Title
How Understanding Novel Compounds is Facilitated by Priming from Similar, Known Compounds

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
Keane, Mark T.
Tagalakis, Georgios

Publication Date
2005

Peer reviewed

How Understanding Novel Compounds is Facilitated by Priming from Similar, Known Compounds

Georgios Tagalakis (georgios.tagalakis@ucd.ie)
Department of Computer Science, University College Dublin
Dublin 4, Ireland

Mark T. Keane (mark.keane@ucd.ie)
Department of Computer Science, University College Dublin
Dublin 4, Ireland

Abstract

We examine how the understanding of novel noun-noun compounds is facilitated by priming from similar, previously-encountered compounds. Similarity between a set of familiar primes and two sets of target phrases was defined using a computational model. This similarity classification was further validated with a rating study. Two reaction time experiments showed that people find it easier to understand target compounds similar to their primes than target compounds dissimilar to their primes, irrespectively of whether the compounds invite relational or property interpretations. Taken together, these findings have important implications: The use of an 'objective' similarity measure improves methodologically on previous work in the field. The empirical substantiation of the similarity metric provides us with some extra comfort for its use as a control tool for experimental studies and as a component of concept combination models. The picture of the similarity priming becomes more complete given the evidence collected for both major compound types (property and relational). The findings also suggest that those models providing mechanisms for producing property and relational interpretations utilizing previous knowledge are on the right track.

Keywords: concept combination; priming; semantic similarity.

Introduction

Imagine that *entry plan* is a novel compound that you have heard for the first time. To what extent might the meaning you develop for it be based on a related known compound, like *exit strategy* (a compound developed in US politics during the 1950s)? In this paper, we examine some theoretical attempts that have been made to explain this sort of phenomenon and present some new empirical tests of these ideas. This research falls under the general rubric of concept combination research - research that examines the processes by which people form new meanings for combinations of words (see, e.g., Costello & Keane, 2000; Gagné & Shoben, 1997; Wisniewski & Gentner, 1991). Concept combination is important to cognition because it is a test-tube case of the generativity of language; for instance, about 55% of the new terms in English come from combinations of existing words (Cannon, 1987).

In general, it has been shown that most of the meanings people generate to novel noun-noun compounds can be characterized in two broad classes: (i) property interpretations, in which a property of one concept is asserted on the other (e.g, *finger cup*: "a thin cup"), and (ii) relational interpretations, in which a relation is found to connect the two concepts (e.g., *finger cup*: "a cup in which to clean fingers") (cf. Wisniewski & Gentner, 1991). Indeed, it has been found that some compounds consistently yield property interpretations (property compounds), while other consistently yield relational interpretations (relational compounds). In the present work, we look at whether known compounds (e.g., *egg cup*) can prime the understanding of a novel compound (e.g., *finger cup*) using both classes of compounds (i.e., property and relational).

Background

Some of the theories of concept combination predict that known compounds should, in some way, support the understanding of novel compounds, though the mechanisms by which this is achieved are of varying specificity (see, e.g., Costello & Keane, 2000; Gagné & Shoben, 1997). Gagné and Shoben (1997) have argued that people retain distributional knowledge about the relations that have previously been associated with constituents of a compound and build interpretations by selecting the most frequently occurring ones. So, for example, the novel compound *mountain bungalow* is interpreted as "a bungalow located in a mountain region" because other known compounds of the *mountain something* variety (e.g., *mountain lake*, *mountain farm*) use the *located-in* relation. Gagné (2001, 2002) has corroborated this by showing that recent exposure to a similar combination (e.g., *oil moisturizer* or *surgery treatment*) influences the ease of processing of a subsequent combination (e.g., *oil treatment*) by increasing the availability of the lexical entries for the head and the modifier and the relation used to link the two nouns.

