Spatial Information and Representations of Word Meaning: Accessing Semantic Size during Reading

Renske S. Hoedemaker

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Psychology

Chapel Hill
2012

Approved by:
Peter C. Gordon
Neil Mulligan
Jennifer Arnold
ABSTRACT

RENSKE S. HOEDEMAKER: Spatial Information and Representations of Word Meaning: Accessing Semantic Size during Reading (Under the direction of Peter C. Gordon)

Theories of embodied cognition have proposed that language is understood through perceptual simulation of the sensorimotor characteristics of its meaning. This claim suggests the activation of sensorimotor representations is encoding-based. Participants in three eye-tracking experiments were presented with triplets of either numbers or object and animal names. When the task was to compare the referent size of the items, word and number decision times showed a symbolic distance effect, such that response time was inversely related to the size difference between the items. When semantic size was irrelevant to the task it had no effect on word encoding times. Number encoding times showed a distance priming effect: encoding time increased with numerical difference between items. Together these results suggest that while activation of numerical magnitude information is both encoding-based and goal-driven, size information associated with words is goal-driven and does not occur automatically during encoding.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Analog Representations of Meaning in Contemporary Cognitive Psychology</td>
<td>2</td>
</tr>
<tr>
<td>Embodied Activation of Word Meaning</td>
<td>4</td>
</tr>
<tr>
<td>Analog Representations of Magnitude</td>
<td>9</td>
</tr>
<tr>
<td>Current study</td>
<td>12</td>
</tr>
<tr>
<td>Experiment 1</td>
<td>14</td>
</tr>
<tr>
<td>Method</td>
<td>15</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>21</td>
</tr>
<tr>
<td>Experiment 2</td>
<td>30</td>
</tr>
<tr>
<td>Method</td>
<td>31</td>
</tr>
<tr>
<td>Results and Discussion</td>
<td>32</td>
</tr>
<tr>
<td>Experiment 3</td>
<td>36</td>
</tr>
<tr>
<td>Method</td>
<td>36</td>
</tr>
<tr>
<td>Results</td>
<td>38</td>
</tr>
<tr>
<td>Discussion</td>
<td>47</td>
</tr>
<tr>
<td>General Discussion</td>
<td>50</td>
</tr>
<tr>
<td>Sensorimotor Representations and Action-Oriented Language Comprehension</td>
<td>53</td>
</tr>
<tr>
<td>Future Directions</td>
<td>52</td>
</tr>
<tr>
<td>References</td>
<td>59</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. Presentation of stimuli and dependent measures in Experiment 1 and 3 ...............20

Figure 2. Decision and encoding times in Experiment 1 .....................................................24

Figure 3. Individuals’ mean middle-last ratio by the slope of the size-based effect on the middle word in Experiment 1 .....................................................................................28

Figure 4. Multiple regression of single fixations duration in Experiment 2 .........................34

Figure 5. Individuals’ mean middle-last ratio by the slope of the size-based effect on the middle word in Experiment 3 .....................................................................................42

Figure 6. Encoding and decision times for words and numbers in Experiment 3 ...............45
Introduction

When reading a word, we are able to access different aspects of its meaning. On the one hand, we may understand a word that has a concrete referent by virtue of its connection to other internal symbols or associations that provide information about categories, semantic relations, or linguistic co-occurrence. On the other hand, we may understand that word’s meaning through our experience with its referent’s sensory characteristics such as color, shape and size. Recently, the field of embodied cognition has assigned a central role to sensorimotor experiences as the source of all conceptual knowledge. It is argued that knowledge of concepts emerges from perceptual experience and is mentally represented in the very same sensorimotor systems that originated the experiences rather than in an amodal semantic faculty (Barsalou, 1999; Glenberg & Robertson, 2000; Zwaan, 2004; Zwaan & Madden, 2005).

Embodied cognition as a field argues against the view that perception and cognition reflect independent systems in the mind (Barsalou, 1999). Instead, it is based on the assumption that cognitive processes consist entirely of sensorimotor processes. Cognition is not the manipulation of abstract symbols, but the interaction with analog, modal representations. Even in the absence of a direct interaction with the environment, sensorimotor functions are recruited to support cognitive activity in the form of a ‘mental simulation,’ in order to represent information or support inferences (Wilson, 2002). Within
this framework, language comprehension is established through the multimodal mental simulation of the situation described in an utterance.

**Analog Perceptual Representation in Contemporary Cognitive Psychology**

In contrasting embodied theories of language comprehension with traditional or ‘disembodied’ approaches, it is tempting to present the latter as allowing no role whatsoever for modality specific representations of meaning. Within traditional approaches of the study of cognition, mental representations have sometimes been conceptualized as arbitrary, abstract symbols that are inherently amodal (Pylyshyn, 2003). Within this framework, language functions as a system that connects external representations, such as words and sentences, to internal representations of meaning. Importantly, these internal representations are qualitatively distinct from representations in our sensorimotor system (e.g. Fodor, 1985).

