In search of the frog's tail: Investigating the time course of conceptual knowledge activation

Conference Paper · May 2007

3 authors, including:

Rebecca Maguire
National University of Ireland, Maynooth

Arthur Cater
University College Dublin

In Search of the Frog's Tail: Investigating the Time Course of Conceptual Knowledge Activation

Phil Maguire (phil.maguire@ucd.ie)
Department of Computer Science, NUI Maynooth
Co.Kildare, Ireland

Rebecca Maguire (rebecca.maguire@dbs.edu)
Department of Psychology, Dublin Business School
Dublin 2, Ireland.

Arthur. W.S. Cater (arthur.cater@ucd.ie)
School of Computer Science and Informatics,
University College Dublin, Dublin 4, Ireland.

Abstract

Slot-filling theories of conceptual combination assume that both constituent concepts are activated before they are combined. However, these theories have difficulty in explaining why combined phrase features are sometimes more available than the features of the constituent nouns. In this study, we investigate the time course of conceptual knowledge activation. Using three verification tasks of varying complexity we demonstrate that basic taxonomic knowledge is retrieved more quickly than modal specific conceptual features. Applying this finding to conceptual combination, we demonstrate that participants take longer to reject combinations requiring the activation of instance specific features (e.g. *frog tail*) than those that can be rejected based on more generalized taxonomic knowledge (e.g. *daffodil tail*). These findings provide convergent evidence that conceptual knowledge is activated dynamically and selectively rather than all at once. We discuss the implications for existing theories.

Keywords: Conceptual combination; noun-noun compounds; knowledge representation; knowledge activation.

Introduction

The combination of two words is a technique commonly adopted by speakers in order to refer to novel concepts and ideas (e.g. *holiday tension*, *picnic bee*). Although people have a well developed means of understanding these novel compounds, the associated comprehension process is not trivial, requiring many levels of understanding. Accordingly, the study of conceptual combination is important, both because it is intimately associated with the generativity and comprehension of natural language and because it is important for understanding how people represent concepts. In English, a language in which compounding is particularly productive, 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, *kitchen chair* is interpreted as a particular type of chair, and more precisely as the type that is located in kitchens.

Thus far, theories of conceptual combination have generally assumed that the comprehension of a compound phrase is dependent on both concepts being fully activated. For example, the Concept Specialization model (Murphy, 1988) assumes a schema structure for concepts, consisting of a series of slots. This theory proposes that during the combination process the modifying concept fills one or more of the slots in the head noun concept. First, the appropriate slot is selected based on world knowledge about the constituent concepts and subsequently this combined concept is elaborated (e.g. realizing that a *car magazine* is likely to have a picture of a car on the front cover).

According to the slot-filling view, an identical set of features is activated whenever a particular concept is used in combination, regardless of the noun it is paired with. Clearly though, people cannot retrieve all associated knowledge about a concept every time it is encountered. Much of that information would be irrelevant and would impair rather than aid comprehension. A more economical approach would be for conceptual information to be activated selectively, thereby avoiding the need for additional processes to suppress irrelevant information. However, current theories of conceptual combination offer no clue as to how a selective activation process might operate.

The inadequacy of slot-filling theories is highlighted by their inability to explain key observations relating to knowledge availability. Springer and Murphy (1992) compared the time taken to verify a property that was true of the head versus a property that was true of the phrase. For instance, the feature *green* applies to both *celery* and *boiled celery* (noun feature). In contrast, the feature *soft* is only valid for *boiled celery* (phrase feature). Based on the idea that concepts must be fully activated before being combined, Springer and Murphy expected that the noun property would be verified more quickly than the phrase property. However, the opposite findings emerged, with participants being quicker to verify the phrase property (i.e.

that boiled celery is soft). According to Springer and Murphy, these findings are paradoxical because they suggest that emergent features of the combined concept are activated before the features of the constituent concepts.

One possible explanation for this result is that people become aware of a compound phrase structure before activating the constituent nouns and are therefore in a position to activate only the conceptual knowledge that is relevant to the combination. The idea that word meanings emerge gradually rather than all at once is well supported. For example, Till, Mross & Kintsch (1988) identified clear stages in word comprehension, with sense selection occurring around 400ms and further semantic inferences following around 1,000ms. Eye-tracking measures show that eye fixations last on average 200ms during the reading of linguistic text (Rayner, 1988), suggesting that people will be able to retrieve preferentially those features that are relevant to the combination. In this case, the instantiation of the concept *boiled celery* should proceed in much the same way as if it was referenced by a single label, in that the properties of ordinary celery that are not pertinent to boiled celery should not be activated.

