An Investigation of Combinational Productivity for Abstract and Concrete Nouns

Phil Maguire (phil.maguire@ucd.ie)
School of Computer Science and Informatics,
University College Dublin, Ireland

Edward J. Wisniewski (edw@uncg.edu)
Department of Psychology,
University of North Carolina at Greensboro, USA

Gert Storms (gert.storms@psy.kuleuven.ac.be)
Department of Psychology,
University of Leuven, Belgium

Abstract

Although various theories of conceptual combination have been proposed in the past, these models have addressed interpretation issues rather than addressing the circumstances in which combinations are used. As a result, existing theories make no explicit predictions about what kind of nouns will be used more often in combination and why this might be the case. In this study we address the issue of combination use and investigate whether differences in productivity exist for two very different types of noun, namely abstract and concrete nouns. The slot-filling view of conceptual combination (e.g. Wisniewski, 1997) motivated our hypothesis that differences in combinational preference for both the modifier and head role would be observed between these two types of noun. We extracted combinations involving a sample of abstract and concrete nouns from the British National Corpus and obtained type and token frequencies for how often each noun occurred in both the modifier and head positions. Results revealed that abstract nouns were more likely to occur in the head position while differences for concrete nouns were only partially supported. We also discovered that abstract nouns occurring in the BNC were far less likely to occur in combination than concrete nouns. We interpret these results in light of what they reveal about the use of compound phrases and discuss the implications for existing theories of conceptual combination.

Keywords: Conceptual combination; noun-noun compounds; abstract concepts; slot-filling theory.

Introduction

Conceptual combination is commonly adopted by speakers in order to refer to novel concepts and ideas and as a result, compound phrases are abundant in everyday conversation (e.g. *holiday tension*, *picnic beer*). This ubiquity is also evidenced by the number of lexicalised combinations entering the English language (e.g. *memory stick*, *laptop computer*). Recently, noun-noun compounding has generated much interest within cognitive science because of the way it crosscuts such diverse disciplines as linguistics, psychology and artificial intelligence. This phenomenon has the potential to reveal much about conceptual representation as well as language production and comprehension in general.

In English, a language in which compounding is particularly productive, the simplest combinations consist of a modifier followed by a head noun. Usually, the head noun denotes the main category while the modifier implies a relevant subcategory or a modification of that set's typical members. In this way, a *cereal box* is interpreted as a particular type of box, and more precisely as one that contains cereal as opposed to other types of boxes. While existing theories of conceptual combination have dealt primarily with the interpretation process, little focus has been directed towards understanding the circumstances in which combinations are used. As a result, existing theories do not make explicit predictions about which types of nouns are more likely to be used in combination or whether particular nouns will exhibit a preference for either the modifier or head position.

Relation-based theories such as the Competition Among Relations In Nominals (CARIN) theory (Gagné & Shoben, 1997) assert that interpretation occurs when a relation linking the constituent nouns is selected. According to CARIN, the difficulty associated with this selection process is determined by how modifiers have been used in previously encountered combinations. However, the theory makes no predictions about combinational productivity as it is concerned solely with ease of interpretation and makes no reference to why combinations are used. On the other hand, slot-filling theories such as the dual-process theory (Wisniewski, 1997) and the concept specialization model (Murphy, 1988) relate combination function to conceptual content. These theories assume that during combination, the modifier concept fills a 'slot' in the head concept, thereby modifying or 'specializing' that concept. Thus, in the case of *coffee bowl*, the *coffee* concept fills the <CONTAINS> slot in *bowl* and describes the purpose of the bowl. Although slot-filling theories do not make explicit predictions about how nouns will be used in combination, they suggest that the most productive head concepts will be those that have many slots. Furthermore, they imply that the most productive nouns in the modifier position will be those that can modify slots in a wide variety of head concepts.

Abstract versus Concrete Nouns

While concrete concepts represent physical entities defined by spatial boundaries and perceivable attributes, abstract concepts are elusive to sensory experience. Such concepts can include references to personality traits, emotions, cognitive processes, events or thoughts. Due to their lack of a perceptual representation, abstract concepts can exhibit considerable variability of their manifestations across situations (Galbraith & Underwood, 1973). Wiemer-Hastings & Xu Xu (2005) found that the properties of abstract nouns were significantly less specific than those for concrete nouns and that these properties were more likely to express subjective experiences.

