In: Kokinov, B., Karmiloff-Smith, A., Nersessian, N. J. (eds.) ***European Perspectives on Cognitive Science***.
 ISBN 978-954-535-660-5

Are Word Meanings Atomic or Complex? An Investigation of Conceptual Knowledge Activation in Context

Phil Maguire (pmaguire@cs.nuim.ie)
Rebecca Maguire (rebecca.maguire@dbs.ie)
Department of Computer Science, NUI Maynooth
Co. Kildare, Ireland

Abstract

While some studies have suggested that conceptual knowledge can be activated selectively (e.g. Potter & Faulconer, 1979; Springer & Murphy, 1992), others have suggested that the same default set of conceptual features is activated, regardless of the context in which a word is used (e.g. McElree, Murphy & Ochoa, 2006; Swinney, Love, Walenski & Smith, 2007). This study investigated whether context-inappropriate features are brought to mind when people interpret modified concepts (e.g. *nursery bear*; *paper boat*; *birdcage door*). In the contextual condition participants evaluated the weight of a modified concept, while in the compositional condition they first evaluated the weight of the unmodified concept before evaluating that of the modified concept. The results revealed that items involving a mismatch between modified and unmodified weight were more difficult to interpret in the compositional condition but not in the contextual condition. These findings indicate that words are interpreted differently in context than in isolation; specifically, they imply that the activation of context-inappropriate features is avoided when words are interpreted in context.

Keywords: Compositionality, contextuality, conceptual activation, concepts, language comprehension, word meaning

Introduction

An important question concerning the process of language comprehension is how words reference meaning. One theory, known as compositionality, is that words index default schemas which are triggered each time a word is encountered (e.g. McElree, Murphy & Ochoa, 2006; Swinney, Love, Walenski & Smith, 2007). According to this view, people initially activate context-insensitive representations for concept words. This pooled knowledge is then sorted out by suppressing inappropriate features and retrieving additional emergent properties. Compositionality assumes that default conceptual schemas form the building block of language: meanings are parceled into atomic units and, in order to understand a phrase or sentence, people must first activate all of the units individually before putting them together.

The alternative theory, known as contextuality, is that semantic memory is addressable not only by words but also by word combinations and sentences (e.g. Barsalou, 1982, 1987; Potter & Faulconer, 1979; Springer & Murphy, 1992). Rather than being linked to fixed, static schemas, people's knowledge of word meanings includes information about how those words interact in context to index mental representations. Accordingly, inappropriate features need not be brought to mind when words are encountered in context (e.g. *fluffy* for *stone squirrel*).

Support for Contextuality and Compositionality

Several studies have supported the idea of contextuality. Potter and Faulconer (1979) found that participants responded more quickly to a picture of a modified concept than an unmodified concept following the presentation of a word combination. They aurally presented participants with sentences such as "the man saw the *burning house* ahead of him". A picture was presented immediately after the critical concept and participants had to verify whether it referred to something in the sentence. In one condition, the picture depicted the modified concept (e.g. a burning house) while in the other it depicted the unmodified concept (e.g. an ordinary house). When presented with an unmodified noun, participants responded quicker to the unmodified picture. However when presented with a modified noun, they responded quicker to the modified picture: the addition of a modifier like *burning* seemed to have an immediate effect on how the concept *house* was interpreted. Based on these observations, Potter and Faulconer (1979) posited that the activation of a contextually-modified concept proceeds as if that concept was directly referenced, in that the properties of the constituent concepts that are not relevant to the combined concept are not activated: when people hear a phrase like *burning house* they are able to home in on the appropriate mental representation without needing to first activate the default schema for *house*.

Another study by Springer and Murphy (1992) revealed a similar effect. They compared the time taken to verify a property that was true of an unmodified noun versus a property that was true of a combined phrase. For instance, the feature *green* is one that applies equally to both *celery* and *boiled celery*. In contrast, the feature *soft* is an emergent phrase feature because it is only valid for *boiled celery*. They found that participants were quicker to verify the phrase feature than the noun feature, supporting the view that context-appropriate meanings are retrieved directly, without recourse to default schemas.