In this work, we are theoretically guided by computational models of the combination process that we have built (Lynott, Tagalakis, & Keane, 2004; Tagalakis & Keane, 2004). These models are based on the Constraint theory of concept combination which maintains that a unitary combination process generates possible meanings that have to satisfy the pragmatic constraints of diagnosticity, plausibility and informativeness (Costello & Keane, 2000). In these models, known compounds would contribute predicates to the combination process in much the same way as the individual constituents. However, the more similar the known compound is to the novel compound, the greater the likelihood that the novel compound will adopt its meaning. The reasons for this prediction basically hinge on the fact that the plausibility constraint promotes new meanings for the novel compound that semantically overlap with already-known meanings. The models also have the added benefit that they can explain the interpretation of property and relational compounds.

Problems Description

Our constraint models do not see relational similarity as the main driver of meaning generation from known compounds, but they rather see similarity as the key factor. Unfortunately, Gagné's (2001, 2002) priming evidence does not decide between these two accounts. In the present experiments, we try to separate the two factors by holding relational similarity constant while varying the similarity of the constituents; more specifically, the modifiers of the novel and known compounds. Also, her studies are limited to the examination of similarity effects on relational compounds only. In this study, we examine similarity priming effects for relational and property compounds (see Experiment 1 and 2).

However, there is a problem with similarity. Gagné (2001, 2002) controlled similarity with a 'subjective' rating study carried out by a small number of participants. The main problem when one relies on consensual ratings is that they may well reflect other variables (e.g., sensicality). Therefore, we had to use an 'objective', computational measure of similarity. Unfortunately, computational measures are far from perfect and can be subject to problems (e.g., dependance on heuristic methods that ignore the rich world knowledge humans possess or emerging properties and relations when concepts combine). Hence, we had also to validate the objective measure with a subjective measure (see Rating Study). We use a combination of the two measures to develop materials to test in our two experiments. Experiment 1 uses a sensicality judgement task in conjunction with our priming paradigm. Experiment 2 replicates the first experiment using a comprehension task instead of a sensicality judgements.

Semantic Similarity Modelling

To properly characterize the similarity between novel and known compounds in our subsequent experiments we adapted a model that uses WordNet (Fellbaum, 1998). WordNet is one of the most influential computational lexical resources and very well suited for similarity measures, since it organizes nouns and verbs into hierarchies based on *is-a* hypernymy relations. One way to compute similarity between two concepts is to use the information content of the least common subsumer of the two concepts compared and the information content of the two concepts themselves. Information content is a measure of the specificity of a concept. Seco, Veale, and Hayes (2004) have developed a method for defining information content that does not require any resources other than WordNet. The main guiding assumption for computing information content values is that concepts with many hyponyms are less informative than concepts with fewer hyponyms. The semantic similarity formulation is:

$$sim(c_1, c_2) = 1 - \frac{ic_{wn}(c_1) + ic_{wn}(c_2) - 2 * sim_{res}(c_1, c_2)}{2}$$

where c_1 and c_2 are the two concepts compared, $ic_{wn}(c_1)$ and $ic_{wn}(c_2)$ are the information content values expressed as functions of the hyponyms each of them has, and sim_{res} corresponds to Resnik's (1995) similarity function that accommodates the information content values. The formula is linearly normalized to constrain the output to values between 0 and 1 (for details, see Seco, Veale and Hayes, 2004).

A preliminary evaluation they conducted showed that this metric correlated highly with some existing similarity ratings gathered from people ($N = 30, \rho = .84, p < .001$) and performed better than previous WordNet-based similarity models. We adopted a simple approach to extending this formula to compounds. For each compound it was assumed that its modifier and head noun as such are equally important for determining similarity. The similarity between the two modifiers and the two heads was computed separately, the scores were divided by two (to constrain the output to values between 0 and 1) and were added up. Formally, this can be expressed as:

$$sim'(cc_1, cc_2) = \frac{sim(m_{cc_1}, m_{cc_2})}{2} + \frac{sim(h_{cc_1}, h_{cc_2})}{2}$$

Rating Study: Compounds Similarity

To empirically validate the objective similarity scores derived from the model, a rating study was carried out to determine the degree of correspondence between the model scores and people's similarity ratings for the same items.