Experiment

Because of this distinction in the processing of abstract and concrete nouns, we predicted that these two types of noun would show differences in combinational productivity. Slot-filling theories imply that nouns with many slots should be productive in the head noun position while nouns with fewer slots should be less productive. Abstract concepts exhibit greater variability in manifestation, therefore we would expect such concepts to have an abundance of slots and hence to be more productive in the head role. Presumably, the ambiguity of abstract nouns should also reduce their effectiveness in specializing other concepts. In contrast, concrete nouns are far more specific and should have a smaller set of associated slots. For example, the concrete noun *banana* is relatively specific and hence will not benefit from specialization to the same extent as an abstract concept like *problem*, which is less specific.

In light of this, we investigated differences in combinational productivity for abstract versus concrete nouns, with the slot-filling view motivating our hypothesis that different patterns in combinational productivity would be observed for these two types of noun. Our prediction was thus that abstract nouns would appear significantly more often as heads than as modifiers and that the opposite pattern would be observed regarding concrete nouns.

Method

Materials We selected abstract and concrete nouns from the MRC2 database (Coltheart, 1981), which contains a large sample of nouns rated by concreteness. We sampled 748 nouns from the extremes of the concreteness ratings, selecting the 374 least concrete and the 374 most concrete nouns. The least concrete nouns sampled referred to situation-embedded instances of events, traits, actions and thoughts such as *ingratitude*, *disquiet* and *chaos*. The most concrete nouns referred to tangible entities like *cigarette*, *garden* and *wind*.

Procedure A corpus analysis was required in order to determine the combinational productivity of the 748 nouns contained in our sample. For this purpose we availed of the British National Corpus (BNC), a tagged, annotated corpus containing over 100 million words. The BNC is designed to represent a wide cross-section of modern English and therefore includes a comprehensive sample of both written and spoken language. Although the corpus contains part-of-speech tagging, this feature is not adequate for accurately extracting combinations. Lapata and Lascarides (2003) estimated that up to 30% of all noun-noun phrases extracted in this way are not true conceptual combinations. This is because the BNC tagging suffers from several weaknesses, most notably the high rates of error in assigning parts of speech (e.g. "the *mountain rose* up before them"), and the inability to differentiate between legitimate combinations and noun-noun co-occurrences (e.g. "last *year houses* were cheaper"). In order to increase our accuracy, we obtained a version of the BNC that had been parsed using the Charniak parser[1]. Tagging simply assigns part-of-speech by means of statistical frequencies applied over a limited window of words, with no regard for overt grammatical rules. In contrast, parsing involves the structural analysis of each sentence as a whole, with respect to a given formal grammar. Consequently, the parsed version of the BNC was better able to distinguish between compound phrases and other noun-noun co-occurrences, thereby reducing error rates. All compound noun phrases consisting of two nouns were extracted from the parsed output using an algorithm programmed in LISP[2].

Some acronyms, misspellings, common nouns and errors remained and additional filtering was required to eliminate these. We discarded all combinations containing proper nouns, plural modifiers and nouns made up of fewer than three letters. We also removed any nouns containing hyphens, numerical digits or any form of punctuation. In order to guarantee that all remaining combinations consisted of valid nouns, we compared our set with the database of nouns included in WordNet, the online lexical reference system. Following this procedure, much of the remaining error could be attributed to what we termed 'noise sponges', words that, because of their nature, triggered a disproportionate number of false positives. Many such sponges were nouns that could double as adjectives, verbs or adverbs (e.g. "it was a *light snack*", "the children *dread school*", "give me my *umbrella back*"). In order to eliminate these false positives, we discarded any combinations involving a word contained in WordNet's database of adjectives, verbs and adverbs.

[1] We thank Mirella Lapata for providing this
[2] We thank Arthur Cater for his contribution to data filtering

Remaining noise sponges (e.g. “good *value* meal”, “fifteenth *century* houses”, “second *hand* car”, “low *risk* venture”) were identified by manually searching through the most productive modifiers and heads and checking a sample of combinations for each. As a result of this procedure, a further 35 nouns were removed from the list of legal nouns. The entire filtering process reduced the total number of combinational types from 320,430 to 252,127, a reduction of 21%. Although some legitimate combinations are likely to have been removed by applying these filtering measures, we had no reason to believe that their elimination was nonrandom.