McElree et al. (2006) highlighted some methodological issues in Potter and Faulconer's (1979) and Springer and Murphy's (1992) studies. They argued that, because participants were given as long as they wanted to respond, these experiments do not tell us about the early stages of processing. They also pointed out that Potter and Faulconer

(1970) did not include a condition in which the picture matched both the modified and unmodified concept. For instance, while a picture of an ordinary house matches the concept *house*, it mismatches the phrase *burning house*. Thus, it is not possible to draw firm conclusions regarding the differences in response times: participants may have taken longer to identify the picture of the ordinary house for *burning house* because of the mismatch between them, rather than because of selective conceptual activation. Springer and Murphy's (2002) experiment also failed to include a condition in which the feature to be verified was true of the noun yet false of the phrase (e.g. toy cars are fast). As a result, the experiment did not shed light on the question of whether context-inappropriate features are initially brought to mind, this being the principal issue separating the compositional and contextual views.

McElree et al. (2006) conducted an experiment which addressed these limitations. They re-used Springer and Murphy's (1992) original paradigm of verifying properties for compound phrases while also including two new conditions, one in which the verification property was false of the noun and false of the phrase (e.g. water pistols have string) and another in which the property was true of the noun and false of the phrase (e.g. water pistols are dangerous). In addition, they employed a speed-accuracy trade-off task in which participants were required to respond within six predefined time windows, allowing them to trace the accumulation of knowledge as a function of time.

McElree et al. (2006) found that by two seconds after the presentation of the verification property, participants were extremely accurate in verifying true features, whether they were based on the noun alone or required integration with the modifier. However, at earlier stages in processing, participants judged noun properties more accurately than those requiring integration of the modifier and noun, contradicting the findings of earlier studies. Based on these observations McElree concluded that context-irrelevant properties are initially brought to mind and that "sorting-out of relevant information occurs at a later, integrative stage of sentence or discourse interpretation" (p. 853). For example, when people encounter the concept *water pistol*, they initially activate properties such as 'is dangerous' and 'has bullets'; these irrelevant properties are later suppressed once the pooled schematic knowledge for *water* and *pistol* has been integrated.

Another on-line study by Swinney et al. (2007) provided further support for compositionality. They examined the time course of integration of modified nouns using cross-modal lexical priming. In this experiment, sentences involving modifier-noun phrases were presented aurally and participants simultaneously made a lexical decision about strings displayed on a computer screen. Priming strings were presented at two points during the sentence (e.g. "the executive delivered a speech proposing the *peeled*[1] *banana*[2] as the company's new logo") and were either associated with the default concept (e.g. *yellow*) or associated with the combined phrase (e.g. *white*). Swinney et al. found that the default noun property elicited priming at the earlier test point, whereas the combined phrase property only elicited priming at the later test point, suggesting that the default properties of a noun are activated before its emergent features, as maintained by the compositional view. Swinney et al. also argued that the context-irrelevant properties which are initially brought to mind are subsequently suppressed, as evidenced by a decay in priming for the noun feature at the later test point.

Methodological Issues

Although the balance of evidence might seem to favor the compositional view, certain methodological issues in both McElree et al.'s (2006) and Swinney et al.'s (2007) studies may undermine the validity of their conclusions.

One significant issue concerning McElree et al.'s (2006) study is that participants are likely to have relied on lexical heuristics to guide their responses, rather than conceptual knowledge. Solomon and Barsalou (2004) argued that participants can use two strategies in property verification tasks, namely a shallow linguistic strategy and a deeper conceptual simulation strategy. In their study they demonstrated that when task conditions allow, participants spontaneously adopt a superficial linguistic strategy; only when deeper conceptual processing is required is simulation used. In McElree et al.'s (2006) experiment, participants are likely to have exploited the fact that the materials were not controlled for lexical association, especially when under the pressure of making a speeded response. For example, the words *pistol* and *dangerous* co-occur with high frequency ('pistol * dangerous' yields 111 million Google hits), while *pistol* and *harmless* less so (13.6 million Google hits).

