Learning to interpret novel noun-noun compounds: evidence from a category learning experiment

Barry Devereux & Fintan Costello
School of Computer Science and Informatics, University College Dublin,
Belfield, Dublin 4, IRELAND
{barry.devereux, fintan.costello}@ucd.ie

Abstract

The ability to correctly interpret and produce noun-noun compounds such as WIND FARM or CARBON TAX is an important part of the acquisition of language in various domains of discourse. One approach to the interpretation of noun-noun compounds assumes that people make use of distributional information about how the constituent words of compounds tend to combine; another assumes that people make use of information about the two constituent concepts' features to produce interpretations. We present an experiment that examines how people acquire both the distributional information and conceptual information relevant to compound interpretation. A plausible model of the interpretation process is also presented.

1 Introduction

People frequently encounter noun-noun compounds such as MEMORY STICK and AUCTION POLITICS in everyday discourse. Compounds are particularly interesting from a language-acquisition perspective: children as young as two can comprehend and produce noun-noun compounds (Clark & Barron, 1988), and these compounds play an important role in adult acquisition of the new language and terminology associated with particular domains of discourse. Indeed, most new terms entering the English language are combinations of existing words (Cannon, 1987; consider FLASH MOB, DESIGNER BABY, SPEED DATING and CARBON FOOTPRINT).

These noun-noun compounds are also interesting from a computational perspective, in that they pose a significant challenge for current computational accounts of language. This challenge arises from the fact that the semantics of noun-noun compounds are extremely diverse, with compounds utilizing many different relations between their constituent words (consider the examples at the end of the previous paragraph). Despite this diversity, people typically interpret even completely novel compounds extremely quickly, in the order of hundredths of seconds in reaction time studies.

One approach that has been taken in both cognitive psychology and computational linguistics can be termed the relation-based approach (e.g. Gagné & Shoben, 1997; Kim & Baldwin, 2005). In this approach, the interpretation of a compound is represented as the instantiation of a relational link between the modifier and head noun of the compound. Such relations are usually represented as a set of taxonomic categories; for example the meaning of STUDENT LOAN might be specified with a POSSESSOR relation (Kim & Baldwin, 2005) or MILK COW might be specified by a MAKES relation (Gagné & Shoben, 1997). However, researchers are not close to any agreement on a taxonomy of relation categories classifying noun-noun compounds; indeed a wide range of typologies have been proposed (e.g. Levi, 1977; Kim & Baldwin, 2005).

In these relation-based approaches, there is often little focus on how the meaning of the relation interacts with the intrinsic properties of the constituent concepts. Instead, extrinsic information about concepts, such as distributional information about how often different relations are associated with a concept, is used. For example, Gagné & Shoben's CARIN model utilizes the fact that the modifier MOUNTAIN is frequently associated with the LOCATED relation (in compounds such as MOUNTAIN CABIN or MOUNTAIN GOAT); the model does not utilize the fact that the concept MOUNTAIN has in-

Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, pages 89–96,
Prague, Czech Republic, June 2007

trinsic properties such as *is large* and *is a geological feature*: features which may in general precipitate the LOCATION relation.

An approach that is more typical of psychological theories of compound comprehension can be termed the concept-based approach (Wisniewski, 1997; Costello and Keane, 2000). With such theories, the focus is on the intrinsic properties of the constituent concepts, and the interpretation of a compound is usually represented as a modification of the head noun concept. So, for example, the compound ZEBRA FISH may involve a modification of the FISH concept, by asserting a feature of the ZEBRA concept (e.g. *has stripes*) for it; in this way, a ZEBRA FISH can be understood as a fish with stripes. Concept-based theories do not typically use distributional information about how various relations are likely to be used with concepts.

The information assumed relevant to compound interpretation is therefore quite different in relation-based and concept-based theories. However, neither approach typically deals with the issue of how people acquire the information that allows them to interpret compounds. In the case of the relation-based approaches, for example, how do people acquire the knowledge that the modifier MOUNTAIN tends to be used frequently with the LOCATED relation and that this information is important in comprehending compounds with that modifier? In the case of concept-based approaches, how do people acquire the knowledge that features of ZEBRA are likely to influence the interpretation of ZEBRA FISH?