Of the 748 nouns under consideration, 16 abstract and 24 concrete nouns were among those nouns already discarded in the filtering process (e.g. *fun*, *risk*, *tweezers*). We also removed a further 4 nouns because they never appeared in the BNC (*gramercy*, *inclemency*, *tush*, *lye*). Thus, in the end 355 abstract and 349 concrete nouns remained in our analysis. Using our filtered corpus of combinations, we obtained type and token frequency counts for each noun in both the modifier and head positions.

Results

63 abstract nouns and 10 concrete nouns failed to occur in any combinations and were consequently excluded from the analyses comparing modifier and head productivity. As a result, these were based on the remaining 292 abstract nouns and 339 concrete nouns. All t-tests reported below are two-tailed.

The abstract nouns occurred as a modifier in 17,330 combinations and as a head in 25,699 combinations, an average of 59.3 modifier combinations versus 88.0 head combinations per noun. This difference between modifier tokens and head tokens was not statistically significant, $t(291) = 1.6$, $p = .11$. There were 4,359 modifier versus 8,287 head types observed, an average of 14.9 modifier types and 28.4 head types per noun. This difference was statistically significant, $t(291) = 4.16$, $p < .01$

A separate analysis examined the *proportions* for each noun of modifier and head tokens and types. The difference between the proportions of abstract head tokens (.61) and abstract modifier tokens (.39) was statistically significant, $t(291) = 5.85$, $p < .01$. The difference between the proportions of abstract head noun types (.63) and abstract modifier types (.37) was also statistically significant, $t(291) = 7.60$, $p < .01$. Therefore, all significant differences observed for the abstract nouns were in line with our hypothesis that these nouns would occur more frequently as heads than as modifiers.

We carried out the same analyses for the concrete nouns. For this set of nouns, there were 68,747 modifier combinations and 65,126 head combinations, an average of 202.8 modifier combinations versus 192.1 head combinations per noun. This difference between modifier tokens and head tokens was not statistically significant, $t < 1$. There were 16,001 modifier versus 12,439 head types observed, an average of 47.2 modifier types and 36.7 head types per noun. This difference was statistically significant, $t(338) = 2.67$, $p < .01$.

The difference between the proportions of concrete head tokens (.53) and concrete modifier tokens (.47) was not statistically significant, $t(338) = 1.69$, $p = .09$. The difference between the proportions of concrete head types (.52) and concrete modifier types (.48) also was not statistically significant, $t(338) = 1.04$, $p = .30$. Therefore, our hypothesis that concrete nouns would appear more often in the modifier position was only partially supported. The results are summarized in Tables 1 and 2.

TOKENS	Average Number		Average Proportion	
	MOD	HEAD	MOD	HEAD
Abstract	59.3	88.0	.39	.61*
Concrete	202.8	192.1	.47	.53

Table 1. Data analysis by combination tokens (* $p < .01$)

TYPES	Average Number		Average Proportion	
	MOD	HEAD	MOD	HEAD
Abstract	14.9	28.4*	.37	.63*
Concrete	47.2*	36.7	.48	.52

Table 2. Data analysis by combination types (* $p < .01$)

An additional analysis examined the relative frequency with which abstract nouns and concrete nouns are used in combination. This analysis included the 73 nouns that failed to occur in any combinations. The number of occurrences of the 355 abstract nouns in the BNC was 836,034 with 43,029 of these occurrences being in combinations (5.1%). The number of occurrences of the 349 concrete nouns was 1,128,311 with 133,873 of the occurrences appearing in combinations (11.9%). On average, an abstract noun occurred as part of a combination 3.8% of the time, 1.7% as a modifier and 2.1% as a head. This difference between the likelihood of an abstract noun appearing as a modifier versus as a head noun was marginally statistically significant, $t(354) = -1.84$, $p = .07$. Concrete nouns were far more likely to appear as part of a combination. On average they occurred in combinations 11.9% of the time, 6.9% as modifier and 5.1% as a head noun. The difference between the likelihood of a concrete noun being used as a modifier versus as a head noun was statistically significant, $t(348) = 13.12$, $p < .01$. We compared the likelihood of an abstract noun versus a concrete noun being used as part of a combination. The difference was statistically significant for modifiers, $t(702) = 3.00$, $p < .01$, for heads, $t(702) = 3.02$, $p < .01$, and overall, $t(702) = 8.47$, $p < .01$. These results suggest that abstract nouns are far less likely to appear as part of a combination than are concrete nouns, a finding that had not been anticipated. These results are summarized in Table 3.