When participants rely on shallow processing heuristics, the verification of properties which require conceptual knowledge retrieval is disadvantaged. Emergent phrase properties which are false of the noun (e.g. *harmless* for *water pistol*) will have lower levels of lexical association than those which are true of the noun (e.g. *has a trigger*). Similarly, features which are true of the noun but false of the phrase (e.g. *dangerous*) will have higher levels of lexical association than those which are false of the noun (e.g. *has string*). Therefore, it is not surprising that, when participants were under greatest time pressure, these properties were verified less accurately. Rather than providing evidence of compositionality, as McElree et al. concluded, these differences in accuracy may simply reflect differences in depth of processing.

Another problematic issue associated with McElree et al.'s (2006) experimental paradigm concerns the process of feature verification. In their study, a different property was verified in each trial. This design is undesirable because it requires an additional concept to be activated before any meaningful response can be given. For example, in order to verify whether "water pistols have string", one must first activate the concept for *string* and then relate that concept to *water pistol*. A wide range of verification properties were

employed: some were abstract adjectives (e.g. dangerous) while others were concrete nouns (e.g. string).

McElree et al.'s use of this paradigm is based on the assumption that all features should be equally straightforward to retrieve. The compositional view suggests that conceptual schemas consist of a set of default propositional features. For example, the schema for *basketball* might have an associated list of features such as *bounces*, *is round*, *weighs 600g*, *floats* et cetera. In this case, it takes just as long to verify that a basketball bounces as it does to verify that it floats, since both properties are stored as labels in the same look-up list. Thus, any differences in ease of verification can be attributed to factors such as conceptual knowledge activation rather than the difficulty of the verification task itself.

However, the contextual view takes a different stance. According to proponents of this theory, the property of, for example, floating is not some propositional label that comes attached with a concept. Instead, to verify whether an object floats or not one must construct, simulate and scrutinize a modality-specific representation in a variety of contexts. Barsalou (1982) asked participants to read two different scenarios (e.g. the basketball was used when the boat sank / the basketball was well worn from much use) and then to verify a property such as "basketballs float". He found that properties were verified more quickly following scenarios to which they were pertinent, arguing against the idea that people store lists of features in propositional form. Because McElree et al.'s (2006) stimuli were not controlled for ease of verification, it is not possible to tell how much of the effect they observed was due to conceptual knowledge activation and how much was due to differences in ease of verification.

A further problem with McElree et al.'s materials is that combinations were not controlled for degree of lexicalization. For example, the fact that *water pistol* has a defined WordNet synset and a Wikipedia page suggests that it is well known concept. Because people are likely to have extensive experience interpreting lexicalized combinations (e.g. *traffic light*, *boiled egg*, *mobile phone*, *garden centre*) the activation of such concepts is less likely to involve the integration of conceptual knowledge.

The methodology of Swinney et al.'s (2007) study is also problematic. One of the key desiderata of scientific investigation is that the process of measurement should be designed so as to limit its effect on that which is being measured. For example, if one observes a fridge light by opening the door, one will erroneously conclude that the light is always on; the error arises because the process of observation affects the state of the light. We suggest that Swinney et al.'s experimental procedure may have directly influenced the results obtained. Their main finding was that contextually inappropriate features were primed at an early stage in processing. For example, participants were quicker to verify *yellow* than *white* while hearing *peeled banana*. However, in this case, the observation of the word *yellow* is part of the context in which the word *banana* is interpreted. When people hear *banana* and see *yellow*, the activation of conceptual knowledge proceeds differently than when *banana* is presented without *yellow*. Studies on backwards affective priming have shown that words presented immediately following a target influence how that word is accessed, revealing that stimulus evaluation is a continuous, dynamic process that does not end with the presentation of the target (Fockenberg, Koole & Semin, 2006). In light of this, we suggest that Swinney et al.'s (2007) priming paradigm may not be suited to examining conceptual knowledge activation: primes presented with short onset asynchronies may contaminate the ongoing activation of the target.