Method

Materials Forty-six noun-noun combinations were selected from the WordNet knowledge base to be used as familiar primes. These combinations were judged to be familiar by the authors. Additionally, all the combinations were found to be present in the British National Corpus (BNC; Burnard, 1995). Half of these compounds were relational, half property compounds. Two versions of each of these prime compounds were created: one with a common head and a similar modifier (e.g., *coal tank* to the original *gas tank*), and one with a common head and a dissimilar modifier (e.g., *sand tank* to the original *gas tank*). Operationally, similar modifiers were defined as pairs of concepts with a similarity score of 0.25 or higher. Dissimilar modifiers were defined as pairs of concepts with a similarity score lower than 0.25. Given that the common head creates a base similarity score of 0.50, as a whole the similar compounds had a score of equal to or greater than 0.75, and the dissimilar compounds had a score of lower than 0.75. Thus, we had 3 sets of 46 compounds: 46 primes, 46 similar compounds to these primes, and 46 dissimilar compounds to the primes.

These similar and dissimilar target compounds also met several other constraints: (i) they were judged as sensible combinations by the authors and two independent judges (i.e., compounds for which a ready meaning could be found), (ii) they were verified to be novel by not occurring at all in the BNC, (iii) for relational compounds, it was established that they could be interpreted using the same relation as the prime and that this relation occurred with the same frequency in the different compounds (to control for Gagné and Shoben's, 1997, relational frequency effect); this was carried out by querying the BNC with all the concepts in the compound. A large sample of sequences (50 noun-noun sequences from the BNC written corpus and 50 from the BNC spoken corpus, when available) involving each concept were extracted, from which associated relations were identified and counted, (iv) the frequency of modifiers of the relational and property compounds was not reliably different in the set of similar targets ($M_R = 26.04, M_P = 20.96$; Mann-Whitney $U = 206, p = .20, 2-tailed$) and dissimilar targets ($M_R = 25.91, M_P = 21.09$; Mann-Whitney $U = 209, p = .22, 2-tailed$), and (v) it was verified that there was not a reliable difference between the frequency of modifiers in the two target sets (Wilcoxon's matched pairs signed ranks test, $Z = -.44, p = .66, 2-tailed$).

Subjects Thirty-five English-speaking, undergraduate students from UCD volunteered to participate in the study.

Procedure The rating study was done via a web interface on a PC. Ninety-two pairs of one prime and one target compound were presented on the screen after the instructions and 2 examples. Below them was a similarity rating scale running from 1 (Very Dissimilar) to 7 (Very Similar) (radio boxes). Participants were asked to rate how similar or dissimilar each target compound is to its prime. They were told that they should not base their judgements on how sensible or familiar they find any of the compounds. Compound pairs were presented in a random order, one by one. The rating took about 10 minutes to complete.

Analysis, Results and Discussion

The ratings were treated as raw scores and as categories. For the latter, the mean ratings for a given pair of compounds and the number of people who judged it as similar, dissimilar or neutral were found. If the mean similarity rating was < 4 and most people deemed it to be dissimilar it was classified as dissimilar; if it was rated > 4 by most people it was classified as similar.

Overall, there was a close agreement between the similarity scores derived from WordNet and the ratings given by people. The agreement between the consensual and model classifications was roughly 90%. Participants classified 42 of the 46 model-defined similar compound pairs as similar ($M = 4.99$, $SD = .67$), and 41 of the 46 model-defined dissimilar pairs as dissimilar ($M = 3.37$, $SD = .50$). The remaining compounds were differentially classified.

Pearson's correlation coefficients were calculated between the WordNet-based similarity scores and the proportions of people rating an item as similar/dissimilar. The correlation was strong in the tests; the higher the score, the higher the similarity rankings ($N = 92, r = .66, p < .001$); the lower the score, the higher the dissimilarity rankings ($N = 92, r = -.63, p < .001$). A reliable correlation was also found between the mean ratings and the model scores ($N = 92, r = .63, p < .001$). This empirical substantiation of the similarity metric derived from WordNet provides us with some extra comfort for its use in the experiments that follow.

Experiment 1: Similarity and Sensicality

We have already seen that previous work has shown that known compounds can prime interpretations of novel compounds (Gagné, 2001, 2002). However, this work did not provide a clear basis for similarity or distinguish it from relational similarity. Neither did it address the property compounds, as it was solely directed at relational ones. In the present experiment, we re-examine these priming results holding relational similarity constant and systematically varying similarity. We also widened the net to handle property compounds and relational compounds in a single study. A sensicality judgement task, where people were asked to decide quickly how sensible they find a particular compound, was used to tap people's understanding of novel compounds.

