lexicalized compound refers to a compound that is familiar and has a common usage. The meaning of lexicalized compounds likely can be retrieved directly from the lexicon. In contrast, the meaning of novel or less familiar combinations must be computed because the meaning has not yet been established. Although we make a distinction between lexicalized and novel compounds throughout this paper, we do not intend to imply that all modifier–noun phrases can be classified into one of these two categories. Instead, we propose that compounds are best viewed as a continuum that ranges from highly familiar compounds (e.g., *snowman*) to novel combinations (e.g., *sandpie*) rather than a dichotomy of lexicalized versus novel compounds. Our primary claim is that there are important commonalities between the processing of lexicalized and novel compounds.

We will begin by reviewing some factors that influence the processing of lexicalized compounds. First, processing is affected by whether there is a space between the constituents of a compound. Inhoff, Radach, and Heller (2000) examined the effect of spacing on eye-fixations for German compounds and found that interword spacing facilitated access to

* This work is supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

constituent word forms, but hampered the creation of a unified compound meaning, as indicated by longer final fixation times for spaced compounds. Inhoff et al. (2000) suggested that two processes are involved in the interpretation of compounds. One process involves accessing the constituent word forms (and is aided by the presence of a space) and the other process involves integrating the constituents (and is hindered by the presence of a space). Juhaz, Inhoff, & Rayner (2005) demonstrated that these results extend to English compounds. The insertion of a space facilitated the participants' ability to process the constituents, whereas the lack of a space benefited the specification of the compound's meaning.

Second, and most generally, there is evidence of decomposition in compound processing. Properties associated with the constituents influence processing of the compound. For example, recognition of a compound seems to involve the recognition of the constituent word forms (Andrews 1986; Inhoff, Briihl, & Schwartz 1996; Lima & Pollatsek 1983; Zwitserlood 1994). Recent exposure to a compound word facilitates the subsequent processing of its constituents (Masson & MacLeod 1992; Weldon 1991; Whittlesea & Brooks 1988). For example, Weldon (1991) found that the identification of the target word *black* was facilitated by prior exposure to either *blackbird* or *blackmail*. Likewise, Masson and MacLeod (1992) demonstrated that a constituent (e.g., *break*) is more accurately identified when it has been seen as part of a noun phase (e.g., *coffee break*) than when it has not been previously studied. In addition, processing of compounds is faster when they have been preceded by one of the compound's constituents (Libben, Gibson, Yoon, & Sandra 2003; Jarema, Busson, Nikolova, Tsapkini, & Libben 1999). These findings suggest that processing of the compound involves accessing the lexical representation of constituent words and that compound words can be identified via their constituents.

Is it necessarily the case that the conceptual representation of the constituents are also accessed? The answer appears to depend on whether the compound is semantically transparent. Transparency refers to the extent to which a compound's meaning is predictable from its constituents. For example, the contribution of *blue* and *berry* to the meaning of *blueberry* is clear. Although the contribution of *berry* to the meaning of *raspberry* is clear, the contribution of *rasp* is not so transparent. Thus, *blueberry* is fully transparent transparent compound, whereas *raspberry* is a partially transparent compound. The compound *humbug* is completely opaque because neither *hum* nor *bug* is related to the compound's meaning.

Semantic transparency affects the amount of cross-activation between the constituent representations and the compound representation. Sandra (1990) preceded Dutch compounds with a word that was semantically related to one of the compound's constituents; for example, *milkbottle* was preceded by *cow*. Transparent compounds were aided by prior exposure to a semantically related word, but opaque compounds were not. Likewise, Zwitserlood (1994) examined whether exposure to compound words affects the ease of processing semantic associates of either the first or second constituents. Semantically related words were faster following transparent and partially opaque compounds but not following fully opaque compounds.