In light of these observations, we propose the following recommendations for experiments investigating the time course of conceptual knowledge activation. First, property verification tasks should require deep processing. Preventing participants from relying on lexical association is crucial, since emergent features will tend to have lower lexical association with concept words than context-independent ones. Solomon and Barsalou (2004) have shown that by carefully constructing a task so that shallow heuristics do not allow above chance performance, participants are motivated to use the conceptual system. Ideally, verification properties should be inexpressible using a single word, thereby avoiding all issues of lexical association (e.g. *does the concept fit in this box that you see in front of you?*). Second, the number of properties to be verified should be minimized and carefully controlled as their introduction presents an additional confounding factor. Third, studies should be based only on novel combinations, as lexicalized phrases can be interpreted directly without requiring the integration of conceptual knowledge. Finally, studies should strive to investigate conceptual knowledge activation under natural conditions: they should not impose artificial response time limits which could alter participants' strategies, or introduce priming concepts which could contaminate the process of activation. In the following section we describe an experiment designed to fulfil these desiderata.

Experiment

In this experiment we investigated whether context-inappropriate features are initially retrieved for modified concepts. The experiment involved two conditions, one which mimicked the assumptions of the compositionality view and the other involving a natural process of combination interpretation. For all trials, the feature to be verified was whether the weight of the concept was above or below 1kg. We reasoned that weight information is likely to form part of the default schema for most concrete concepts, given that it plays a crucial role in determining the interactional affordances of an object. The fact that this feature cannot be expressed using a single word avoids the issues of lexical association associated with previous experiments. The consistency of the task also avoids any issues of erratic verification difficulty.

In the compositional condition, participants first evaluated the weight of the unmodified concept (e.g. *battery*), followed by the weight of the modified concept (e.g. *vehicle battery*). This presentation was intended to mimic a compositional process: when people first encounter a concept like *battery* they initially activate all of its default features, including the property of being lighter than 1kg. In the contextual condition, participants were presented directly with the combined phrase, and were thus not exposed to the unmodified concept.

The experiment investigated whether there was a difference in ease of activation between the two conditions. For one half of the stimuli, the weight of the modified concept (i.e. greater or less than 1kg) was the same as that of the unmodified concept (e.g. *cellar door*), whereas for the other half, the weight of the modified concept was the opposite to that of the unmodified concept (e.g. *birdcage door*). When concepts are presented in a compositional manner, mismatch stimuli should be more difficult to verify, because of the interference introduced by activating incorrect out-of-context weight information (e.g. if people have already activated the context-inappropriate feature *heavy* for door, then it should be more confusing for them to assess the weight of *birdcage door*). In contrast, the contextual view suggests that, if concepts are presented in context, then there should be no difference between match and mismatch conditions, as inappropriate features are not brought to mind.

Because the compositional view maintains that concepts are always activated independently, it therefore predicts that mismatch stimuli will be more difficult to activate for both the compositional and contextual conditions. In contrast, the contextual view predicts that there will only be a difference in the compositional condition, since inappropriate weight information is not activated in the contextual condition. The contextual view therefore predicts a two-way interaction between presentation (i.e. compositional and contextual) and matching (i.e. match and mismatch).

Method

Materials We identified a set of concepts which were reliably evaluated as lighter or heavier than 1kg. Candidate concepts were presented to a group of 63 undergraduate participants who rated them as either light or heavy. Any concept which failed to garner a two-thirds majority consensus was discarded. Subsequently, we generated a set of candidate modifier-head combinations, intended to be either light or heavy. These were presented to another group of 93 undergraduate participants who rated them using the same paradigm. Again, only those garnering a two-thirds majority consensus were considered for inclusion.