FREQUENCY	Modifier	Head	Total
Abstract	1.7%	2.1%	3.8%
Concrete	6.9%	5.1%	11.9%

Table 3. Combination frequency as a percentage of total

General Discussion

Our results clearly support the view that different types of noun manifest different behavioral patterns in combination and hence that noun properties affect how a noun will be used in combination. In general, findings of differences in productivity are in line with our hypotheses, even if the predicted difference for concrete nouns was not as well supported as that for abstract nouns. These results are therefore compatible with the slot-filling approach which motivated our predictions but are not accounted for by the CARIN theory.

Differences in modifier and head productivity

We hypothesized that abstract nouns would be better suited to acting as heads because of their ambiguity and that concrete nouns would be less suited to acting as heads because of their specificity. Although these trends were partially supported, the preferences of both types of noun exhibited a considerable amount of variance.

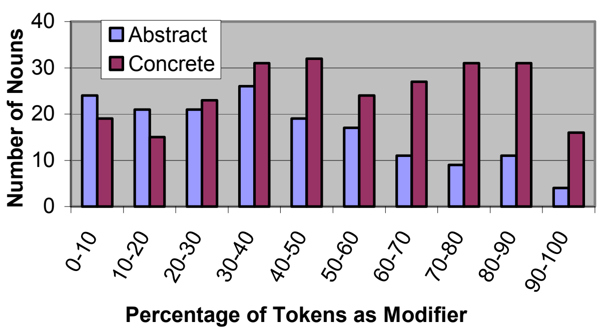

Figure 1. Number of nouns by level of modifier preference

Figure 1 reveals the different levels of modifier preference for the 163 abstract nouns and 249 concrete nouns involved in more than 10 combinations. The different levels correspond to the overall percentage of times that nouns appeared as modifiers as opposed to as heads. Although the majority of abstract nouns (68%) combined more often in the head position, a considerable proportion still exhibited a preference for the modifier position. Illustrating the extent of this variability, abstract terms such as *requirement* (99%) and *behavior* (93%) appeared in combination far more often as head nouns whereas *security* (92%) and *intelligence* (89%) occurred predominantly as modifiers.

Concrete nouns also exhibited preferences at both extremes, but in this case, the number of nouns with a preference for the head position (48%) was very close to the number with a preference for the modifier position (52%). Concrete nouns like *shed* (100%) or *tray* (99%) appeared nearly always as head nouns whereas *mahogany* (100%) and *silver* (97%) appeared nearly always as modifiers.

As our selection of concrete concepts from the MRC2 database included only the most concrete concepts, we expected that these concepts would be highly specific and therefore that they would not benefit from specialization. However, the variability in the preference of these concrete nouns challenges this assumption. Investigating the issue further, we inspected the concrete nouns contained within our sample. It emerged that although these concepts had been sampled at the extremes of concreteness, many were relatively ambiguous and hence their associated set size was larger than expected. For example, the sample included concepts such as *men*, *dweller*, *performer* and *retailer*, none of which are highly specific. As a result, these nouns *can* benefit from specialization and are thus liable to appear as head nouns (e.g. a *dweller* can be a *cave dweller*, *city dweller* etc.) Even a seemingly concrete concept like *bowl* exhibits a level of variability which allows for considerable specialization (e.g. *sugar bowl*, *plastic bowl*, *coffee bowl*). In light of this, the lack of support for the predicted bias in productivity regarding concrete nouns does not necessarily oppose the slot-filling view: the nature of the concrete nouns contained in our study might explain why our hypothesis was not more clearly supported. We speculated that an analysis containing only highly specific concrete nouns might have been more in line with our predictions. For example, *carrot* (19 modifier types, 5 head types), *violin* (35 modifier types, 4 head types), and *steam* (178 modifier types, 11 head types) are among the more specific nouns in our sample and these appear to support the hypothesis that concrete nouns prefer the modifier role.

Frequency of Use in Combination

Our results have shown that the relative frequency with which abstract nouns appear in combination is less than that of the concrete nouns. While slot-filling theories imply that the properties of a concept affect its productivity in either the head or modifier position, they do not make any predictions regarding the relative frequency with which concepts will be used in combination. Therefore, the finding that the concrete nouns in our study were over three times more likely to appear as part of a combination was unexpected.

In order to investigate this issue and develop a further understanding of the link between abstractness and productivity, we computed the ten most productive modifiers and heads appearing in the BNC. Table 4 presents these nouns by token frequency.