However, several fields within contemporary cognitive psychology have traditionally recognized the role of analogue perceptual representations. Baddely (1986) assigned a specialized system to visuospatial information within working memory, and long term episodic memory is considered to rely crucially on multimodal simulation of previous events (Paivio, 1986; Conway, 1990). Within the study of memory and learning, implicit learning (Jacoby, 1983) repetition priming (Kirsner et al., 1989) and transfer appropriate processing (Morris, Bransford & Franks, 1977) are considered to depend largely on the amount of overlap in perceptual circumstances of the learning and testing episode. In addition, mental imagery is considered to play an important role in the representation of concepts, as has been found in studies of categorization and property verification (Kosslyn, 1980). Within language research, spatial representations are considered crucial to the construction of mental models (Bower & Morrow, 1990; Bransford, Barclay & Franks, 1972; Glenberg, Meyer & Lindem,
Research focused on reasoning and problem solving has shown we often rely on analog visuospatial imagery when confronted with problems that involve physical reasoning (Hegarty, 2004; Schwartz & Black, 1999) or mental rotation of objects (Shepard & Metzler, 1971).

This selection of examples demonstrates an established tradition within cognitive psychology focusing on the role of analogue representations of meaning. The real difference between the embodied and symbolic accounts can be framed in terms of the level of language processing at which these representations are considered to play a role. In traditional models of language processing, linguistic input progresses through several representational stages of increasing complexity. Word meaning at the lexical level is represented in terms of abstract, arbitrary symbols as well as the connections between symbols such as semantic networks, feature lists, and schemas. Once the word’s meaning has been accessed at this level, it may be introduced into the discourse model (Johnson-Laird, 1983), which represents the discourse referents, patterns of coreference and how the referents relate to each other (Kintsch & Van Dijk, 1978). Creating an effective discourse model requires representation of logical relations among the referents, including but not limited to causal, temporal and spatial relations. The reader’s background or world knowledge is also incorporated at this level, allowing for elaborative and predictive inference generation (Bower & Morrow, 1990). Therefore, representations of the meaning of a text or sentence at the discourse or task level are likely to involve at least some analog, modality specific components.

The embodied approach, however, considers word meaning to consist of analog, modality specific sensorimotor representations at all levels of conceptual processing. According to Barsalou’s perceptual symbol systems hypothesis (1999), words activate
simulators associated with the perceptual properties of their referents, creating a perceptual simulation of a text or utterance that preserves the analog properties of what is represented. Similarly, the immersed experiencer framework (Zwaan, 2004) describes language comprehension as the ‘vicarious experience of the described situation’ (p. 36). This simulation of perceptual experience is explicitly theorized to take place at the word and even the morpheme level of comprehension. The vicarious experience is created when words activate networks of meaning or ‘functional webs’ that are also activated when the word’s referent is experienced perceptually. The simulated experience therefore involves more perceptual information than an abstract network of semantic meaning may be able to represent at the word level. Importantly, perceptual simulation within the embodied framework is not an optional aid to comprehension. Instead, representation of concepts is inherently perceptual and comprehension consists of perceptual simulation or immersed experience.

Embodied Activation of Word Meaning

Demonstrations of activation of sensorimotor properties of word meaning during reading and spoken language comprehension provide important empirical support for the embodied account of language processing. Richardson, Spivey, Barsalou and McRae (2003) found that during comprehension, verbs can activate spatial representations based on the spatial properties of the expressed action. Participant responded to visual stimuli presented in different areas of the visual field, while listening to sentences containing verbs that typically express either horizontal (e.g. push) or vertical motion (e.g. lift). This study found that processing of the visual stimuli was inhibited when presented in the area of the visual field congruent with the verb’s direction of motion. Similar results were found for subject nouns
that canonically refer to objects in the upper (e.g. *sky*) and lower (e.g. *grass*) portion of the visual field (Bergen Lindsay, Matlock & Narayanan, 2007). In both studies, the spatial properties of the sentence were not considered relevant to the visual task, which suggests that the observed effects occurred during the online, automatic comprehension process.

The notion that encoding of linguistic stimuli involves rapid activation of spatial information has been explored intensively in a series of picture-sentence mapping studies by Zwaan and colleagues on what has been named the ‘appearance-compatibility effect.’ These authors found that after reading a sentence about a specific object, participants were faster to respond to pictures that matched the object’s shape or orientation as it had been described in the sentence. For example, reading ‘*the carpenter pounded the nail into the wall*’ resulted in faster response times for a picture of a nail in a horizontal compared to a vertical position (Stanfield & Zwaan, 2001; Wassenburg & Zwaan, 2010) and after reading about an eagle in the sky participants were quicker to respond to a picture of an eagle with its wings spread out than one with its wings folded in (Zwaan, Stanfield & Yaxley, 2002). Similarly, knowledge of the canonical shape of objects can affect eye-movements within a visual-display during visual search tasks, so that upon hearing the word ‘Frisbee,’ a distracter picture of a slice of pizza drew more looks than unrelated distractor pictures (Yee, Huffstetler & Thompson-Schill, 2011).

Besides visuospatial information, language can also interact with motor processing. Glenberg & Kaschak (2002) found that during a sentence plausibility judgment task, participants responded more quickly when the action required to respond matched the direction of movement described in the sentence. For example, reading the sentence ‘*close the drawer,*’ which implies an action oriented away from the body, participants responded
more quickly when the correct response required pushing a button away from the body. This finding, coined the ‘action-compatibility effect,’ suggests that language is understood at least in part through the activation of a sensorimotor representation of the utterance.

Together, these studies and many others effectively demonstrate activation of visuospatial and motor representations during sentence or word comprehension. Nevertheless, they do not provide specific evidence about the level of language processing at which this activation takes place while also leaving an unclear picture of the role of task-goals in relation to these findings. According to the embodied language comprehension account, analog symbols extracted from perceptual experience constitute conceptual meaning (Barsalou, 1999), and activation of ‘experiential representations’ takes place at the word level of language comprehension (Zwaan, 2004). “Whereas such knowledge is cumbersome and brittle in amodal symbol systems…it ‘rides for free’ in perceptual symbol systems” (Barsalou, 1999, p. 604, quotations in original). As such, the embodied cognition account predicts encoding-based activation during the reading of words, irrespective of higher level processes related to context or task-specific goals.