A total of 80 modifier-noun combination phrases were generated for the experiment, including 20 matching combinations, 20 mismatching combinations and 40 control combinations. Of the 20 matching combinations, 10 were light and 10 were heavy. Of the 20 mismatching combinations, 10 were altered from light to heavy by the modifier and 10 were altered from heavy to light. Each item in the matching condition was paired with another item in the mismatching condition using the same head noun (e.g. *moth net* and *trawler net*). The 40 control combinations were included to balance out the modifiers. Each combination in the matching and mismatching conditions was paired with another combination in the control condition. These controls used the same modifier but had the opposite weight (e.g. *compost fork* and *compost pellet*). As recommended by Solomon and Barsalou (2004), this experimental design ensured that no heuristic based on modifier or head words alone could result in better than chance performance: each modifier and each head was associated with precisely one light combination and one heavy combination. There was no significant difference in the log Google frequency of the combinations in the match and mismatch conditions, $F(1,19) = .32$, $p = .58$, $MS_e = .46$. In addition, none of the stimuli was lexicalized, as verified using a Google definition search (e.g. 'define: dog ball'). Although Google is by no means a definitive source, it nevertheless provides a strict criterion for lexicalization. A set of sample stimuli are shown in Table 1.

Table 1: Sample experimental stimuli

	Match	Mismatch	Control	
Light → heavy	*dog ball*	*iron ball*	*dog cage*	*iron nail*
Heavy → light	*lawn tractor*	*toy tractor*	*lawn daisy*	*toy slide*

The experiment involved a two-way mixed model. The independent variable of presentation (i.e. compositional or contextual) was a between-participants within-items measure while the independent variable of matching (i.e. match or mismatch) was a within-participants within-items measure. The dependent variable of response time was used to infer ease of activation.

Procedure The stimuli were divided into two lists, each of which included 5 light matching, 5 heavy matching, 5 light mismatching and 5 heavy mismatching items. Each list included the full set of head concepts and a balancing set of modifiers. The presentation of each list was counterbalanced across the match and mismatch conditions.

Participants were randomly assigned to either the compositional or contextual condition and were presented with one of the two lists. Examples of objects that weighted 1kg were provided, including a liter of milk and a 1kg bag of sugar. Participants were then instructed to place 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, pressing F for light and J for heavy. They conducted a number of practice trials in order to familiarize them with the task.

Combinations appeared in the middle of the screen and participants had to make a decision by pressing the appropriate key. In the compositional condition, the unmodified head concept initially appeared on the screen in

isolation and participants evaluated its weight. Immediately after responding, the modifier then appeared to the left of the head and participants evaluated the weight of the modified concept. In the contextual condition the modifier and head appeared together. Each trial was separated by a blank screen lasting for one second.

Participants 80 undergraduate students from NUI Maynooth participated voluntarily in this experiment. All were native English speakers.

Results and Discussion

The mean response times for the compositional and contextual conditions were 1,214 and 1,452 ms respectively while those for the match and mismatch conditions were 1,345 and 1,320 ms respectively. Figure 1 provides a breakdown.