This paper presents an experiment which examines how both distributional information about relations and intrinsic information about concept features influence compound interpretation. We also address the question of how such information is acquired. Rather than use existing, real world concepts, our experiment used laboratory generated concepts that participants were required to learn during the experiment. As well as learning the meaning of these concepts, participants also built up knowledge during the experiment about how these concepts tend to combine with other concepts via relational links. Using laboratory-controlled concepts allows us to measure and control various factors that might be expected to influence compound comprehension; for example, concepts can be designed to vary in their degree of similarity to one another, to be associated with potential relations with a certain degree of frequency, or to have a feature which is associated with a particular relation. It would be extremely difficult to control for such factors, or investigate the aquisition process, using natural, real world concepts.

2 Experiment

Our experiment follows a category learning paradigm popular in the classification literature (Medin & Shaffer, 1978; Nosofsky, 1984). The experiment consists of two phases, a training phase followed by a transfer phase. In the training phase, participants learned to identify several laboratory generated categories by examining instances of these categories that were presented to them. These categories were of two types, conceptual and relational. The conceptual categories consisted of four "plant" categories and four "beetle" categories, which participants learned to distinguish by attending to differences between category instances. The relational categories were three different ways in which a beetle could eat a plant. Each stimulus consisted of a picture of a beetle instance and a picture of a plant instance, with a relation occurring between them. The category learning phase of our experiment therefore has three stages: one for learning to distinguish between the four beetle categories, one for learning to distinguish between the four plant categories, and one for learning to distinguish between the three relation categories.

The training phase was followed by a transfer phase consisting of two parts. In the first part participants were presented with some of the beetle-plant pairs that they had encountered in the training phase together with some similar, though previously unseen, pairs. Participants were asked to rate how likely each of the three relations were for the depicted beetle-plant pair. This part of the transfer phase therefore served as a test of how well participants had learned to identify the appropriate relation (or relations) for pairs of conceptual category exemplars and also tested their ability to generalize their knowledge about the learned categories to previously unseen exemplar pairs. In the second part of the transfer phase, participants were presented with pairs of category names (rather than pairs of category items), presented as noun-noun compounds, and were asked to rate the appropriateness of each relation for each compound.

In the experiment, we aim to investigate three issues that may be important in determining the most appropriate interpretation for a compound. Firstly, the experiment aims to investigate the influence of concept salience (i.e. how important to participants information about the two constituent concepts are, or how relevant to finding a relation that information is) on the interpretation of compounds. For example, if the two concepts referenced in a compound are identical with respect to the complexity of their representation, how well they are associated with various alternative relations (and so on), but are of differing levels of animacy, we might expect the relation associated with the more animate concept to be selected by participants more often than a different relation associated equally strongly with the less animate concept. In our experiment, all three relations involve a beetle eating a plant. Since in each case the beetle is the agent in the EATS(BEETLE,PLANT) scenario, it is possible that the semantics of the beetle concepts might be more relevant to relation selection than the semantics of the plant concepts.

Secondly, the experiment is designed to investigate the effect of the ordering of the two nouns within the compound: given two categories named A and B, our experiment investigates whether the compound "A B" is interpreted in the same way as the compound "B A". In particular, we were interested in whether the relation selected for a compound would tend to be dependent on the concept in the head position or the concept in the modifier position. Also of interest was whether the location of the more animate concept in the compound would have an effect on interpretation. For example, since the combined concept is an instance of the head concept, we might hypothesize that compounds for which the head concept is more animate than the modifier concept may be easier to interpret correctly.

Finally, were interested in the effect of concept similarity: would compounds consisting of similar constituent categories tend to be interpreted in similar ways?