Taken together, these studies suggest that the meaning of the constituents of opaque compounds might not be available during the processing of compound words, even though the lexical forms are retrieved. In contrast, the constituents of transparent compounds are represented (and available) at both the lexical and conceptual level. This highlights the need to distinguish between lexical and conceptual information in compound word processing. Indeed, such a distinction has already been incorporated into several theories of compound processing (e.g., Libben 1998; Zwitserlood, Bolwiender, & Drews 2005).

In summary, the literature on the processing of compounds appears to indicate that the constituents of the compound have important effects on the processing of the compound. This literature suggests that compounds are not processed in a way that is completely analogous to matched monomorphemic words, even for quite familiar lexicalized compounds. Let us now turn to the question of how the constituents might affect the ability to derive meanings for compounds.

An obvious advantage of compounding as a way of introducing new words is that the comprehender is likely to know the words that are now the constituents of the compound, and knowledge of a compound's constituents can be used to help determine the meaning of the compound as a whole. For example, the head noun (the second constituent in English compounds) often provides knowledge about the category, so a person can easily guess that a *beach ball* is a kind of ball. However, people know more than just the category label. After all, if all they knew was the category of the head noun, there would be no reason to use a modifier --they could just use *ball*. They also seem to know more than just that it is a ball that is in some way related to the concept beach. In particular, they derive a more detailed meaning, such as "a ball to use at the beach".

In the remainder of this paper, we will focus on what knowledge people might be drawing on to derive this more detailed representation. In particular, we propose that people are identifying a relation that is used to create a unified representation, and that deriving the meaning of a compound is affected by the ease which which this relation can be identified. We will briefly review some findings on the use of relations in novel compounds, then we will discuss some recent findings that suggest that similar processes occur in the processing even of lexicalized compounds, and we will end with a brief discussion of how the novel to lexicalized shift might best be conceptualized for compounds.

Novel compounds are compounds that are not part of the language, but can be interpreted using knowledge of the constituents as well as knowledge about how the concepts corresponding to the constituents can be combined. An important aspect of the processing of compounds and modifier-noun phrases is establishing the relation that forms the basis for the representation of the whole word. To illustrate, the novel compound *grasscord* is constructed using the relation *noun MADE OF modifier* to link the constituents *grass* and *cord*. The resulting representation denotes a subcategory of *cord*. To combine concepts, we propose that people draw on knowledge about how concepts can be combined with other concepts and that availability of a particular relation affects the ease with which novel compounds and phrases can be interpreted (Gagné 2000; 2001; 2002; Gagné & Shoben 1997; 2002). In particular, the more highly available a relation is, the easier it is to interpret the phrase. In this way, relation availability provides an important constraint on what conceptual knowledge is used during conceptual combination.

This process of selecting a relation and constructing a unified representation is obligatory for novel compounds and phrases (e.g., *beach beverage*) because the compound is not part of the lexicon. Recently, we (Gagné & Spalding 2004; 2006) have argued that this process is also obligatory for familiar compounds and have proposed that accessing familiar compounds involves exactly the same processing as computing the meaning of completely novel compounds, despite the widely accepted notion that the meaning of familiar compounds is retrieved from the mental lexicon but novel compound meanings are not. That is, we do not believe that the relation selection process disappears completely, even for familiar, lexicalized compounds. Thus, we propose that when an individual first encounters *teapot* they must also construct the meaning using the relation *FOR* (assuming they know *tea* and *pot*). Furthermore, we suggest that this process continues to occur whenever *teapot* is encountered, even after *teapot* is lexicalized.

One reason to suspect that the interpretation of familiar compounds might involve an obligatory conceptual combination process is the work on semantic transparency discussed above. The fact that transparent compounds (in which the overall meaning is clearly related to the meaning of the constituents) are processed differently than non-transparent compounds implies that the derivation of the meaning of the whole is assisted in some way by the meaning of the constituents. Relation selection is one mechanism that could be involved.