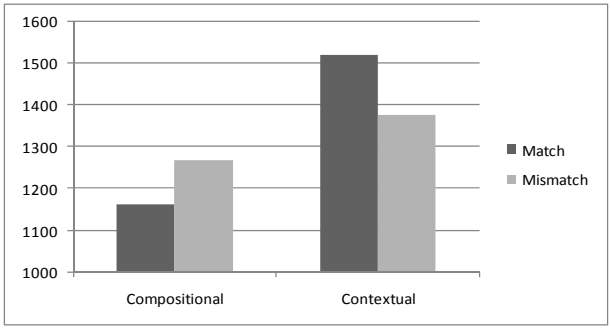

Figure 1: Mean response times by condition

We carried out a series of ANOVAs analyzing the effects of the independent variables on response time. As predicted by the contextual view, there was a reliable interaction between presentation and matching for response time both by-items, $F(1,19) = 8.36$, $p = .009$, $MS_e = 33752.00$ and by-participants, $F(1,39) = 9.04$, $p = .004$. There was a main effect of presentation on response time both by-items, $F(1,19) = 49.36$, $p < .001$, $MS_e = 22704.62$ and by-participants, $F(1,39) = 1361.24$, $p < .001$, $MS_e = 204960.08$, due to the priming effect exerted by the unmodified concept in the compositional condition. There was no reliable effect of matching on response time either by-items, $F(1,19) = .012$, $p = .92$, $MS_e = 138969.96$, or by-participants, $F(1,39) = .03$, $p = .87$, $MS_e = 38843.16$. We conducted a series of planned contrasts investigating the effects of matching on ease of activation for each presentation condition. As predicted by the contextual view, there was a reliable effect of matching on response times in the compositional condition, $F(1,39) = 5.76$, $p = .02$, $MS_e = 27167.52$, but no other reliable effects (all ps $> .05$). An unexpected finding was that, in the contextual condition, the mean response time was higher for matching than mismatching items (though the difference was not significant). It may be the case that mismatching items had a faster response time because these were necessarily more distinctive than the matching items (having features which differed from the default features for the head word). This greater distinctiveness may have led to such items being retrieved more quickly relative to the less distinctive matching items.

The results reveal reliable differences in how stimuli were processed in the compositional and contextual conditions. Specifically, participants in the compositional condition found it more difficult to assess the weight of mismatching combinations: assessing the weight of an unmodified concept first made it more difficult to assess the weight of a modified concept with the opposite weight (e.g. assessing the weight of *fork* made it more difficult to assess the weight of *compost fork*). In contrast, participants in the contextual condition were able to avoid this confound: they did not find it more difficult to assess the weight of mismatching combinations. These results indicate that participants in the contextual condition did not activate context-inappropriate weight information for modified concepts (e.g. when evaluating the weight of *manhole lid* they did not activate the weight of *lid*). Accordingly, the results argue in favor of the contextual view: when words are presented individually they are processed differently to when they are presented in combination.

General Discussion

Communicating a comprehensive account of even the simplest sensory experience using words is a difficult task. According to Shannon's (1948) information theory, the amount of information which must be transmitted in order to convey a message is dependent on the amount of information shared by the sender and the receiver. The effectiveness of language is therefore dependent on the assumption that speakers share a considerable volume of knowledge. For example, if a label is agreed upon by two speakers to reference a particular shared mental representation, then the use of that label by one speaker has the capacity to index that same mental representation for the other speaker. In this way, a complex concept like *apple* can be communicated using a single word. Information theory reveals that words by themselves carry no inherent informational content; instead, the meaning of a word emerges from the way in which an interpreter uses it to diagnose a subset of their own personal experience.

An important question regarding the process by which speakers derive meaning from words is whether the significance of a word is static or whether it is context-sensitive. The compositional view proposes that words index a fixed subset of experiential knowledge (e.g. encountering the word *apple* always causes the same set of knowledge to be retrieved). On the other hand, the contextual view proposes that people's understanding of a word is more accurately viewed as an elaborate function which takes contextual factors as inputs in order to compute a more precise referent. For example, people's understanding of the word *apple* is likely to take into account how it can serve to denote different mental representations in different situations (e.g. *apple* in the context of a pie can be a soft filling; *apple* in a business

context can refer to a company). Given that the ultimate interpretation of a phrase is non-compositional, it would seem more efficient for speakers to take contextual effects into account immediately, rather than activating context-inappropriate interpretations and then applying a subsequent 'cleaning up' process. Much cognitive effort could be avoided by applying even the most intuitive of heuristics describing how context affects meaning.