learn	*trans.*	Nr	Rel	Bcat	Pcat	B1	B2	B3	P1	P2	P3
l		1	1	1	3	4	1	1	3	2	3
l		2	1	1	3	4	4	1	2	3	3
l	t	3	1	1	3	1	1	1	3	3	2
l	t	4	1	1	3	4	1	2	3	3	3
l	t	5	2	2	2	2	2	2	2	2	3
l		6	2	2	2	2	2	1	2	3	2
l		7	2	2	2	2	3	2	2	2	1
l	t	8	2	2	2	2	2	3	2	2	2
l	t	9	3	3	1	3	3	3	4	1	2
l	t	10	3	3	1	3	3	2	1	1	1
l		11	3	3	1	2	3	3	4	4	1
l		12	3	3	1	3	2	3	4	1	1
l	t	13	1	4	4	1	1	4	4	4	4
l	t	14	2	4	4	4	1	4	4	1	4
l	t	15	3	4	4	4	4	4	1	1	4
	t	16	-	1	1	4	1	1	4	1	1
	t	17	-	3	3	3	3	3	3	3	3
	t	18	-	2	4	2	2	2	4	1	4
	t	19	-	4	2	4	1	4	2	2	2

Table 1: The experiment's abstract category structure

2.1 Method

2.1.1 Participants

The participants were 42 university students.

2.1.2 Materials

The abstract category structure used in the experiment is presented in Table 1. There are 19 items in total; the first and second columns in the table indicate if the item in question was one of the 15 items used in the learning phase of the experiment (l) or as one of the 13 items used in the transfer stage of the experiment (t). There were four beetle categories (Bcat), four plant categories (Pcat) and three relation categories used in the experiment. Both the beetle and plant categories were represented by features instantiated on three dimensions (B1, B2 & B3 and P1, P2 & P3, respectively). The beetle and plant categories were identical with respect to their abstract structure (so, for example, the four exemplars of Pcat1 have the same abstract features as the four exemplars of Bcat1).

Beetles and plants were associated with particular relations; Bcat1, Bcat2 and Bcat3 were associated with Relations 1, 2 and 3, respectively, whereas Pcat1, Pcat2 and Pcat3 were associated with Relations 3, 2 and 1, respectively. Bcat4 and Pcat4 were not associated with any relations; the three exemplar instances of these categories in the learning phase appeared once with each of the three relations. The features of beetles and plants were sometimes diagnostic of a category (much as the feature *has three wheels* is diagnostic for TRICYCLE); for example, a particular feature associated with Bcat1 is a 1 on the B3 dimension: 3 of the 4 Bcat1 training phase exemplars have a 1 on dimension B3 while only one of the remaining 11 training phase exemplars do. Also, the intrinsic features of beetles and plants are sometimes diagnostic of a relation category (much as the intrinsic feature *has a flat surface raised off the ground* is diagnostic for the relational scenario *sit on*); values on dimensions B1, P1, B2 and P2 are quite diagnostic of relations. Participants learned to identify the plant, beetle and relation categories used in the experiment by attending to the associations between beetle, plant and relation categories and feature diagnosticity for those categories.

The beetle and plant categories were also designed to differ in terms of their similarity. For example, categories Bcat1 and Bcat4 are more similar to each other than Bcat3 and Bcat4 are: the features for Bcat1 and Bcat4 overlap to a greater extent than the features for Bcat3 and Bcat4 do. The aim of varying categories with respect to their similarity was to investigate whether similar categories would yield similar patterns of relation likelihood ratings. In particular, Bcat4 (and Pcat4) occurs equally often with the three relations; therefore if category similarity has no effect we would expect people to select each of the relations equally often for this category. However, if similarity influences participants' relation selection, then we would expect that Relation 1 would be selected more often than Relations 2 or 3.

The abstract category structure was mapped to concrete features in a way that was unique for each participant. Each beetle dimension was mapped randomly to the concrete dimensions of beetle shell color, shell pattern and facial expression. Each plant dimension was randomly mapped to the concrete dimensions of leaf color, leaf shape, and stem color. The three relations were randomly mapped to *eats from leaf*, *eats from top*, and *eats from trunk*.