Our expectation was that the similar to the primes compounds would be judged sensible faster than the dissimilar to the primes compounds. On the compound type dimension (property/relational), Gagné (2000) predicts that relational compounds should be judged faster than property compounds because they are considered more 'natural'. Notably, Tagalakis and Keane (2003) have not found consistent evidence for this expectation.

Method

Materials Thirty-eight familiar compounds from the rating study were used as primes, 19 of which were property compounds and 19 relational compounds. The novel targets were 38 matching compounds with a common head and a similar modifier (as defined above), and 38 matching compounds with a common head and a dissimilar modifier (as defined above). Thirty-eight filler compounds were also used. They were noun combinations with no obvious interpretations, with common heads to the prime and target compounds, and modifiers with zero similarity scores to the prime and target modifiers.

The target compounds met several constraints: (i) none of them was compound on which people's classification differed from the model classification, (ii) there was no reliable variance of similarity between the property and relational compounds in the two target sets according to the people's ratings ($F(1, 72) = 1.07$, $p = .31$) and the model's scores ($F(1, 72) = .68$, $p = .41$), (iii) there was a significant variance between the two sets of target items regarding their degree of similarity with the human ($F(1, 72) = 135.02$, $p < .001$) and model ($F(1, 72) = 176.83$, $p < .001$) metrics, (iv) there was no reliable variance in word length, (v) the frequency of modifiers of the relational and property compounds was not reliably different in the set of similar targets ($M_R = 22.00, M_P = 17.00$; Mann-Whitney $U = 133, p = .17, 2-tailed$) and dissimilar targets ($M_R = 21.74, M_P = 17.26$; Mann-Whitney $U = 138, p = .22, 2-tailed$), (vi) only 3 modifiers had also an adjective sense (2 in the set of similar targets and 1 in the set of dissimilar targets), and (vi) there was not a reliable variance between the frequency of modifiers in the two target sets, as a Wilcoxon's matched pairs signed ranks test showed ($Z = -.33, p = .74, 2-tailed$).

Subjects Thirty-four English-speaking, undergraduate students from UCD participated in this experiment for partial course credit.

Procedure There were two parts to the experiment. First, participants were asked to study a simple compound and type in its meaning; the familiar primes were presented during this part. Second, when they had completed interpreting the familiar primes, they were presented with a new set of compounds, the novel ones, and asked to judge whether they were sensible or not. The appropriate instructions were presented prior to both parts. In the interpretation task, they were instructed to read each combination carefully before typing in an interpretation. Participants placed the index finger of their predominant hand on a function key of the keyboard ("/" for the right-handed; "\" for the left-handed). When they were ready, they pressed the function key once. The phrase they were being asked to read was displayed immediately in the middle of the screen. Once they had read and understood the word pair, they pressed the function key, typed in the interpretation, and pressed the function key again to continue with the next compound. In the sensicality judgement task, participants were asked to read the compounds presented on the following screens and indicate whether they was sensible or non-sensible as accurately and quickly as possible. They pressed one of two keys to indicate whether they thought the item to be sense or nonsense. For half of the participants, the J key corresponded to "sense" and F key to "nonsense". For the other half, the meaning of the keys were reversed. The response time from presentation of the compounds to pressing the key in judgement was measured. In between all trials the word "Ready?" appeared first on the screen and proceeded by pressing the space bar once. There was a practice session for each part. Stimulus order was randomized for each participant and with each session lasting about 30 minutes.

Analysis, Results and Discussion

The data of 6 participants with error rates greater than 33% were excluded from the analyses. Trials greater than $3SDs$ from each of the remaining participants' grand mean ($N = 44$) were also excluded.

Response Times A two-factor, repeated measures ANOVA for the variables of similarity (similar or dissimilar) and compound type (property or relational) was carried out on the response times to the targets judged as sensible. Analyses were carried out by-subjects and by-items using the General Linear Model procedure. In the by-subjects analysis (F_1), similarity and compound type were treated as fixed factors, whereas subjects was the random factor. For the by-items analysis (F_2), the items nested within the similarity and compound type conditions was the random factor.