A study by Barsalou (1987) offers support for the idea that context influences how the meaning of a word is derived. He instructed participants to adopt different points of view while making judgments about the typicality of different members of a category. For example, participants were asked to judge a list of vehicles for how typical they would seem from the point of view of a suburban housewife as opposed to that of a farmer. Barsalou found that the ranking of typicalities changed radically according to point of view: participants consistently recognized that *vehicle* was more likely to reference a tractor for the farmer than for the housewife. Barsalou concluded that word meanings are unlikely to be stored as detached descriptions involving fixed sets of propositional features. Instead, he suggested that understanding a word entails the ability to produce a wide variety of situated conceptualisations that support goal achievement in specific contexts.

The idea that people's understanding of words is more than the sum of their individual interpretations can also explain the observation that compound phrases often activate features which are not associated with the individual words. For example, Hampton (1987) asked participants to produce lists of 30 common properties for birds, for pets, and for the conjunction 'birds that are pets'. A striking result was that people generated attributes for the conjunction that had not been considered true of either class considered alone. For examples, birds and pets do not live in cages, yet pet birds do. Similarly, birds and pets do not talk, yet pet birds do. Such observations are more naturally explained by a contextual view of language understanding.

Why then does it seem so intuitive to associate a word's meaning with an atomic definition? We propose that this tendency arises because of a confusion of word *meaning* and word *interpretation*. When people are asked to provide a definition for a given word, they are inclined to describe the mental representation that it serves to index in a context-free setting. However, people's understanding of a word goes much further than this, allowing them to arrive at a wide variety of different interpretations across a range of contexts. For example, if one is asked to define one's understanding of the word *red*, one will typically identify a prototypical red colour. However, one's actual understanding of the word *red* extends to identifying different forms of red in different contexts, such as *red hair* or *red skin*. Similarly, one's understanding of the word *fruit* is not simply the mental representation activated by the word *fruit* in isolation: it also encompasses the ability to construct different mental representations across a variety of contexts (e.g. *tropical fruit, inedible fruit*).

In sum, we propose that word meanings are complex rather than atomic: the understanding of a word is more accurately viewed as a theory which describes how it is likely to be used in different situations, rather than a static representation which is invariably called to mind regardless of the context. While it is straight-forward to provide an interpretation for a word, it is far more challenging to comprehensively state one's understanding of it. People are thus largely unaware of the rich, complex understanding they have of how words and context combine to index meaning.

Conclusion

Evidence in support of the compositional and contextual views has thus far been equivocal. In this article we have described an experiment designed to avoid some of the methodological issues which have arisen in previous studies. The results provide clear evidence in support of contextuality: we have shown that the way in which people interpret words in context differs reliably from the way in which they interpret the same words in isolation. In light of this finding, we have suggested that language understanding is complex, in that people are aware of how words can serve to address different aspects of semantic memory in different contexts, thus avoiding the activation of context-inappropriate features.

References

Barsalou, L.W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82-93.

Barsalou, L.W. (1987). The instability of graded structure in concepts. *Concepts and Conceptual Development,* 101-140. New York: Cambridge University Press.

Fockenberg, D.A, Koole, S.L. & Semin, G.R. (2006). Backward affective priming: Even when the prime is late, people still evaluate. *Journal of Experimental Social Psychology, 42*(6), 799-806.

Hampton, J.A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory & Cognition, 15*, 55-71.

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

Potter, M.C., & Faulconer, B.A. (1979). Understanding noun phrases. *Journal of Verbal Learning and Verbal Behavior, 18,* 509-521.

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

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

Swinney, D., Love, T., Walenski, M. & Smith, E.E. (2007). Conceptual combination during sentence comprehension: Evidence for compositional processes. *Psychological Science, 18, 5*, 397-400.