2.1.3 Procedure

The experiment consisted of a training phase and a transfer phase. The training phase itself consisted

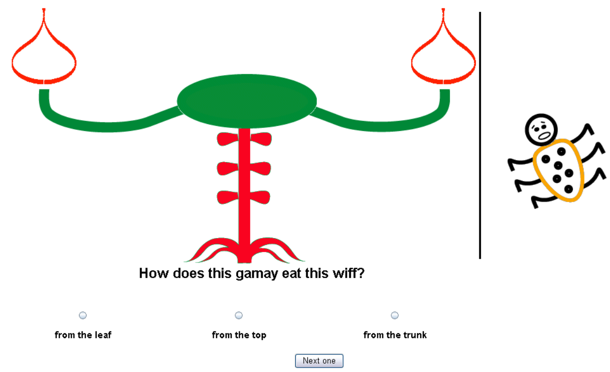

Figure 1: Example of a relation learning stimulus

of three sub-stages in which participants learned to distinguish between the plant, beetle and relation categories. During each training sub-stage, the 15 training items were presented to participants sequentially on a web-page in a random order. Underneath each item, participants were presented with a question of the form "What kind of plant is seen in this picture?", "What type of beetle is seen in this picture?" and "How does this $\langle Bcat \rangle$ eat this $\langle Pcat \rangle$?" in the plant learning, beetle learning, and relation learning training sub-stages, respectively (e.g. Figure 1). Underneath the question were radio buttons on which participants could select what they believed to be the correct category; after participants had made their selection, they were given feedback about whether their guess had been correct (with the correct eating relation shown taking place). Each of the three substages was repeated until participants had correctly classified 75% or more of the items. Once they had successfully completed the training phase they moved on to the transfer phase.

The transfer phase consisted of two stages, an exemplar transfer stage and a compound transfer stage. In the exemplar transfer stage, participants were presented with 13 beetle-plant items, some of which had appeared in training and some of which were new items (see Table 1). Underneath each picture was a question of the form "How does this $\langle Bcat \rangle$ eat this $\langle Pcat \rangle$?" and three 5-point scales for the three relations, ranging from 0 (unlikely) to 4 (likely).

The materials used in the compound transfer stage of the experiment were the 16 possible noun-noun compounds consisting of a beetle and plant category label. Participants were presented with a sentence of the form "There are a lot of $\langle Pcat \rangle$ $\langle Bcat \rangle$s around at the moment." and were asked "What kind of eating activity would you expect a $\langle Pcat \rangle$ $\langle Bcat \rangle$ to have?". Underneath, participants rated the likelihood of each of the three relations on 5-point scales. One half of participants were presented with the compounds in the form "$\langle Bcat \rangle$ $\langle Pcat \rangle$" whereas the other half of participants saw the compounds in the form "$\langle Pcat \rangle$ $\langle Bcat \rangle$".

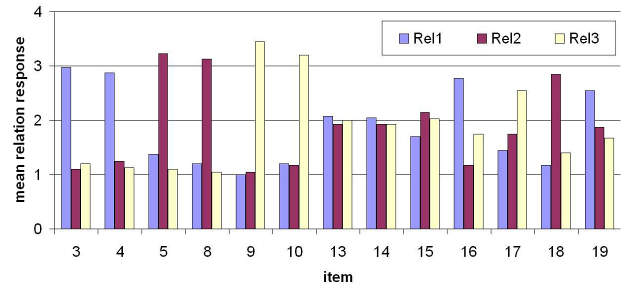

Figure 2: Participants' mean responses for the exemplar transfer items.

2.2 Results

2.2.1 Performance during training

Two of the participants failed to complete the training phase. For the remaining 40 participants, successful learning took on average 5.8 iterations of the training items for the plant categories, 3.9 iterations for the beetle categories, and 2.1 iterations for the relation categories. The participants therefore learned to distinguish between the categories quite quickly, which is consistent with the fact that the categories were designed to be quite easy to learn.

2.2.2 Performance during the exemplar transfer stage

Participants' mean ratings of relation likelihood for the nine previously seen exemplar items is presented in Figure 2 (items 3 to 15). For each of these items there was a correct relation, namely the one that the item was associated with during training. The difference between the mean response for the correct relation ($M = 2.76$) and the mean response for the two incorrect relations ($M = 1.42$) was significant ($t_s(39) = 7.50$, $p < .01$; $t_i(8) = 4.07$, $p < .01$). These results suggest that participants were able to learn which relations tended to co-occur with the items in the training phase.