MODIFIERS	Freq	HEADS	Freq
Family	983	System	1353
Government	843	Work	858
Business	813	Group	849
School	773	Level	839
Water	752	Problem	783
Police	691	Area	778
Party	674	Programme	675
Metal	651	Type	618
Market	639	Unit	615
Group	595	Process	610

Table 4. BNC's 10 most productive modifiers and heads

This selection reveals that the most productive heads in the BNC are quite abstract. Markman (1999) measures abstractness by the range of different items that can be represented by a concept. Clearly, words like *type* and *unit* have a very large possible range of representation. While it appears that the top modifiers are somewhat less abstract (*water* and *metal* are tangible entities, as can be *family*, *school* and *police*), the nouns contained in this selection are nevertheless relatively abstract. This is very surprising given our finding that our abstract concepts were much less likely to be used in combinations.

We explain this apparent contradiction by considering the differences between the abstract concepts listed in Table 4 and those in our own sample. While the above nouns are abstract because of their range and ambiguity, the nouns in our selection were abstract by virtue of their intangibility. Concepts like *type*, *unit* and *group* frequently boast physical manifestations, whereas this is far less likely for concepts such as *discipline*, *competence* or *insight*.

In light of this, our claim that abstract concepts are less likely to be involved in combination should be refined: abstractness related to range of representation actually increases productivity while abstractness related to intangibility has the opposite effect. As an example, let us compare the productivity of a superordinate category with a member of that same category: the noun *sofa* has 6 modifier types and 17 head types while *furniture* has 110 modifier types and 69 head types. This is because *furniture* refers to a broader range of perceptually diverse entities while the category extension of *sofa* is relatively small. Thus, *furniture* can be associated with a variety of combinational types that are impossible for *sofa* (e.g. *furniture polish*). In contrast, the abstract nouns in our sample did not involve superordinate categories and hence this form of abstractness was unrelated to range of representation or category extension. Instead, they referred to events, traits, emotions and thoughts (e.g. *ingratitude*, *disquiet*, *chaos*), concepts which are unlikely to have physical manifestations or a wide range of representation.

Based on these observations we propose an account of modifier and head productivity which acknowledges the reasons why combinations are used. According to slot-filling theories, the purpose of compound phrases is to specify a particular subcategory of a head noun concept. This function is used in cases where the head noun by itself would not adequately delineate the intended referent. The modifier that is selected is chosen because the required concept is most precisely and diagnostically distinguished from within its category by the relationship it has with that modifier. For example, *picture book* is a particular type of book, but more specifically it is one with pictures. Because a suitable one-word expression which can index the *picture book* concept does not exist, its relationship with the concept *picture* is used to distinguish it from other members of the *book* category.

In light of this, we propose that the principal variable influencing head productivity will be the extent to which subcategories can be identified within it. Thus, very specific concepts like *turnip* will not typically be specialized: it is difficult for the average person to imagine different types of turnip that could be identified as separate subcategories within the general category. Since *turnip* refers to a relatively homogenous group of entities, the addition of a modifier is typically unnecessary. More unusual specializations, such as *dustbin turnip* or *fungus turnip*, are only required in highly specific situations that do not arise very often. On the other hand, concepts with a greater range of representation will have more potential for supporting distinct subcategories and will therefore exhibit higher productivity as a head. Indeed, *vegetable* is six times more productive as a head noun than *turnip*, mainly because subcategories of *vegetable* are easily identified (e.g. *root vegetable*, *leaf vegetable*). As shown in Table 4, the most productive head noun in the BNC is *system*, no doubt due to the many different types of system that can be identified.

We propose that the principal variable influencing modifier productivity is the extent to which subcategories can be delineated by virtue of their relationship with that modifier concept. As a result, concepts that can fill slots in a wide range of head nouns will be the most productive as modifiers. In this way, the noun *mountain* is very productive since it can fill the <LOCATED> slot of a wide variety of entities that live on or are located in the mountains. *Metal* is productive because of the many things that can be made out of metal. Similarly, *summer* is productive because of its ability to fill the <DURING> slot of numerous entities that can occur during the summer. In contrast, a noun like *turnip* has less potential for acting as a modifier because of the limited set of concepts defined by their interaction with turnips. Possibilities exist perhaps in cooking (*turnip stew*), in production (*turnip field*), or describing the parts of a turnip (*turnip head*). As shown in Table 4, the most productive modifier is *family*, due to the fact that many concepts can have subcategories which relate to families (e.g. *family car*, *family holiday*).