Critics of the embodied account have countered that demonstrations of sensorimotor activation during conceptual processing do not provide evidence that sensorimotor representations are constitutive of the concept or even conceptually relevant (Adams, 2010; Mahon & Caramazza, 2008; Weiskopf, 2010), which would be expected if the activation were task-independent and encoding-based. As argued by Mahon and Caramazza (2008), the observed sensorimotor effects may not be related to word meaning accessed during encoding, but instead result from activation originating in abstract, disembodied representations spreading into systems capable of sensorimotor representation. This cascading activation
may result in fast, somatotopic activation of the sensorimotor system upon processing words referring to concrete objects or physical action, but this observation does not in itself provide evidence that such activation is constitutive of the word’s conceptual representation.

In support of this argument, Mahon and Caramazza (2008) provide an example from speech production studies. Naming latency measurements have shown that when participants are asked to name pictures (e.g. a picture of a hammock), phonologically related distractor pictures (e.g. a hammer) result in faster target naming times compared to unrelated distractor pictures (e.g. a button) (Navarrete and Costa, 2005). These results are taken as evidence that even though the participants intend to ignore the distractor picture, its presence results in fast automatic activation of phonological information associated with the word ‘hammer.’

Similarly, studies of embodied language comprehension may show that seeing a picture of a hammer activates motor programs associated with handling tools, or areas of the motor cortex involved in actions related to the hand and arm. Within embodied cognition, such motor activation upon seeing a hammer is taken as evidence that the concept hammer is represented within these sensorimotor systems. However, the automatic online activation of phonological aspects of unproduced words is never taken to suggest that the part of the conceptual meaning of hammer is represented in the phonological system. As such, the analogy clarifies that the observation of motor activation related to handling hammers upon seeing one pictured does not provide sufficient evidence that this activation is critically related to understanding the concept HAMMER.

An embodied cognition rebuttal might of course point out that the phonologically related distractor picture results in activation of the label ‘hammer’ rather than the concept itself. Observed motor activation related to actually using a hammer is arguably a
semantically meaningful component of the concept HAMMER, whereas phonological or
label-based activation is not. Nevertheless, the example illustrates the need to reconsider the
idea that observed sensorimotor activation during language processing must be conceptually
relevant information activated as part of the encoding process. A disembodied account of
language comprehension can explain appearance- and action-compatibility effects promoted
as evidence for the embodied account by assuming flow of activation within and between
cognitive systems, urging a careful consideration of how this activation originates as well as
what drives its spread into particular sensorimotor and conceptual systems.

Importantly, the hammer-analogy also draws attention to the importance of task-goals
during linguistic and semantic processing. Phonological distractor pictures lead to
phonological activation during word-naming tasks, while sensorimotor activation may be
observed during semantic categorization. In both cases, the type of information accessed
facilitates performance on the task at hand. As emphasized by several critics of the embodied
account, it is likely that goal-driven recruitment of sensorimotor areas facilitates performance
on conceptual tasks (Mahon & Carammaza, 2008, Weiskopf, 2010). However, this does not
suffice to claim the information is inherently a part of the conceptual representation
HAMMER, irrespective of task-specific comprehension goals. In other words, there may be
circumstances in which an adequate understanding of HAMMER does not require the
involvement of any sensorimotor components of meaning. Therefore, strong support for the
embodied cognition claim that language comprehension at the word-meaning level consists
of analog mental simulation, requires that activation of sensorimotor information relevant to
the word’s conceptual meaning, is observed during encoding, regardless of task or goal-
related processes. Evidence that would allow for an empirical distinction between embodied
mental simulation and ‘disembodied’ flow of activation accounts of observed congruent sensorimotor activation during language processing requires tasks that distinguish (task-independent) *encoding-based* from (task-related) *goal-driven* activation of sensorimotor representations during language processing. The study of how we encode and compare representations of referent magnitude provides the necessary methodological tools to make such a distinction.

**Analog representations of magnitude**

A variety of useful tasks have been developed that provide insight into the processing of representations of semantic magnitudes associated with symbols such as words and numbers. Tasks involving comparison of magnitudes have consistently shown the *symbolic distance effect* -- the finding that the time required to determine the larger of two visually presented numerals is an inverse function of the difference between the two numerals (Henik & Tzelgov, 1982; Koechlin, Naccache, Block & Dehaene, 1999; Moyer & Landauer, 1967; Van Opstal, Gevers, De Moor & Verguts, 2008). For example, choosing the larger number of the pair 2 and 8 takes less time than making the same choice for the pair 6 and 8. Reaction times depend more strongly on the ratio of the pair of numbers than on their absolute difference and in this way resemble a classic psychophysical function. The same effect occurs when participants compare lexical stimuli instead of numbers, such that choosing the larger or smaller of a pair of animal names is facilitated for pairs with a large difference in size (lobster – cow) compared to a small difference in size (sheep – cow). This finding is impressively robust (Banks & Flora, 1977; Dean et al., 2005; Holyoak, Dumais & Moyer, 1979; Moyer & Bayer, 1976) and occurs for a variety of other types of words and dimensions (Dean et al., 2005; Holyoak & Walker, 1976; Paivio & Te Linde, 1980; Te Linde & Paivio,
The pattern of reaction times is similar to that found for actual perceptual comparisons such as choosing the longer of two lines (Johnson, 1939) or the larger of two circles (Moyer & Bayer, 1976). Therefore, the symbolic distance effect is generally thought to indicate a conversion process of abstract numbers or words into analog representations on the dimension of interest, such as numerical magnitude or animal size. The comparison decision is then made based on these analog representations, constituting an “internal psychophysical judgment” (Moyer, 1973 p.183).