A reliable main effect of similarity was found, with the similar compounds ($M = 1703.67, SD = 757.02$) being judged more quickly to be sensible than the dissimilar compounds ($M = 1881.65, SD = 1002.46$), ($F_1(1, 27) = 19.61, p < .001$; $F_2(1, 77) = 10.08, p = .002$). There was no reliable difference between relational ($M = 1794.18, SD = 802.32$) and property compounds ($M = 1773.48, SD = 965.91$), ($F_1(1, 27) = .34, p = .56$; $F_2(1, 77) = .03, p = .87$), and no reliable interaction between the two factors ($F_1(1, 27) = 1.87, p = .18$; $F_2(1, 77) = .09, p = .77$). The means and standard deviations for the four conditions were as follows: property similar, $M = 1687.93, SD = 809.96$; property dissimilar, $M = 1886.48, SD = 1131.27$; relational similar, $M = 1718.30$, $SD = 704.89$; relational dissimilar, $M = 1877.90, SD = 891.14$. The correlation between response times and model metrics was moderately negative ($N = 76, r = -.32, p = .005$); the higher the similarity between the prime and target, the less time was needed to understand the target.

Taking these analyses together, it is quite clear that modifier similarity has a marked effect on the ease with which people can make sense judgements of novel compounds. There is no evidence to suggest that relational compounds are judged as sensible faster than property compounds.

Proportions of Sense Judgements Typically, in these priming experiments, all response times analyses are based on correct responses only with incorrect responses being dropped (e.g., judging a sensible compound to be non-sensible). Under these conditions, the pattern of proportions accepted can also be informative about participants comprehension of the compounds. Data based on the proportions of times a compound was correctly judged to be sensible was analyzed using a two-way, repeated measures ANOVA for the factors of similarity and compound type. In the by-subjects analysis, the dependent variable was the proportion of combinations that a participant judged to be sensible in each condition. In the by-items analysis, the dependent variable was the proportion of occasions that each combination was judged to be sensible.

This analysis revealed a main effect of similarity, with the similar compounds ($M = 75.47, SD = 12.60$) being judged sensible more often than the dissimilar compounds ($M = 62.97, SD = 14.77$), ($F_1(1, 27) = 39.02, p < .001$; $F_2(1, 72) = 18.93, p < .001$). There was also a main effect of compound type with the relational compounds ($M = 74.53, SD = 13.34$) being judged sensible more often than the property compounds ($M = 63.91, SD = 14.87$), ($F_1(1, 27) = 10.91, p = .003$; $F_2(1, 72) = 13.66, p < .001$). There was also a reliable interaction between compound type and similarity in the by-participants test ($F_1(1, 27) = 9.32, p = .005$;

$F_2(1,72) = 3.24, p = .08$). The means and standard deviations for the four conditions were as follows: property similar, $M = 72.74, SD = 12.09$; property dissimilar, $M = 55.08, SD = 11.98$; relational similar, $M = 78.20, SD = 12.82$; relational dissimilar, $M = 70.86, SD = 13.16$. The correlation between proportions of sense judgements and model metrics was moderately positive ($N = 76, r = .39, p = .001$).

These results clearly demonstrate that similar compounds are more often judged to be sensible than dissimilar ones, which is consistent with the response time results. However, the analyses also give some evidence that relational compounds, particularly dissimilar ones, are more often considered to be sensible than their property compound counterparts. This can be a general matter of difficulty people have with property compounds (Gagné, 2000) or be due to the fact that speeded responses are sensitive to the higher distribution of relational compounds in language (see, e.g., Levi, 1978). These are questions that we try to answer in the next experiment and, at the same time, we examine once more the similarity effects using a comprehension task rather than a sensicality judgement task.

Experiment 2: Similarity and Comprehension

Method

Materials The stimuli were identical with those of Experiment 1.

Subjects Thirty English-speaking, University College Dublin, undergraduate students received partial course credit for their participation in the experiment.