The enabling condition for selective activation is that knowledge retrieval is a gradual incremental process rather than an all at once phenomenon. The existence of a distinction between different levels of conceptual detail is well supported by neurological evidence. For example, Warrington (1975) described a patient with a dementing illness who had lost subordinate attribute information (e.g. knowing that a cabbage was green) yet retained superordinate classification information (e.g. knowing that a cabbage was a plant). Also, several distinct event-related brain potentials have been identified that occur at different time intervals during concept activation (Kumar & Debruille, 2004). These have been linked to various different stages of the knowledge activation process, specifically phonological matching, activation of syntactic word category information, semantic processing, evaluation and finally representation construction.

In light of this, we propose that knowledge activation is a dynamic process and that this phenomenon can successfully explain how phrase features for conceptual combinations can be more available than noun features. In this paper we present two experiments which investigate this possibility. In the first we compare response times for three verification tasks of differing complexity. In the second we apply these findings to conceptual combination and investigate the time taken to reject phrases requiring the activation of different levels of conceptual detail.

Experiment 1

The aim of this experiment was to present participants with a series of words and to analyze the time taken to verify different conceptual features. We wished to ascertain whether the more general taxonomic knowledge about a concept becomes available prior to the retrieval of detailed features. For example, do people realize that a *dog* is a thing before they realize that it is an animal? Do they realize it is an animal before they know what it looks like?

We required a set of verification tasks that would test the availability of different conceptual features. Three tasks were selected, one requiring word-level information (does it name a thing?), a second requiring the activation of basic conceptual knowledge (is this thing alive?) and a third requiring the activation of a specific perceptual feature (is this organism hairy?). In the latter task, the hairiness attribute was selected because this information is not accurately reflected by the conceptual hierarchy (e.g. although many mammals have hair, hippos and rhinos do not). Importantly, for all three tasks, the concepts did not need to be situated within a context in order to verify the relevant property. Given our hypothesis that the activation of conceptual knowledge proceeds from the basic to the more detailed, we predicted the following trend in response times: Object < Animate < Feature.

Method

Participants Twenty-seven first year undergraduate students from University College Dublin participated in the experiment for partial course credit.

Design The experiment used a within-participants design, with three conditions corresponding to the three verification tasks, namely Object, Animate and Feature. In order to facilitate a within-participants design it was necessary to use a separate list of words for each condition. Had the same list been used for all three tasks, then the equal partitioning of true and false responses would not have been possible. Each participant saw the same set of 180 stimuli, comprising the three conditions of 60 items each.

This design improves on that of previous verification tasks (e.g. McElree & Murphy, 2006) involving the introduction of an additional concept (e.g. *boiled celery* is soft). In our experiment, the words under consideration are presented on their own. Participants are already aware of the feature to be verified so they are not required to activate information about other concepts in order to respond. In addition, participants apply the same verification task to a broad variety of words, therefore providing a more reliable measure of feature availability.

Materials We compiled separate lists of 60 different nouns for each of the three conditions. In each we included 30 items which were representative of the feature being verified and 30 items which were not. In the Object condition, half of the items were nouns (e.g. *vase*, *couch*) while the other half were connectives and other parts of speech (e.g. *because*, *when*). In the Animate condition, half of the items were organisms (e.g. *mouse*, *tulip*) while half were artifacts (e.g. *shed*, *pebble*). In the Feature condition, half of the items were haired creatures (e.g. *leopard*, *panda*) while the other half were hairless (e.g. *whale*, *rhinoceros*). All sets of words were controlled for length and familiarity. Analysis revealed no significant differences in average word length between the three conditions (5.2, 5.1 and 5.5 for the Object, Animate and Feature conditions respectively, $F(2, 177) = 1.5$, $p = .22$). There were also no significant differences between average word lengths for the representative and non-representative items in the Object condition (5.3 and 5.0 respectively, $t(58) = .83$, $p = .41$), the Animate condition (5.1 and 5.2, $t(58) = -.22$, $p = .83$) and the Feature condition (5.4 and 5.6, $t(58) = -.57$, $p = .84$). The familiarity of the nouns in the various conditions was compared by taking the log of their BNC frequency. This revealed no significant differences in frequency for the representative items (3.0, 2.8 and 2.9 for the Object, Animate and Feature conditions respectively, $F(2, 87) = .69$, $p = .50$). Furthermore, there was also no significant difference in the log of the frequency for the representative and non-representative items in the Animate (2.8 and 3.0, $t(58) = -1.01$, $p = .32$) and the Feature conditions (2.9 and 2.8, $t(58) = .87$, $p = .39$). We did not include the frequencies of the non-nouns in our analyses, as a comparison of this nature would have been misleading. The relationship between familiarity and frequency of use is not consistent when comparing nouns with other parts of speech. All of the nouns included in our experiment were associated with tangible manifestations (e.g. artifacts and plants), meaning that even those with a relatively low frequency were recognizable (e.g. *walrus* occurs only 64 times in the BNC). In contrast, non-nouns can occur more frequently yet be unfamiliar due to their abstractness (e.g. *trenchant* has a BNC frequency of 74). In order that all of the words in our Object condition be comparatively familiar, the frequency of our non-nouns was necessarily higher (4.5).