While symbolic comparison effects provide insight into the way that we process symbolic stimuli, the trial-final reaction times measured in these tasks do not provide direct insight into the encoding of symbols because they are influenced by other task-relevant processes such as comparison, decision making and response selection. In contrast, distance priming tasks are thought to measure the effects of encoding-based activation. The numerical distance priming effect is the finding that responses to a numerical target are faster when it is preceded by a number that is close in magnitude than by one that is distant. For example, processing the digit 6 is facilitated when preceded by the prime 4 compared to the prime 2 (Brysbaert, 1995; Den Heyer & Briand, 1986; Marcel & Forrin, 1974; Reynvoet, Brysbaert & Fias 2002). This effect is found both when the prime requires a deliberate response and during short SOA masked priming (Koechlin, Naccache, Block, & Dehaene, 1999); it has been found with verbal numerals (*four, six*) as well as Arabic numerals and non-symbolic quantities such as dot patterns (Herrera & Macizo, 2008). In addition, the effect occurs even when successful execution of the task does not require activation of the number’s magnitude (Brysbaert, 1995; Van Opstal & Verguts, 2011). Importantly, comparison to letter stimuli has shown that numerical distance priming relies on activation of
magnitude rather than simply order information, since the (alphabetical) order information associated with letters leads to comparison distance effects, but not distance priming (Van Opstal, Gevers, De Moor & Verguts, 2008).

Few studies have explicitly investigated whether analog representations of size associated with lexical items that refer to objects also lead to size-based priming effects during encoding, and those that do leave us with an unclear picture of the role of task-related and goal-driven processes involved in such effects. Schreuder, Flores d’Aracais, and Glazenburg (1984) showed that words were processed more quickly when preceded by a word with shared perceptual qualities, such that orange could prime ball. However, Pecher, Zeelenberg & Raaijmakers (1998), found that perceptual priming occurred only when the perceptual qualities of the stimulus words were first made salient to the participants through a perceptual judgment task, pointing towards a task or context dependent activation of perceptual properties. However, many of the perceptually related stimulus pairs had different sizes (eg. pizza-coin, bracelet-hoop), which may have eliminated perceptual priming. In addition, the associative priming effect (studied for comparison) was quite small, raising concerns about whether a relatively small perceptual priming effect could be detected in the non-salient condition.
Current Study

In the symbolic distance effect and numerical distance priming, proximity in the semantic magnitude of two stimuli have opposite effects on reaction time, and for this reason the two effects are thought to measure different levels of processing. The symbolic distance effect, or facilitation of comparison decisions for numbers that are numerically distant, is considered to take place at the task or goal-related level of comprehension. At this level, symbolic comparison tasks with lexical items robustly show reaction time patterns similar to numbers, suggesting that both numerical and lexical symbolic comparison decisions are based on analog representations. As with other priming effects, numerical distance priming is observed at encoding, thus providing substantial evidence for the analog representation of numerical semantic magnitude at the encoding or word-meaning level. This provides the opportunity to distinguish between activation at the encoding and goal-driven stages of language comprehension based on distinct patterns of reaction time within a single task.

Brysbaert (1995) created a novel numerical-priming paradigm that is especially well suited to the goal of distinguishing encoding-based priming from goal-driven comparison effects. Participants in an eye-tracking experiment were presented with triplets of numbers and asked to indicate whether the middle number was numerically in between the two outer numbers. In conjunction with eye-tracking, the use of three numbers, rather than the typical two, allowed encoding and decision processes to be isolated in a single task. Encoding of the middle number takes place before a comparison decision must be made, since all information
needed to make a correct decision is available only after the third number has been processed. As such, time spent looking at the middle number provided a measure of encoding time. A gaze-contingent display technique was used in which the numbers were masked except when the participant looked at them during their first-pass reading from left to right. This eliminated preview and rereading effects that might have reduced the value of looking time as a measure of encoding. Brysbaert observed a numerical distance priming effect on the middle number, such that its encoding time was reduced on trials with a small numerical difference between the first and middle number, compared to trials with a larger difference. Crucially, the same effect was observed when, instead of judging magnitude-order, participants simply encoded each triplet into memory for an immediate recognition task, demonstrating encoding-based activation of semantic magnitude associated with numbers independent of goal-based activation related to the reader’s task. The current study extends Brysbaert’s distance-priming paradigm to words referring to objects in order to assess whether activation of magnitude information about semantic size associated with lexical items is encoding- or goal-based.
Experiment 1

This experiment applied Brysbaert’s (1995) gaze-contingent triplet-comparison task to object and animal names. As illustrated in Figure 1, participants read three words, all referring either to objects or animals, and then indicated with a speeded key-press whether the size of the referent of the middle word was between the sizes of the referents of the first and last word. The three object or animal names were shown from left-to-right on a screen, and a gaze-contingent display technique was used so that each name could only be encoded during the eyes’ first-pass over that word. Triplets of words were constructed so that the distance in the semantic size of the second and first items was small, medium or large. The task provides two measures of interest: (1) Decision time is the interval between when the eyes first look at the final name in the triplet and the execution of the manual response; it includes goal-based activation of task-relevant semantic properties since it covers the time during which all information needed to perform the task is available. (2) Encoding time is the gaze duration on the middle word; for numerals in Brysbaert’s study, this measure showed distance priming indicating that it reflected encoding-based activation of numerical magnitude information.