Participants' mean ratings of relation likelihood for the four exemplar items not previously seen in training are also presented in Figure 2 (items 16 to 19). Each of these four items consisted of a prototypical example of each of the four beetle categories and each of the four plant categories (with each beetle and plant category appearing once; see Table 1 for details). For these four items there was no correct answer; indeed, the relation consistent with the beetle exemplar was always different to the relation suggested by the plant exemplar. For each trial, then, one relation is consistent with the beetle exemplar (r_b), one is consistent with the plant exemplar (r_p) and one is neutral (r_n). One-way repeated measures ANOVAs with response type (r_b, r_p or r_n) as a fixed factor and either subject or item as a random factor were used to investigate the data. There was a significant effect of response type in both the by-subjects and by-items analysis ($F_s(2, 39) = 19.10$, $p < .01$; $F_i(2, 3) = 24.14, p < .01$). Pairwise differences between the three response types were investigated using planned comparisons in both the by-subject and by-items analyses (with paired t-tests used in both cases). The difference between participants' mean response for the relation associated with the beetle exemplar, r_b ($M = 2.68$), and their mean response for the neutral relation, r_n ($M = 1.44$) was significant ($t_s(39) = 5.63$, $p < .001$; $t_i(3) = 5.34$, $p = .01$). These results suggest that participants were strongly influenced by the beetle exemplar when making their category judgments. However, the difference between participants' mean response for the relation associated with the plant exemplar, r_p ($M = 1.62$), and their mean response for the neutral relation was not significant ($t_s(39) = 1.11$, $p = .27$; $t_i(3) = 0.97$, $p = .40$). These results suggest that participants were not influenced by the plant exemplar when judging relation likelihood. Since the beetle and plant categories have identical abstract structure, these results suggest that other factors (such as the animacy of a concept or the role it plays in the relation) are important to interpretation.

The data from all 13 items were also analysed taken together. To investigate possible effects of category similarity, a repeated measures ANOVA with beetle category and response relation taken as within subject factors and subject taken as a random factor was undertaken. There was a significant effect of the category that the beetle exemplar belonged to on participants' responses for the three relations (the interaction between beetle category and response relation was significant; $F(6, 39) = 26.83$, $p < .01$. Planned pairwise comparisons (paired t-tests) were conducted to investigate how ratings for the correct relation (i.e. the relation consistent with training) differed for the ratings for the other two relations. For Bcat1, Bcat2 and Bcat3, the ratings for the relation consistent with learning was higher than the two alternative relations ($p < .01$ in all cases). However, for the Bcat4 items, there was no evidence that participants we more likely to rate Relation 1 ($M = 2.09$) higher than either Relation 2 ($M = 1.97$; $t(39) = 0.54$, $p = .59$) or Relation 3 ($M = 1.91$; $t(39) = 0.69$, $p > .50$). Though the difference is in the direction predicted by Bcat4's similarity to Bcat1, there is no evidence that participants made use of Bcat4's similarity to Bcat1 when rating relation likelihood for Bcat4.

In summary, the results suggest that participants were capable of learning the training items. Participants appeared to be influenced by the beetle exemplar but not the plant exemplar. There was some evidence that conceptual similarity played a role in participants' judgments of relation likelihood for Bcat4 exemplars (e.g. the responses for item 19) but over all Bcat4 exemplars this effect was not significant.

2.2.3 Performance on the noun-noun compound transfer stage

In the noun-noun compound transfer stage, each participant rated relation likelihood for each of the 16 possible noun-noun compounds that could be formed from combinations of the beetle and plant category names. Category name order was a between subject factor: half of the participants saw the compounds with beetle in the modifier position and plant in the head position whilst the other half of participants saw the reverse. First of all, we were interested in whether or not the training on exemplar items would transfer to noun-noun compounds. Another question of interest is whether or not participants' responses would be affected by the order in which the categories were presented. For example, perhaps it is the concept in the modifier position that is most influential in determining the likelihood of different relations for a compound. Alternatively perhaps it is the concept in the head position that is most influential.