Procedure Participants sat in front of a computer screen and placed the index finger of their left hand on the F key of the computer keyboard and the index finger of their right hand on the J key. They were informed that a series of words would be displayed on the screen and that the objective was to decide whether the words were representative of the feature in question, pressing J for 'yes' and F for 'no'. For the Object condition, the task presented to participants was to verify whether the word in question referred to a thing or not. For the Animate condition, the task was to decide if the item in question was alive or not. For the Feature condition, the task was to decide if the animal in question was covered in hair. Prior to the start of the experiment, participants were provided with several worked examples in order to demonstrate the nature of the verification task. During the experiment, words appeared in the middle of the screen and participants had to make a decision by pressing the appropriate key. Trials were separated by a blank screen lasting for one second.

Each condition began with 10 practice trials which did not form part of the experiment. The purpose of these trials was to allow participants to adjust to the task, although they were not aware that the trials in question would not be included. Subsequently, the 60 experimental stimuli followed seamlessly. In each condition, the words were presented in a random order to each participant. Furthermore, the three conditions were randomized so that participants performed the tasks in a different order.

Results and Discussion

A total of 6.5% of the data were omitted from the analysis. 3.7% of responses were incorrect and hence these trials were not considered. Additionally, response times deemed unreasonably fast (< 400ms, 0.1%) or unreasonably slow (> 4000, 0.6%) were also excluded. After this initial elimination process, any remaining response times which were more than three standard deviations outside each participant's mean for that condition were also excluded. This eliminated a further 2.0% of trials.

The mean response times were 738, 802 and 884 ms for the Object, Animate and Feature conditions respectively. Further analysis revealed that the mean response times for the representative items in the three conditions were 727, 768 and 873 ms while the mean response times for the non-representative items were 749, 837 and 897 ms respectively. These data are illustrated below.

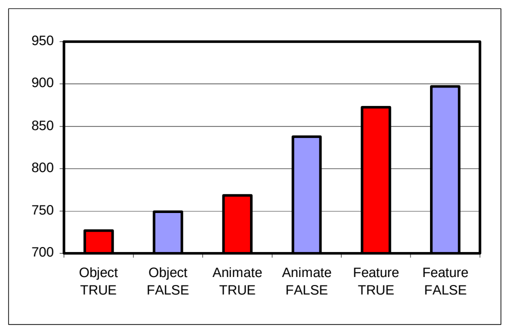

Figure 1. Mean positive and negative response times (ms)

We conducted a series of ANOVAs in order to examine the differences between the various conditions. For the by-participants analysis we computed a two-way repeated measures ANOVA, with three levels corresponding to the different tasks and two levels corresponding to the appropriate response type, all within-participants. For the by-items analysis we computed a non-repeated measures ANOVA with two fixed factors. There was no significant interaction between task and response, F_1 (2, 52) = 2.48, $p = .09$, MSE = 3652.76; F_2 (2, 174) = 1.28, $p = .28$, MSE = 6816.40. However, there was a significant main effect of task, F_1 (2, 52) = 21.89, $p < .001$, MSE = 16321.42; F_2 (2, 174) = 54.32, $p < .001$, MSE = 370306.12. A Page's L trend analysis revealed a significant increasing trend in response times according to Object < Animate < Feature, $L(2) = 364$, $p < 0.01$. In other words, participants were able to verify that a word was a thing before they were able to verify that it was alive or that it had hair. This pattern of results supports our view of knowledge activation as constituting a dynamic, incremental process.