To our knowledge, theories of embodied cognition have not previously been applied to distance-priming or the symbolic distance effect. However, these theories have proposed that sensorimotor representations are the foundation of the meaning of language and that language comprehension involves perceptual simulations of the meaning of what is being understood (Barsalou, 1999; Glenberg & Roberts, 2000; Zwaan, 2004; Zwaan & Madden, 2004).
2005), rather than having sensorimotor representations activated based on task goals as critics have contended (Mahon & Caramazza, 2008; Adams, 2010). As such, we believe that the embodied cognition approach leads naturally to the prediction for this task that distance priming should be found for encoding time on the middle object or animal name as Brysbaert (1995) found for numerals, because encoding a name should involve creation of a representation that captures the perceptual properties of its referent. In contrast, the goal-driven activation account suggests that the visuospatial characteristics of a word’s referent are not activated automatically during encoding but that instead this information is accessed in a goal-driven manner at a later stage of language processing. Therefore, the goal-driven activation account does not predict distance priming on encoding time and instead suggests that any effects of semantic distance should be understood in terms of how semantic distance affects the processes that are invoked by the task goals. Both the embodied-cognition and goal-driven activation approaches lead to the prediction that the decision-time measure should show a symbolic distance effect since the difficulty of comparing the size of perceptual representations is inversely related to the difference in their sizes (Moyer & Bayer, 1976; Banks & Flora, 1977; Holyoak et al., 1979; Dean et al., 2005).

**Method**

*Participants.* Twenty-four undergraduates at the University of North Carolina at Chapel Hill participated for course credit. All were native English speakers with normal or corrected-to-normal vision and were naïve about the research goals.

*Stimuli.* Experimental stimuli consisted of triplets of object and animal names selected from size-rating norms for animal (Dean et al., 2005) and object names (Holyoak et al., 1979). Objects and animals were selected with two constraints: (1) Across items there
was a substantial range in rated size that was evenly distributed so that equal numbers of items could be placed in each of three size groups (small, medium and large), and (2) the average standard deviation in size ratings for an item was low, indicating that size representations were similar across raters. Cross-individual consistency in size estimates was further validated by examining the reliability of ratings for the 21 animal names that were normed in both Dean et al. and Holyoak et al. Although the two studies used different rating scales, had different rating instructions, and were collected more than twenty years apart, the two sets of norms were highly correlated, \( r = .956, p < .001, R^2 = .91 \).

The stimulus triplets were constructed as follows. Twenty-one animal names referring to animals across a range of sizes served as targets appearing in the middle position of a triplet, while the remaining 25 animal names served as flankers. Triplets were constructed so that the semantic size difference between the first and the second animal was small (about 1 target standard deviation), medium (about 2 target standard deviations) or large (3 or more target standard deviations). In addition, the first item was either larger or smaller than the target item, so that there were an equal number of ascending and descending triplets, resulting in a \( 3(\text{size difference}) \times 2(\text{direction of the size difference}) \) within-subjects design. The third word in the triplet was added so that for half of the trials, the size of the middle animal was in between the size of the two outer animals, resulting in a ‘consistent’ triplet (e.g. eagle – cow – elephant). On the other half of the trials, the size of the middle item was not in between the two outer items, resulting in an ‘inconsistent’ triplet (e.g. giraffe – cow – elephant). The same procedure was followed for the object triplets, using 21 target and 37 flanker objects.
A norming study was conducted to measure the semantic relatedness of all the word pairs occurring within the triplets. Thirty participants were presented with the pairs one at a time and asked to rate the association between the items in the pair (“Indicate the degree to which you feel these two items are associated. For example, how often do you think of the two items together, do they occur together in the real world, or are they mentioned in the same context?”), answering on a scale of 1 (not at all associated) to 7 (entirely associated). The order of the words in each pair was counterbalanced across participants. None of the participants in the norming study took part in any of the eye-tracking experiments reported here. Overall, the pairs were rated low on semantic relatedness, with a mean rating of 2.52 for small difference pairs ($SD = .62$), 2.48 for medium difference pairs ($SD = .62$) and 2.18 for large difference pairs ($SD = .53$). The slight difference in mean similarity ratings for pairs with small and medium size differences was significant by subjects but not by items, $t_1(29) = 2.82, p = .009; t_2(41) = .32, p = .749$. For pairs with medium and large size differences, the difference in similarity ratings was significant, $t_1(29) = 8.50, p < .001; t_2(41) = 2.65, p = .011$.

Previous research on semantic or associative priming suggests that words are processed more quickly when preceded by a semantically related or associated prime (Meyer & Schvaneveldt, 1971). Therefore, if the results of the experiment show reduced reading times for pairs within the small or medium size difference condition compared to the large size difference condition, it will be important to distinguish between effects caused by the size difference manipulation and differences in semantic relatedness.

Target items were counterbalanced across six lists, each consisting of six blocks of 42 trials, resulting in 252 trials per participant. In each block, all target words appeared once in random order, with animal and object triplets randomly intermixed. Across the blocks within
a list, each word appeared once in each condition, so that all participants saw each target in each condition. To control for possible effects of target repetition, each list started with a different block. Each experimental session started with twelve warm-up trials, which did not contain any of the target words used in the experimental trials. The warm-up trials were excluded from all analyses. To make sure participants fully understood the task, they were also presented with several practice triplets verbally by the experimenter before they were set up for eye-tracking.