To answer such questions a $4 \times 4 \times 3 \times 2$ repeated measures ANOVA with beetle category, plant category and response relation as within subject factors and category label ordering as a between subject factor was used to analyze the data. The interaction between beetle category and response relation was significant ($F(6, 38) = 59.79$, $p < .001$). Therefore, the beetle category present in the compound tended to influence participants' relation selections. The interaction between plant category and response relation was weaker, but still significant ($F(6, 38) = 5.35$, $p < 0.01$). Therefore, the plant category present in the compound tended to influence participants' relation selections. These results answer the first question above; training on exemplar items was transferred to the noun-noun compounds. However, there were no other significant interactions found. In particular, the interaction between category ordering, beetle category and response relation was not significant ($F(6, 38) = 1.82$, $p = .09$). In other words, there is no evidence that the influence of beetle category on participants' relation selections when the beetle was in the modifier position differed from the influence of beetle category on participants' relation selections when the beetle was in the head-noun position. Similarly, the interaction between noun ordering, plant category and response relation was not significant ($F(6, 38) = 0.68$, $p = .67$); there is no evidence that the influence of the plant category on relation selection differed depending on the location of the plant category in the compound.

Planned pairwise comparisons (paired t-tests) were used to investigate the significant interactions further: for Bcat1, Bcat2 and Bcat3, the ratings for the relation consistent with learning was significantly higher than the two alternative relations ($p < .001$ in all cases). However, for Bcat4, there were no significant differences between the ratings for the three relations ($p > .31$ for each of the three comparisons). For the plants, however, the only significant differences were between the response for Relation 1 and Relation 2 for Pcat2 ($t(39) = 2.12$,

$p = .04$[1]) and between Relation 2 and Relation 3 for Pcat2 ($t(39) = 3.08$, $p = .004$), although the differences for Pcat1 and Pcat3 are also in the expected direction.

In summary, the results of the noun-noun compound stage of the experiment show that participants' learning of the relations and their associations with beetle and plant categories during training transferred to a task involving noun-noun compound interpretation. This is important as it demonstrates how the interpretation of compounds can be derived from information about how concept exemplars tend to co-occur together.

2.3 Modelling relation selection

One possible hypothesis about how people decide on likely relations for a compound is that the mention of the two lexemes in the compound activates stored memory traces (i.e. exemplars) of the concepts denoted by those lexemes. Exemplars differ in how typical they are for particular conceptual categories and we would expect the likelihood of an exemplar's activation to be in proportion to its typicality for the categories named in the compound. As concept instances usually do not happen in isolation but rather in the context of other concepts, this naturally results in extensional relational information about activated exemplars also becoming activated. This activated relational information is then available to form a basis for determining the likely relation or relations for the compound. A strength of this hypothesis is that it incorporates both intensional information about concepts' features (in the form of concept typicality) and also extrinsic, distributional information about how concepts tend to combine (in the form of relational information associated with activated exemplars). In this section, we present a model instantiating this hybrid approach.

The hypothesis proposed above assumes that extensional information about relations is associated with exemplars in memory. In the context of our experiment, the extensional, relational information about beetle and plant exemplars participants held in memory is revealed in how they rated relational likelihood during the exemplar transfer stage of the experiment. For each of the 13 beetle and plant exemplars, we therefore assume that the average ratings for each of the relations describes our participants' knowledge about how exemplars combine with other exemplars. Also, we can regard the three relation likelihood ratings as being a 3-dimensional vector. Given that category ordering did not appear to have an effect on participants' responses in the compound transfer phase of the experiment, we can calculate the relation vector $\vec{r}_{B,P}$ for the novel compounds "B P" or "P B" as

$$\vec{r}_{B,P} = \frac{\sum_{e \in U} (typ(e_b, B) + typ(e_p, P))^{\alpha} \cdot \vec{r}_e}{\sum_{e \in U} (typ(e_b, B) + typ(e_p, P))^{\alpha}}$$

where e denotes one of the 13 beetle-plant exemplar items rated in the exemplar transfer stage, $typ(e_b, B)$ denotes the typicality of the beetle exemplar present in item e in beetle category B and $typ(e_p, P)$ denotes the typicality of the plant exemplar present in item e in plant category P. U is the set of 13 beetle-plant exemplar pairs and α is a magnification parameter to be estimated empirically which describes the relative importance of exemplar typicality.