Procedure The experiment had two parts as in Experiment 1, the only difference being that in the second part participants were asked to make a comprehension judgement of the targets. Specifically, they read the presented compound on the screen, pressed a function key when they had comprehended it - a condition added to minimize the possibility that priming effects are purely lexical, but are rather about the interpretation of compounds - and typed in the interpretation that first came to mind before passing on to the next compound. The response time was measured from the presentation of the compound to the pressing of the function key. For those compounds they could not comprehend, they typed in the word "no" and continued to the next compound. The order of stimuli was randomized for each participant and each session took about 50 minutes to complete.

Analysis, Results and Discussion

The data of 4 participants with error rates greater than 33% were excluded from the analyses. Trials greater than $3SDs$ from each of the remaining participants' grand mean ($N = 21$) were also excluded. Two judges analyzed participants' interpretations of the compounds ($N = 1,420$) classifying them as relational, property or "other" (minority interpretation types). Two dissimilar, property compounds were relisted as relational because the majority of people interpreted them us such.

Response Times The analyses were conducted on those trials in which participants comprehended the compound and produced a property interpretation ($N = 566$) or a relational interpretation ($N = 700$). The model specifications for the ANOVA tests were exactly the same as those in Experiment 1.

A reliable main effect of similarity was found again, with the similar compounds ($M = 3481.29, SD = 1613.70$) being comprehended more quickly than the dissimilar compounds ($M = 4287.00, SD = 1804.24$), ($F_1(1,26) = 36.70, p < .001$; $F_2(1,651) = 24.68, p < .001$). There was no reliable difference between comprehension times for the relational ($M = 3888.37, SD = 1748.47$) and property compounds ($M = 3767.89, SD = 1741.30$), ($F_1(1,26) = 1.73, p = .20$; $F_2(1,651) = .77, p = .38$), and no reliable interaction between the two factors ($F_1(1,26) = .89, p = .35$; $F_2(1,651) = .68, p = .41$). The means and standard deviations for the four conditions were as follows: property similar $M = 3382.14, SD = 1660.17$; property dissimilar, $M = 4348.21, SD = 1702.44$; relational similar, $M = 3572.16, SD = 1566.66$; relational dissimilar, $M = 4244.95, SD = 1872.36$. The association between response times and model metrics was negative ($N = 76, r = -.62, p < .001$). The results are consistent with those of Experiment 1.

Proportions of Comprehension Judgements Following the same method used in Experiment 1, analyses of variance on the proportions of compounds comprehended were also conducted. A reliable main effect of similarity was observed, with the similar compounds ($M = 79.05, SD = 14.60$) being comprehended more often than the dissimilar compounds ($M = 64.68, SD = 14.77$), ($F_1(1,25) = 50.37, p < .001$; $F_2(1,72) = 18.14, p < .001$). There was no reliable difference between the proportions of relational compounds ($M = 73.73, SD = 15.04$) comprehended over property ones ($M = 69.67, SD = 17.59$), ($F_1(1,25) = 1.75, p = .20$; $F_2(1,72) = .73, p = .40$), neither was the interaction reliable ($F_1(1,25) = .16, p = .70$; $F_2(1,72) = 1.25, p = .27$). The means and standard deviations for the four conditions were as follows: property similar, $M = 79.57, SD = 13.88$; property dissimilar, $M = 61.34, SD = 16.24$; relational similar, $M = 78.67, SD = 15.42$; relational dissimilar, $M = 68.02, SD = 12.70$. A correlation between the proportions of comprehended compounds and the model metrics was found to be moderately positive ($N = 76, r = .34, p = .002$).

Again, we see the same marked similarity effect found in Experiment 1. However, the partial effects of compound type that were observed in the proportions of sensicality judgments in Experiment 1 disappear when people are asked to comprehend the compound targets in this experiment. This observation is consistent with Tagalakis and Keane's (2003) study and their claim that when people are making sensicality judgments they are probably just developing a rough feel for whether or not an interpretation is likely to be found.