Given the above perspective, how can we explain why our abstract nouns appeared in combination so infrequently? For a start, they were sampled at the extreme of abstractness and thus referred to intangible entities with no physical manifestation. At this level of abstractness, the potential for identifying subcategories (head productivity) and for interacting with other concepts (modifier productivity) is severely limited. For example, a concept like *ingratitude* does not support distinct variations in the way that a concrete noun like *shoe* does (e.g. *leather shoe* versus *gift ingratitude*). *Ingratitude* is a vaguely defined mental construct that encompasses a general spectrum of behavior and consequently, distinct manifestations of this concept are poorly recognized. On the other hand, concrete concepts exist in two or three dimensions, and thus can vary in their constituents and location. *Ingratitude* enjoys none of these properties and the lack of a physical manifestation also reduces its potential for interaction with other concepts, a problem that concrete concepts like *shoe* do not have (e.g. *shoe box*, *shoe shop*). Indeed, *ingratitude* cannot be used for anything, cannot be used by anything and other concepts are rarely defined by their relationship with *ingratitude*. Consequently, this noun's level of abstractness severely constrains its productivity as both a modifier and a head: of the 55 times *ingratitude* occurred in the BNC, it never once occurred as part of a combination.

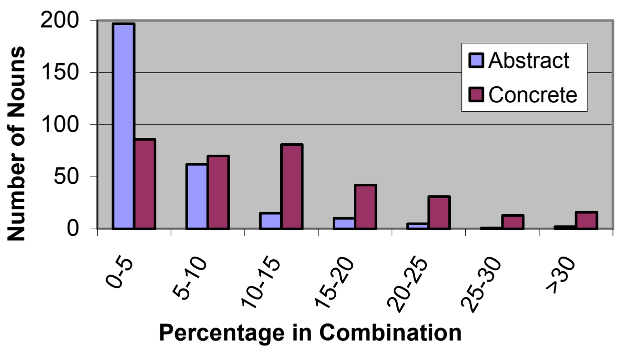

Figure 2. Number of nouns as a function of the percentage of occurrences in the BNC as part of a combination

Despite the fact that the abstract nouns in our study appeared in combination far less often than concrete nouns, there were still a number of exceptions to this rule. Although concepts like *criterion* or *requirement* are intangible, they can still support subcategories: there are many different types of criteria and many different types of requirements. As a result, these nouns exhibited high head productivity. In contrast, concepts like *security* and *mystery* have much potential for specializing other concepts: many things can be used for security purposes and many different situations can constitute a mystery. Accordingly, these nouns were very productive as modifiers. In light of this variability, abstractness alone should not be interpreted as a determinant of combinational productivity, only as a general guide. Figure 2 provides a breakdown of the number of abstract and concrete concepts in our study falling into different levels of combinational productivity.

Conclusion

Previous studies in conceptual combination have focused primarily on the interpretation process. Although slot-filling theories imply that conceptual content might affect how nouns are used in combination, this had not been empirically tested. Addressing this issue, we examined a representative sample of how nouns occur in combination. As suggested by slot-filling theories, our results demonstrated a link between the abstractness of a concept and its frequency of use in the modifier or head position. We also found that overall, abstract concepts were far less likely to be used in combination. This study has provided strong empirical evidence for the view that concept properties dictate their use in combination. As a result, theories of conceptual combination should be extended to account for this relationship. In particular, our results suggest that measures of combinational frequency used in theories such as CARIN might actually be an epiphenomenon of conceptual content. In order to achieve a better understanding of conceptual combination, future study should investigate the factors influencing modifier and head productivity in more detail.

References

Colheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental. Psychology, 33A,* 497-505.

Gagné, C. L. & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, 71-87.

Galbraith, R. C., & Underwood, B. J. (1973). Perceived frequency of concrete and abstract words. *Memory and Cognition*, *1*, 56-60.

Lapata, M. & Lascarides, A. (2003). Detecting Novel Compounds: The role of distributional evidence. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, 235-242.

Markman, A. B. (1999). *Knowledge Representation*. Lawrence Erlbaum Publishers Mahwah, New Jersey.

Murphy, G. L. (1988). Comprehending complex concepts. *Cognitive Science, 12,* 529-562.

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science*, 29, 719-736.

Wisniewski, E. J. (1997). When concepts combine. *Psychonomic Bulletin & Review*, 4, 167-183.