Procedure. Eye movements were recorded from the participants’ right eye using an SR EyeLink 1000. Stimuli were presented on a 20 inch ViewSonic G225f Monitor at a distance of 61 cm with a display resolution of 1024 x 768. At the beginning of each session the tracker was calibrated using a 9-point procedure; calibration was checked between trials and the tracker was recalibrated when necessary. Participants sat in a well-lit room with a chin- and forehead rest minimizing their head movements. They were instructed to read the triplets silently and decide for each triplet whether the size of the middle object or animal was in between the two outer (first and last) objects or animals. Participants answered by indicating ‘yes’ or ‘no,’ using a handheld console. The experimenter monitored eye movements throughout the session.

Each trial started with a fixation point placed on the left side of the screen on the horizontal axis. Once this point was fixated, the next screen appeared, with the first word of the triplet slightly to the right of the fixation point. The middle and last word appeared in the center and on the right side of the screen on the same horizontal axis, masked by hash marks. Gaze-contingent invisible boundaries were placed approximately 120 pixels to the left of the middle and last word’s mask. Gaze contingencies were set up so that each word was visible
only when the eyes entered the word’s region from left to right, and was no longer visible after the eyes left its region to the right (see Figure 1). This method of stimulus presentation prevented potential parafoveal preview or rereading of the first and middle word. When the participants moved their gaze across the invisible boundary between the first and middle word, the middle word was unmasked and the first word was masked. The same event was repeated when the eyes crossed the invisible boundary between the middle and last word, so that the middle word was masked and the third word unmasked. Once the eyes left the first or middle word to the right, these items did not become visible again upon regressive eye-movements. The last word remained visible until a response was made. Participants received feedback on their accuracy after every trial.
Figure 1. Presentation of stimuli and dependent measures in Experiment 1 and 3. Encoding time is measured as the gaze duration on the middle word. Decision time is the time from the onset of the last word until the indication of a response by a speeded key-press. A gaze-contingent display technique was used in which the numbers were masked except when the participant looked at them during their first reading pass from left to right. This eliminated preview and rereading of the first and middle number.
Analysis of eye movements. Fixations shorter than 80 ms and within 1 degree of a longer, immediately subsequent fixation were merged with the longer fixation by an automatic procedure in the Eyelink software.

Data analysis focused on two standard eye-movement measures. Single-fixation duration was the average of the duration of the initial, first-pass fixation on a word given that the word was fixated on first-pass only once. Gaze duration was the average of the sum of all first-pass fixation durations on a word (Rayner, 1998). Following Brysbaert (1995), gaze duration on the middle word was taken as a measure of encoding time for this item, reflecting encoding-based activation of word properties. In addition to these widely used measures, Decision Time was the time from the onset of the third word until the button press indicating a response, reflecting goal-based activation. We also computed Middle-Last Ratio, which was the gaze duration of the middle word over the decision time. This provided a measure of the proportion of time spent on the middle word compared to the time spent on the last word and the task-decision on each trial.

Results and Discussion

The detailed data analyses reported below are based on the first three blocks performed by each participant. This is done to allow close comparison to the results of Experiment 3, where the addition of another task meant that only three blocks of data could be obtained in this part of the experiment. Analyses of all six blocks in Experiment 1 showed very similar patterns of results and all effects of interest remained significant\(^1\). Overall accuracy was 87%, and no participant had an average accuracy below 71%. Inaccurate trials were excluded from all analyses involving response time measures.
Mean decision times are shown in the left panel of Figure 2. They showed a strong symbolic distance effect, with times decreasing as the difference in semantic size between the first and middle word increased, $F_1(1,23) = 24.43, p < .001; F_2(1,41) = 17.82, p < .001$, for the linear effect of size difference. There was no difference in decision times between ascending and descending trials, $F_1(1,23) = 3.12, p = .09; F_2(1,41) = 2.18, p = .148$, though there was a marginally significant tendency for the effect of size difference to be stronger for descending than ascending trials, $F_1(1,23) = 4.48, p = .045; F_2(1,41) = 2.57, p = .117$. Decision times were faster for consistent trials (requiring a ‘yes’ response, $M = 2049$ ms) than for inconsistent trials (requiring a ‘no’ response, $M = 2241$ ms), $F_1(1,23) = 10.43, p = .004; F_2(1,41) = 10.39, p = .003$. Consistency did not interact with size difference, $F_1(1,23) = 1.14, p = .295; F_2(1,41) = .539, p = .467$.

As shown in the right panel of Figure 2, middle word encoding times also showed a symbolic distance effect, such that mean encoding time on the middle word was shorter when the semantic size difference between the first and middle word was larger, $F_1(1,23) = 11.73, p = .002; F_2(1,41) = 6.02, p = .018$ for the linear effect of size difference. Middle word encoding time did not show a difference between ascending and descending trials, $F_1(1,23) = 1.95, p = .674; F_2(1,41) = .002, p = .969$, nor was there an interaction between size difference and direction of the difference, $F_1(1,23) = .181, p = .674, F_2(1,41) = .919, p = .343$.