There was also a significant main effect of response, F_1 (1, 26) = 23.20, p < .001, MSE = 3987.85; F_2 (1, 174) = 13.78, p < .001, MSE = 93960.26. Thus, across all three conditions, participants were quicker at verifying word features than they were at discounting them. This pattern of results suggests that property verification involves some kind of active search process which terminates as soon as confirming information is identified but which otherwise continues until a certain threshold of certainty is reached. This challenges the notion that people store information about concepts in a propositional format, as assumed by the schema-structures used in slot-filling models. For example, if the knowledge as to whether an animal is hairy or not is explicitly stored with that concept then there should be no difference between the time needed to confirm or discount the feature. The fact that we observed a difference suggests that the verification process involves more than simply accessing propositional knowledge.

The idea of a gradual knowledge activation process may explain some features of conceptual combination which could not be accounted for by schema-based theories, such as the fact that phrase features can be verified more quickly than noun features. If the most basic knowledge about concepts is activated first, then this provides a means by which more detailed information can be activated selectively (i.e. only the combined concept itself need be simulated). In the following experiment we investigate whether dynamic knowledge activation can contribute to the understanding of how concepts are combined.

Experiment 2

The aim of this experiment was to investigate whether dynamic knowledge activation can explain some of the counterintuitive effects observed involving conceptual combination (e.g. Springer and Murphy, 1992). Specifically, we wished to ascertain whether the combination process begins before the constituent concepts have been fully activated. In order to do this, we created two conditions of implausible combinations, one where explicit featural knowledge was required in order to reject the combination (e.g. *frog tail*), and another where more basic taxonomic knowledge was sufficient (e.g. *daffodil tail*). Our hypothesis was that participants would reject combinations from the Basic condition more quickly than those from the Detailed condition, based on the differences in time taken to activate the requisite knowledge.

Existing theories of conceptual combination have difficulties in explaining how combinations can be rejected as implausible. The Competition Among Relations in Nominals (CARIN) theory (Gagné & Shoben, 1997) proposes that combinations are interpreted by applying one of a small set of possible relations to the constituent nouns. The theory therefore implies that combinations can only be rejected when every single possible relation has been applied and none result in a satisfactory interpretation. People cannot know if a relation will be successful or not until they apply it, meaning that they have to option but to apply them all. According to the CARIN theory then, there should be no difference in the time taken to reject combinations from either condition. Slot-filling theories suggest that the modifier fills a slot in the head noun and that this process is guided by general knowledge about the two concepts. These theories assume that both concepts are completely activated prior to their combination (e.g. Murphy, 1988). If people have retrieved the concepts *frog* and *tail* before attempting to combine them, then they will be aware that *tail* cannot fill the <has as body part> slot in *frog* since frogs do not have tails. On the other hand, the link between *daffodil* with *tail* is less obvious, suggesting that a more extensive search for plausible relationships will be required before this combination can be ruled out. Thus, the slot-filling view predicts that *frog tail* should be easier to dismiss as the appropriate slot and filler are clear yet obviously incompatible.

According to our dynamic activation view, the combination process will begin *before* both nouns have been fully activated. *Daffodil tail* can be rejected as soon as the basic semantic categories of the constituent concepts becomes available since the pairing <plant-body part> does not match a productive pattern. On the other hand, *frog tail* matches a very productive pattern and can only be rejected when the precise knowledge that frogs do not have tails becomes available. We propose that this detailed knowledge about frogs will only be activated when people attempt to visualize the combined concept and fail. Accordingly, we propose that combinations in the Detailed condition will take longer to reject than those in the Basic condition.

Method

Participants Twenty-six first year undergraduate students from University College Dublin participated in the experiment for partial course credit.

Design A within-participants design was used for the experimental manipulation of condition. Each participant saw the same set of 80 stimuli, comprising the two conditions of 20 items each and the 40 sensible fillers.

Materials Twenty combinations were generated for each of the conditions. For the Detailed condition, this set constituted a series of combinations that were exemplars of a productive pattern of combination (e.g. <animal-body part>). However, all happened to be implausible by virtue of some instance-specific detail of one of the constituents. For example, *frog tail* is an implausible combination since frogs do not have tails, yet many other reptiles and animals do. *Raspberry peel* is implausible since raspberries cannot be peeled, yet many other fruits can. Also, *train tyres* are implausible because trains do not have tyres, yet many other vehicles have tyres. In the Basic condition we created a matching set of combinations which substituted the concept for which detailed knowledge was required. This substituted concept was too far removed in the conceptual hierarchy to yield a sensible combination. For example, both daffodil and frog are organisms. However, *daffodil* is a plant and the lowest common abstraction of entities that tend to have tails is *animal*. Our hypothesis was that *daffodil tail* would be rejected before *frog tail* because the knowledge that daffodils are not animals would be activated before the knowledge that frogs do not have tails. In other examples, *raspberry peel* was paired with *doughnut peel* and *train tyres* was paired with *vase tyres*.