A second reason comes from research on compound processing by those with language impairments. For example, aphasiological evidence discussed by Jarema (2006, see also Libben 1998) is consistent with the proposal that various meanings compete for selection even for known compounds. It appears that the mixed aphasic in the study reported in Libben (1998) was unable to inhibit conceptual representations during the processing of compound words. For example, the person paraphrased *blueprint* as "a print that is blue." Note that in this case, the compound is being processed as though it were a novel compound and that the relation *noun IS modifier* is highly frequent for the modifier *blue*. We suggest that this combination process always occurs during the processing of compounds, even for non-aphasics, but that the conventional (i.e., lexicalized) meaning is more available and, thus, can usually effectively compete with alternative meanings.

A third reason comes from recent demonstrations of relation priming with lexicalized compounds. Gagné and Spalding (2004) manipulated relation availability by presenting target combinations after a prime combination containing the same modifier. The same relation prime (e.g., *snowfort*) used the same relation as the target (e.g., *snowball*) and the different relation prime (e.g, *snowshovel*) used a different relation. Participants performed a sense-nonsense judgment task for both the primes and the targets. We found that it took less time to respond to the target compound when preceded by the same relation compound than by the different relation compound. We obtained the same result when a lexical decision task was used (Gagné & Spalding 2004). These results are consistent with those found using novel compounds (e.g., Gagné 2001). We argued that relation priming occurs because recent exposure to a compound facilitates the selection of that relation by increasing its relative availability. Thus, even though the compounds used in this experiment presumably are lexicalized and have a representation in the mental lexicon, the ease of processing of the target compound was affected by exposure to the prime compound.

A fourth reason comes from data recently published by Gagné, Spalding, and Gorrie (2005). One aim of this study was to determine whether the ability to interpret familiar phrases (e.g., *bug spray*) was affected by recent exposure to an alternative, innovative meaning. For example, participants viewed the phrase as part of a sentence that was consistent with the established meaning (e.g., *Because it was a bad season for mosquitoes, Debbie made sure that every time she went outside, she wore plenty of bug spray*) or with the innovative meaning (e.g., *As a defense mechanism against predators, the Alaskan beetle can release a deadly bug spray*). Immediately after viewing this prime sentence, the participants viewed the target phrase (*bug spray*) with either the established definition (e.g., *a spray for bugs*) or the innovative meaning (e.g., *spray by bugs*) and indicated whether the definition was plausible. They were told that the definition did not have to be the best definition, but that they should indicate "yes" if the definition was plausible. Filler items that did not have plausible definitions were included in the studies. Of most interest is the effect of the prime sentence on the percentage of plausibility judgments. When the sentence used the established meaning, the established definition was judged plausible 89% of the time. However, when the sentence used the innovative meaning, the established definition was judged plausible only 64% of the time. In terms of response time, participants took longer to indicate that the established definition was plausible when the sentence supported the innovative meaning than when it supported the conventional meaning. These findings suggest that the established meaning was competing with the innovative meaning constructed in the previous sentence and this competition decreased the availability of the established meaning.

Each of the last three points suggests that relation selection processes occur even in the processing of familiar compounds. It may be the case that these processes are normally invisible. That is, accessing the established meaning of the compound may usually be so quick that the relation selection process need not reach completion. Thus, one only sees the relation selection process when the established meaning is difficult to access (e.g., in aphasia or when a conflicting meaning has just been activated) or when the relation selection process is speeded (e.g., in priming paradigms). In any case, it appears that the same relation selection process that occurs for novel compounds is still there, lurking just below the surface during the processing of lexicalized compounds. Consequently, understanding the processing of novel compounds provides valuable insight into the processing of familiar compounds, as well.