Observation of a symbolic distance effect on the middle words shows that the effect of size difference on encoding time for words was the opposite of the distance-priming effect that Brysbaert (1995) found for numbers. This pattern does not support encoding-based activation of size information, which we have argued leads to the prediction that distance priming should be observed for encoding times. The surprising finding of a symbolic
distance effect on encoding times is analyzed below in relation to individual differences in strategies for performing the triplet comparison task.
Figure 2. Decision and encoding times in Experiment 1. Mean decision times (A) and middle word encoding times (B) for trials with small, medium and large differences in semantic size between the middle and first item. Both measures show a symbolic distance effect. Gaze duration on the middle word and decision time are shorter when the difference in semantic size between items is larger.
Variation in individual strategies. Our analysis plan followed Brysbaert (1995) in operationalizing encoding time as gaze duration on the middle word, an approach that is consistent with practice in the reading literature where systematic effects of lexical properties on gaze duration have led to its being considered a primary measure of word recognition (Rayner, 1998). The finding that gaze duration on the middle word showed a symbolic distance effect indicates that in this task this measure may reflect both lexical encoding and goal-driven processing. This possibility was tested by examining whether the magnitude of a participant’s symbolic distance effect on encoding time for the middle word was related to his or her judgment strategy. The magnitude of the symbolic distance effect on encoding time was given by the slope of the linear effect of size difference between the first and middle words on middle-word gaze duration encoding time. More negative slopes indicate stronger symbolic distance effects, while a slope of zero indicates no distance effect and a positive slope would indicate distance priming. Judgment strategy was given by the a participant’s middle-last ratio, defined as the participant’s average encoding time on the middle word divided by his or her average decision time. Mean participant middle-last ratios ranged from .12 to .99. High ratios indicate that the participant spent a relatively long time encoding the middle word compared to the time spent on the last word until the response indicating a decision, a pattern that is consistent with the use of an incremental-judgment strategy where the semantic size of the middle word is mentally compared to that of the first word before the eyes move on to the final word. Lower ratios indicate long decision times relative to middle word encoding times, a pattern that is consistent with the use of a final-judgment strategy where mental comparison of sizes is postponed until the last word is seen.
The scatterplot in Figure 3 shows that the magnitude of a participant’s symbolic distance effect on encoding time was strongly related to his or her judgment strategy, with larger symbolic distance effects observed for participants pursuing an incremental-judgment strategy than for those pursuing a final-judgment strategy. This conclusion is supported by a robust negative correlation between participants’ middle-last ratios and the slopes of their size-difference effects, $r = -.706, p < .001, R^2 = .498$. Judgment strategy, as measured by middle-last ratio, was not related to overall speed measured by total time for trial completion, $r = .115, p = .593$. However, subject middle-last ratio was related to overall accuracy, with use of an incremental strategy showing a positive correlation with mean accuracy in the task, $r = .487, p = .016, R^2 = .237$. This supports the idea that individual participant strategies not only affect the extent to which participants exhibit the symbolic distance effect on the middle word, but also their level of task performance.

The scatterplot in Figure 3 also provides evidence about whether use of an incremental judgment strategy may have caused distance priming for the second word to be masked by mental comparison processes that occurred while the participant looked at the second word. If that were the case, then it would be expected that participant’s using a final-judgment strategy (indicated by a low middle-last ratio) might show distance priming, rather than the symbolic distance effect, on middle-word encoding time. Examination of performance for participants using a final-judgment strategy, shown toward the left in Figure 3, suggests that this was not the case. The slopes of the size-difference effect for these participants appear to be randomly distributed around zero, rather than showing a preponderance of positive slopes of the sort that would be associated with distance-priming effects. Statistical support for this conclusion comes from the absence of a significant effect
of size difference on encoding time (middle-word gaze duration) for the 50 percent of participants with the lowest middle-last ratios \( (M = .21), \ F_1(1,11) = 2.652, p = .132; F_2(1,41) = .839, p = .365 (n = 12). \) The 50 percent of participants with the highest middle-last ratios \( (M = .56) \) showed robust symbolic distance effects on encoding time, \( F_1(1,11) = 12.5, p = .005; F_2(1,41) = 10.6, p = .002, (n = 12). \)

This experiment showed robust symbolic distance effects on decision times across participants, and on encoding times for those participants using an incremental-judgment strategy. Because the symbolic distance effect reflects difficulty of comparing the size of objects that are close in size, these effects support the idea that size information associated with lexical items was activated by the goal of size comparison. The experiment provided no evidence of distance priming as would be expected if size information were automatically activated during lexical encoding.
Figure 3. Individuals’ mean middle-last ratio by the slope of the size-based effect on the middle word in Experiment 1. There is a strong, negative relationship between the individuals’ mean middle-last ratio and mean encoding slope. Individuals with a high middle-last ratio (incremental judgment strategy) show a strong symbolic difference effect during middle word gaze durations. Individuals with a low middle-last ratio (final judgment strategy) show no systematic effect of size difference on middle word gaze durations.
Results of Experiment 1 with 6 blocks: Mean accuracy was 88%. Decision times decreased as the difference in semantic size between the first and middle word increased: $F_1(1,23) = 36.69, p < .001; F_2(1,41) = 27.63, p < .001$ for the linear effect of size difference. There was no difference in decision times between ascending and descending trials: $F_1(1,23) = 1.54, p = .228; F_2(1,41) = .916, p = .344$ and no interaction between size difference and the direction of the difference: $F_1(1,23) = 2.62, p = .119; F_2(1,41) = 3.10, p = .086$. Decision times were faster for consistent trials (1853 ms) than for inconsistent trials (2050 ms), $F_1(1,23) = 40.71, p < .001; F_2(1,41) = 23.49, p < .001$ and there was a marginally significant tendency for the effect of size difference to be stronger for consistent trials: $F_1(1,23) = 5.66, p = .026; F_2(1,41) = 2.85, p = .099$.