The combinations were controlled for length and also for familiarity. The average number of letters in the Detailed and Basic conditions was not significantly different (both 10.75). The average number of syllables between these conditions was not significantly different (both 3.25). The log of the average BNC frequency of the words used in the Detailed and Basic conditions was not significantly different (6.1 and 6.3 respectively, $t(19) = -.99$, $p = .33$). Finally, the log of the Google frequency of the combinations used in both conditions was not significantly different (2.7 and 2.3 respectively, $t(19) = 1.81$, $p = .09$), which was to be expected given that none of the combinations were intended to be sensible.

As well as the 40 implausible stimuli we also included 40 sensible filler items in order to balance the sample (e.g. *tomato sandwich*). We avoided including overtly lexicalized items, in order that participants would be required to actively combine the constituent concepts.

Procedure Participants sat in front of a computer screen and placed the index finger of their left hand on the F key of the computer keyboard and the index finger of their right hand on the J key. They were informed that a series of noun-noun compounds would be displayed on the screen for which they would have to make plausibility judgments, pressing J for plausible and F for implausible. Each trial was separated by a blank screen lasting for one second. The combination then appeared in the middle of the screen and participants had to make a decision by pressing the appropriate key.

Participants were initially given a short practice session where feedback was given regarding their judgments. The aim of this practice session was to familiarize them with the process of making quick plausibility judgments and also to set a reliable threshold for plausibility. After completing the practice session, participants were instructed that they were now beginning the experiment. The stimuli were then presented in a random order to each participant.

Results and Discussion

A total of 14.5% of the data were omitted from the analysis. We eliminated any positive responses to the implausible stimuli (12.5%). Additionally, response times deemed unreasonably fast (< 400ms, 0.1%) or unreasonably slow (> 4000, 1.6%) were also excluded. After this initial elimination process, any remaining response times which were more than three standard deviations outside each participant's mean for that condition were also excluded. This eliminated a further 0.3% of trials.

The average response time for the Detailed condition was 1,503ms while that for the Basic condition was 1,333ms. Repeated measures ANOVAs revealed that this difference was significant both by-items and by-participants, $F_1(1,25) = 27.89$, $p < .001$, MSE = 12360.01; $F_2(1,19) = 9.53$, $p < .01$, MSE = 48560.09. The difference in accuracy for the Detailed and Basic conditions (93% and 78% respectively) was also significant both by-items and by participants, $F_1(1,25) = 17.47$, $p < .001$, MSE = 6.36; $F_2(1,19) = 15.80$, $p < .001$, MSE = 9.14. Only two of the stimuli were incorrectly judged by the majority of participants (both Detailed), namely *liquid ice* (13 correct responses) and *evening sunrise* (9 correct responses).

These results demonstrate that participants were quicker and more accurate in dismissing the Basic combinations than the Detailed combinations. This finding provides converging evidence that knowledge activation is not an all at once phenomenon, therefore providing a means by which conceptual information might be activated selectively in combination. Importantly, the difference in response times between both conditions indicates that the combination process begins *before* all knowledge relevant to the constituent nouns has been activated. Had the participants in our experiment activated both concepts first, then the items in the Detailed condition would certainly have been rejected first: a full representation of *frog* and *tail* would have permitted the speedy realization that the concepts were incompatible.

This experiment has demonstrated that implausible combinations can be quickly and reliably rejected without the need for a long search for potential interpretations. Yet, how could participants be confident that a combination was not sensible before trying every single possibility? Clearly, they must have been relying on some kind of heuristic in order to guide the combination process, or else the more ambiguous items in the Basic condition would have taken longer to reject. Given the finding of Experiment 1 that basic taxonomic knowledge is the first to be activated, we propose that people rely on this information in order to constrain the interpretation process, and that more detailed information is applied selectively, thereby 'homing in' on the precise meaning of the combination. For example, knowing that *frog* is an animal and that *tail* is a body part is enough to strongly suggest the <has> relation, thereby greatly reducing the range of possible interpretations which must be considered. Similarly, the knowledge that daffodil is a plant is sufficient for dismissing *daffodil tail* since the pattern <plant-body part> is highly irregular. In sum, we propose that people are sensitive to how different types of concept tend to interact in combination and that they use this heuristic in order to activate conceptual detail selectively. This guided selective activation process might explain how people can interpret potentially ambiguous combinations so quickly and so reliably, an issue which previous theories have failed to explain satisfactorily.