General Discussion

In this study, we have provided evidence for how the understanding of novel compounds is facilitated by priming from similar, known compounds. We replicated the results in two experiments using response time measures. We have demonstrated that facilitating effects by priming occur for both relational and property compounds. Unlike previous approaches that are based on a mixture of distributional and semantic information of the concepts been involved, this research has focused on a well understood and measured notion of similarity; namely, that of semantic, taxonomic similarity. In this process, an existing similarity model was extended and used to evaluate our material. The model itself has been evaluated through a comparison with raters' assessments and experimental data, and was found to perform quite well.

Going on from here, future work could examine the influence exerted by a known compound with a common modifier and similar/dissimilar head to some novel one (e.g., *bullet car* to the original *bullet train*), or examine the effects of a familiar prime on a novel target by varying the degree of similarity between both their constituent concepts (e.g., *rocket car* to the original *bullet train*). Also, given the differences reported between people's sensicality and comprehension judgements of property versus relational compounds, it seems that an interesting follow up experiment to the present work would be to compare understanding of novel compounds that are primed by compounds that are similar/dissimilar to these targets in terms of either relations or properties.

With regards to computational models of concept combination, our findings suggest that those models providing mechanisms for producing relational and property interpretations utilizing previous knowledge are on the right track. For example, Lynott, Tagalakis, and Keane (2004) argued that familiarity effects can be manifested when people encounter a compound that is, while still novel, nonetheless similar to some existing compound. Tagalakis and Keane (2004) presented a case-based reasoning system that 'comprehends' compounds using information from previous cases and interpretation patterns encoded in its knowledge base, as well as semantic structures and concept definitions provided by WordNet. The incorporation of a direct measure of similarity, like the one we have used here, can further improve its flexibility and efficiency. While semantic similarity of cannot provide a complete account of concept combination, it does not always constitute a necessary or sufficient condition, it is clear that it is a crucial factor that has something very useful to offer in the development and extension of cognitive and computational models.

References

Burnard, L. (Ed.) (1995). Users' reference guide for the British National Corpus. Oxford: British National Corpus Consortium, Oxford University Computing Service.

Cannon G. (1987). Historical change and English word-formation. New York: Lang.

Costello, F. J., & Keane, M. T. (2000). Efficient creativity: Constraint-guided conceptual combination. *Cognitive Science*, *24*, 299–349.

Fellbaum, C. (Ed.) (1998). WordNet: An electronic lexical database. Cambridge, MA: MIT Press.

Gagné, C.L. (2000). Relation-based combinations versus property-based combinations: A test of the CARIN theory and dual-process theory of conceptual combination. *Journal of Memory and Language*, *42*, 365–389.

Gagné, C. L. (2001). Relation and lexical priming during the interpretation of noun-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 236–254.

Gagné, C. L. (2002). Lexical and relational influences on the processing of novel compounds. *Brain and Language*, *81*, 723–735.

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.

Levi, J. (1978). The syntax and semantics of complex nominals. New York: Academic Press.

Lynott, D., Tagalakis, G., & Keane, M. T. (2004). Conceptual Combination with PUNC. *Artificial Intelligence Review*, *22*, 247–267.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. *Proceedings of the 14^{th} International Joint Conference on Artificial Intelligence* (pp. 448-453). San Francisco, CA: Morgan Kaufmann.

Seco, N., Veale, T., & Hayes, J. (2004). An intrinsic information content metric for semantic similarity in WordNet. *Proceedings of the 16^{th} European Conference on Artificial Intelligence* (pp. 1089–1090). Amsterdam: IOS Press.

Tagalakis, G., & Keane, M. T. (2003). Modelling the understanding of noun-noun compounds: The role of familiarity. *Proceedings of the European Cognitive Science Conference 2003* (pp. 319–324). Mahwah, NJ: Lawrence Erlbaum Associates.

Tagalakis, G., & Keane, M. T. (2004). Automating the interpretation of novel noun-noun compounds using WordNet. *Proceedings of the European Conference on Case-Based Reasoning 2004 Workshops* (pp. 225–236). Madrid: Complutense University of Madrid.

Wisniewski, E. J., & Gentner, D. (1991). On the combinatorial semantics of noun pairs: Minor and major adjustments to meaning. In G. B. Simpson (Ed.), *Understanding Word and Sentence.* North Holland: Elsevier.