Finally, we return to the notion that compound words are best viewed as a continuum that ranges from completely novel to highly familiar. In particular, what does this imply for the theories of compound processing? The most directly way of incorporating this continuum is to view compound processing as consisting of two independent parallel processes. Following Logan's (1990) instance model which proposes that responses in tasks such as lexical decision can be based either on direct retrieval of information from memory or on an algorithmic route, we suggest that when a compound is encountered, the language system attempts to locate a unified representation of the compound, as well as to derive the meaning of the compound based on the constituents and conceptual knowledge about how the constituents are related (as though it were a novel compound). In effect, these two methods of understanding a compound word race against each other. For completely novel compounds, the only way to derive the meaning is via the algorithmic (conceptual combination) route and thus this route always "wins" the race. As a compound becomes more familiar, both processing routes are affected. The representation for the whole compound becomes more available and easier to activate, and, thus, the direct access route takes less time to complete. In addition, the compositional route is faster due to prior experience with processing the compound; retrieval of prior processing episodes for the same compound reduces the time required to derive the meaning of the compound via the compositional route.

This viewpoint is similar to dual-route models of lexical access (such as Juhasz, Inhoff, & Rayner 2005, and Pollatsek, Hyönä, & Bertram 2000) which also posit a direct access and a componential route. However, these previous accounts are focused on familiar compounds that already exist in the mental lexicon, whereas our account extends to (and uses the same framework for) both familiar and novel phrases. In addition, for the compositional route, we focus on the role of the conceptual system (and the use of relational information) in the creation of new unified representations rather than primarily on constituents' ability to access the compound's representation in the lexicon. In general, our framework places more emphasis on the role of the conceptual system than do previous frameworks, and in doing so hopes to explain the interpretation of novel and familiar words within the same framework.

To conclude, understanding how novel words are interpreted is important for both conceptual and linguistic purposes. Research on this issue helps researchers to understand how new conceptual representations are created within the conceptual system, and to understand how new representations are constructed within the lexical system. We have outlined one way in which familiar and novel compounds are similar. In particular, both can interpreted via a compositional "meaning creation" route which involves relation selection (as one component). It remains to be seen whether this framework extends to noncompound words; perhaps, the processing of such words, even monomorphemic words, also involves both the retrieval of a lexicalized representation (as is commonly assumed by linguists and psycholinguists) as well as a "meaning creation" process (as we have suggested is the case for familiar compounds). Even though noncompound words often lack internal units that can be combined (except in the case of multi-morphemic words such as *learner, learn* + *er*), meaning creation for noncompound words could be based on factors such as syntactic, semantic, and pragmatic context that are external to the word itself.

References:

ANDREWS, Sally. 1986. Morphological influences on lexical access: Lexical or nonlexical effects? In *Journal of Memory & Language,* vol. 25, pp. 726-740.

GAGNÉ, Christina L. 2000. Relation-based combinations versus property-based combinations: A test of the CARIN theory and the dual-process theory of conceptual combination. In *Journal of Memory & Language,* vol. 42, pp. 2000.

GAGNÉ, Christina L. 2001. Relation and lexical priming during the interpretation of noun-noun combinations. In *Journal of Experimental Psychology: Learning, Memory, & Cognition,* vol. 27, pp. 236-254.

GAGNÉ, Christina L. 2002. Lexical and relational influences on the processing of novel compounds. In *Brain and Language,* vol. 81, pp. 723-735.

GAGNÉ, Christina L., SHOBEN, Edward J. 1997. Influence of thematic relations on the comprehension of modifier-noun combinations. In *Journal of Experimental Psychology: Learning, Memory, & Cognition,* vol. 23, pp. 71-87.

GAGNÉ, Christina L., SHOBEN, Edward J. 2002. Priming relations in ambiguous noun-noun combinations. In *Memory & Cognition,* vol. 30, pp. 637-646.

GAGNÉ, Christina L., SPALDING, Thomas L. 2004. Effect of relation availability on the interpretation and access of familiar noun-noun compounds. In *Brain and Language,* vol. 90, pp. 478-486.

GAGNÉ, Christina L., SPALDING, Thomas L., GORRIE, Melissa C. 2005. Sentential context and the interpretation of familiar open-compounds and novel modifier-noun phrases. In *Language & Speech,* vol. 48, pp. 203-221.

GAGNÉ, Christina L., SPALDING, Thomas L. 2006. Conceptual combination: Implications for the Mental Lexicon, In *The representation and processing of compound words,* LIBBEN, G. and JAREMA G., New York : Oxford University Press, 2006, pp. 145-168.