General Discussion

We have provided evidence that conceptual knowledge is activated dynamically, with detailed features being less available than more basic taxonomic knowledge. One reason for this effect might be that the latter is represented amodally while the verification of specific features requires the manipulation of a representation. Most interpretations of property verification tasks have assumed that participants make use of amodal representations, accessing data structures such as semantic networks, feature lists or schemas in order to find the required information. In contrast, our results have suggested that conceptual knowledge is not stored in this way.

Much evidence has been garnered supporting the idea that a significant part of conceptual knowledge is modality-specific. Barsalou (2005) maintains that during property verification, people scan mental simulations in order to evaluate whether test properties can be perceived. This view is supported by numerous studies showing that variables such as occlusion, size, shape and orientation affect conceptual processing. For example, Solomon and Barsalou (2004) analyzed response times for a property verification task and found that as features became larger, they took longer to verify, suggesting that people must attend to particular regions of a simulation in order to perceive a feature. Supporting this stance, many of our participants reported using visual imagery in order to discriminate between hairy and non-hairy animals, particularly in cases where this information could not be deduced from the conceptual hierarchy. The idea that people instantiate concepts visually in order to verify visual features would explain the differences in response time between our conditions. For example, it may be the case that participants had to instantiate a visual representation of a rhinoceros before being able to tell whether it was hairy or not, thus lengthening the response time for this task. The greater the level of representational manipulation required, the longer the time taken to verify the feature.

Barsalou (2005) proposed that rather than being regarded as a general description of a category, a concept can be more accurately described as the productive ability to generate many different situated conceptualizations. For example, traditional models involving conceptual knowledge view the concept *dog* as a detached collection of amodal facts that becomes active as a whole every time the category is processed. However, this idea cannot provide for the specialized inferences needed in particular situations (e.g. a growling guard-dog as opposed to a playful pup). An understanding of the concept entails the ability to produce a wide variety of situated conceptualizations that support goal achievement in specific contexts. In light of this, the concept for *dog* cannot simply be a detached global description of a fixed set of propositional features.

This view is compatible with the findings of our experiments. Basic taxonomic knowledge is likely to be stored propositionally in order to facilitate conceptualization (e.g. knowing whether a word refers to a thing, or whether that thing is an artefact or an organism). However, for reasons of economy, the number of features stored in this way is likely to be relatively small. More obscure properties are likely to be verified by scrutinizing a modal-specific simulation. However, if basic taxonomic knowledge is sufficient for indicating how the constituent nouns of a combination are related, then this can be used to inform the activation process so that only the combined concept need be simulated.

Conclusion

We have shown that conceptual knowledge is activated dynamically, with generalized taxonomic information being more available than specific features. Our findings are compatible with previous findings in neuroscience (e.g. Kumar & Debruille, 2004; Warrington, 1975) and psycholinguistics (e.g. Solomon & Barsalou, 2004) and suggest a distinction in how people represent different types of information. We have shown that the phenomenon of dynamic knowledge activation may be crucial to the understanding of how concepts are combined. The idea of a selective activation process can successfully explain how people avoid accessing conceptual information that is not relevant to a combination. Future work should investigate in more detail how conceptual knowledge is represented in memory and further analysis of the processes involved in conceptual combination may prove revealing in this regard.

References

Barsalou, L.W. (2005). Situated conceptualization. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (pp. 619-650). St. Louis: Elsevier.

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.

Kumar, N. & Debruille, J.B. (2004). Semantics and N400: Insights for schizophrenia. *Journal of Psychiatry and Neuroscience*, *29*, 89-98.

McElree, B. & Murphy, G.L. (2006). Time course of retrieving conceptual information. A speed-accuracy trade off study. *Psychonomic Bulletin and Review*, *13*, 848-853.

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

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422.

Solomon, K.O. & Barsalou, L.W. (2004). Perceptual simulation in property verification. *Memory and Cognition*, *32*, 244-259.

Springer, K., & Murphy, G.L. (1992). Feature availability in conceptual combination. *Psych. Science*, *3*, 111-117.

Till, R.E., Mross, E.F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory and Cognition*, *16*, 283-298.

Warrington, E.K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635-657.