In this model, we require a measure of how typical of a conceptual category an exemplar is (i.e. a measure of how good a member of a category a particular category instance is). In our model, we use the Generalized Context Model (GCM) to derive measures of exemplar typicality. The GCM is a successful model of category learning that implements an an exemplar-based account of how people make judgments of category membership in a category learning task. The GCM computes the probability Pr of an exemplar e belonging in a category C as a function of pairwise exemplar similarity according to:

$$Pr(e, C) = \frac{\sum_{i \in C} sim(e, i)}{\sum_{i \in U} sim(e, i)}$$

where U denotes the set of all exemplars in memory and $sim(e, i)$ is a measure of similarity between exemplars e and i. Similarity between exemplars is in turn defined as a negative-exponential transforma-

[1]This is not significant if Bonferroni correction is used to control the familywise Type I error rate amongst the multiple comparisons

tion of distance:

$$sim(i,j) = e^{-cdist(i,j)} \quad (1)$$

where c is a free parameter, corresponding to how quickly similarity between the exemplars diminishes as a function of their distance. The distance between two exemplars is usually computed as the city-block metric summed over the dimensions of the exemplars, with each term weighted by empirically estimated weighting parameters constrained to sum to one. According to the GCM, the probability that a given exemplar belongs to a given category increases as the average similarity between the exemplar and the exemplars of the category increases; in other words, as it becomes a more typical member of the category. In our model, we use the probability scores produced by the GCM as a means for computing concept typicality (although other methods for measuring typicality could have been used).

We compared the relation vector outputted by the model for the 16 possible compounds to the relation vectors derived from participants' ratings in the compound transfer phase of the experiment. The agreement between the model and the data was high across the three relations (for Relation 1, $r = 0.84$, $p < 0.01$; for Relation 2, $r = 0.90$, $p < 0.01$; for Relation 3, $r = 0.87$, $p < 0.01$), using only one free parameter, α, to fit the data[2].

3 Conclusions

The empirical findings we have described in this paper have several important implications. Firstly, the findings have implications for relation-based theories. In particular, the finding that only beetle exemplars tended to influence relation selection suggest that factors other than relation frequency are relevant to the interpretation process (since the beetle and plants in our experiment were identical in their degree of association with relations). Complex interactions between concepts and relations (e.g. agency in the EATS(AGENT,OBJECT) relation) is information that is not possible to capture using a taxonomic approach to relation meaning.

Secondly, the fact that participants could learn to identify the relations between exemplars and also transfer that knowledge to a task involving compounds has implications for concept-based theories of compound comprehension. No concept-based theory of conceptual combination has ever adopted an exemplar approach to concept meaning; models based on concept-focused theories tend to represent concepts as frames or lists of predicates. Our approach suggests an exemplar representation is a viable alternative. Also, distributional knowledge about relations forms a natural component of an exemplar representation of concepts, as different concept instances will occur with instances of other concepts with varying degrees of frequency. Given the success of our model, assuming an exemplar representation of concept semantics would seen to offer a natural way of incorporating both information about concept features and information about relation distribution into a single theory.

[2] In the GCM, c was set equal to 1 and the three dimensional weights in the distance calculation were set equal to $1/3$

References

G. Cannon. 1987. *Historical change and English word formation*. New York: Lang.

E. V. Clark and B.J. Barron. 1988. A thrower-button or a button-thrower? Children's judgments of grammatical and ungrammatical compound nouns. *Linguistics*, 26:3–19.

F. J. Costello & M.T. Keane. 2000. Efficient creativity: Constraint guided conceptual combination.. *Cognitive Science*, 24(2):299–349.

C. L. Gagné and E.J. Shoben. 1997. Influence of thematic relations on the comprehension of modifier noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23:71–78.

S. N. Kim and T. Baldwin. 2005. Automatic Interpretation of Noun Compounds Using WordNet Similarity. *Lecture Notes in Computer Science*, 3651:945–956.

J. N. Levi. 1978. *The Syntax and Semantics of Complex Nominals.* New York: Academic Press.

D. L. Medin & M.M. Schaffer. 1978. Context theory of classification learning. *Psychological Review*, 85:207–238.

R. N. Nosofsky. 1984. Choice, similarity, and the context theory of classification.. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1):104–114.

E. J. Wisniewski 1997. When concepts combine. *Psychonomic Bulletin & Review*, 4(2):167–183.