INHOFF, Albrecht W., BRIIHL, Deborah, SCHWARTZ, Jill. 1996. Compound word effects differ in reading, on-line naming, and delayed naming tasks. In *Memory & Cognition,* vol. 24, pp. 466-476.
INHOFF, Albrecht W., RADACH, Ralph, HELLER, Dieter. 2000. Complex compounds in German: Interword spaces facilitate segmentation but hinder assignment of meaning. In *Journal of Memory and Language,* vol. 42, pp. 23-50.

JAREMA, Gonia, BUSSON, Celine, NIKOLOVA, Rossitza, TSAPKINI, Kyrana, LIBBEN, Gary. 1999. Processing compounds: A cross-linguistic study. In *Brain and Language,* vol. 68, pp. 362-369.

JAREMA, Gonia. 2006. Compound representation and processing: A cross-language perspective, In *The representation and processing of compound words,* LIBBEN, G. and JAREMA G., New York : Oxford University Press, 2006, pp. 45-70.

JUHASZ, Barbara J, INHOFF, Albrecht W, RAYNER, Keith. 2005. The role of interword spaces in the processing of English compound words. In *Language & Cognitive Processes,* vol. 20, pp. 291-316.

LIBBEN, Gary. 1998. Semantic transparency in the processing of compounds: Consequences for representation, processing, and impairment. In *Brain and Language,* vol. 61, pp. 30-44.

LIBBEN, Gary, GIBSON, Martha, YOON, Yeo Bom, SANDRA, Dominiek. 2003. Compound fracture: The role of semantic transparency and morphological headedness. In *Brain and Language,* vol. 84, pp. 50-64.

LIMA, Susan D., POLLATSEK, Alexander. 1983. Lexical access via an orthographic code? The basic orthographic syllabic structure (BOSS) reconsidered. In *Journal of Verbal Learning and Verbal Behavior,* vol. 22, pp. 310-332.

LOGAN, Gordon. 1990. Repetition priming and automaticity: Common underlying mechanisms? In *Cognitive Psychology,* vol. 22, pp. 1-35.

MASSON, Michael, MACLEOD, Colin. 1992. Re-enacting the route to interpretation: Enhanced perceptual identification without prior perception. In *Journal of Experimental Psychology: General,* vol. 121, pp. 145-176.

POLLATSEK, Alexander, HYÖNÄ, Jukka, BERTRAM, Raymond. 2000. The role of morphological constituents in reading Finnish compound words. In *Journal of Experimental Psychology: Human Perception & Performance,* vol. 26, pp. 820-833.

SANDRA, Dominiek. 1990. On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. In *The Quarterly Journal of Experimental Psychology,* vol. 42A, pp. 529-567.

WELDON, Mary, S. 1991. Mechanisms underlying priming on perceptual tests. In *Journal of Experimental Psychology: Learning, Memory, & Cognition,* vol. 17, pp. 526-541.

WHITTLESEA, Bruce W., BROOKS, Lee R. 1988. Critical influence of particular experiences in the perception of letters, words, and phrases. In *Memory & Cognition,* vol. 16, pp. 387-399.

ZWITSERLOOD, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. In *Language and Cognitive Processes,* vol. 9, pp. 341-368.

ZWITSERLOOD, Pienie, BOLWIENDER, Agnes, DREWS, Etta. 2005. Priming morphologically complex verbs by sentence contexts: Effects of semantic transparency and ambiguity. In *Language & Cognitive Processes,* vol. 20, pp. 395-415.

Christina Gagné
Department of Psychology
University of Alberta
P217 Biological Sciences Building
Edmonton Alberta Canada
T6G 2E9
e-mail: cgagne@ualberta.ca

and

Thomas Spalding
Department of Psychology
University of Alberta
P217 Biological Sciences Building
Edmonton Alberta Canada
T6G 2E9
e-mail: spalding@ualberta.ca