Mean encoding times were shorter when the semantic size difference between the first and middle word was larger: $F_1(1,23) = 7.74, p = .011; F_2(1,41) = 21.25, p < .001$ for the linear effect of size difference. There was no difference between ascending and descending trials: $F_1(1,23) = .99, p = .330; F_2(1,41) = 1.45, p = .236$ and no interaction between size difference and direction of the difference: $F_1(1,23) = 0.15, p = .904; F_2(1,41) = .011, p = .918$.

Mean participant middle-last ratios ranged from .11 to 1.05. There was a negative correlation between the middle-last ratio and the slopes of their size-difference effects: $r = -.665, p < .001, R^2 = .443$. Middle-last ratio was not related to total trial completion time: $r = .106 p = .623$, but it was related to overall accuracy, with use of an incremental strategy correlating positively with accuracy: $r = .51, p = .011, R^2 = .260$. The 50 percent of participants with the lowest middle-last ratios (M = .25) showed no significant effect of size difference on encoding time: $F_1(1,11) = 1.17, p = .302; F_2(1,41) = .636, p = .430$ (n = 12). The 50 percent of participants with the highest middle-last ratios (M = .63) showed robust symbolic distance effects on encoding time, $F_1(1,11) = 8.72, p = .013; F_2(1,41) = 23.16, p < .001$, (n = 12).
Experiment 2

The results of Experiment 1 did not show a size-based priming effect of the sort that would be expected if activation of semantic size occurs during lexical encoding of object and animal names in the same way that Brysbaert (1995) showed that it did for numerals. However, the results did show comparison-based activation of size information both on decision time and on middle-word encoding time, with the analysis of judgment strategy suggesting that the absence of distance priming was not due to masking by the symbolic distance effect. The current experiment examines whether distance priming between successive object and animal names is found when the task demands do not involve the semantic size of the referents. It uses Brysbaert’s memory paradigm, where participants read three words as in Experiment 1 but then had to judge whether a subsequently presented probe word was among the preceding triplet. Brysbaert found numerical distance priming in this task even though the magnitude of the numbers was not relevant to performing the task. This provided evidence that the magnitude of numbers is activated automatically upon encoding even if magnitude is not relevant to task goals. Examination of whether size-based priming is found for object and animal names in this task provides a further test of whether there is encoding-based activation of semantic size that may perhaps have been masked by comparison processes in Experiment 1.
Method

Participants. Twenty-four undergraduates at the University of North Carolina at Chapel Hill participated for course credit. All were native English speakers with normal or corrected-to-normal vision and were naïve about the research goals. None had participated in Experiment 1.

Stimuli & Procedure. The object and animal names as well as the mode of presentation were identical to those used in Experiment 1. However, participants were instructed to move their gaze to a mask at the bottom-center of the screen after reading the third word in each triplet. Once fixated, the mask revealed a fourth word. On half of the trials, the fourth word had been among that trial’s triplet, while on the other half it was new (but could have appeared earlier in the experiment on one or more different trials). The participants’ task was to indicate whether the probe had been among the triplet or not, indicating ‘yes’ or ‘no’ on a hand-held console. The first, middle and last word of the triplet appeared as ‘old’ probes equally often. No words of the triplet were visible on the screen after the onset of the probe. The probe remained visible until a response was made. Participants received accuracy feedback after each trial.

After completing the experiment, participants were given a questionnaire with increasingly specific questions checking for any awareness of the size difference manipulation.

Analysis of eye movements. Eye movements were analyzed as in Experiment 1. Middle-last ratios were not computed since the triplet and the probe were not considered to take part in the same task process. Due to a stimulus error .9% of all trials were excluded.
Results and Discussion

Detailed results are presented for the first three blocks only as in Experiment 1. Overall accuracy across participants was 97%, with no participant scoring below 92% accuracy. Inaccurate trials were excluded from all analyses involving fixations on the last word or the probe and response time measures.

Average encoding times for the middle words did not vary significantly between triplets with a small (242 ms), medium (247 ms) or large (243 ms) size difference between the first and middle words, F(1,23) < 1; F(2,41) < 1. Encoding times also did not vary as a function of the direction of the difference (ascending: 245 ms or descending: 244 ms) between the first and middle words, F(1,23) < 1; F(2,41) < 1, nor was there a significant interaction between size difference and the direction of the difference, F(1,23) = 1.64, p = .213; F(2,41) = 2.67, p = .110.

The null effect of size difference on middle-word encoding time was found for a measure that was used by Brysbaert (1995) to show distance priming with numbers and which has been widely used as a measure of lexical access in studies of reading (Rayner, 1998). In order to show that reading time on the middle word was a sensitive measure of lexical encoding in this experiment, we investigated its relation to word length (number of letters) and word frequency (SUBTLEXus, Brysbaert & New, 2009), factors that are strongly related to ease of lexical encoding (Rayner, 1998). Multiple-regression analyses showed that on the middle-word, single-fixation duration was strongly related to word length and frequency, r = .729, R^2 = .532, F(2,39) = 22.15, p < .001 (see Figure 4); word length significantly predicted single fixation duration, β = .42, p < .01, as did the log of word
frequency within the 51 million word SUBTLEXus corpus, $b = -.396$, $p < .01$. Neither word length nor word frequency interacted with size difference.

In sum, this experiment showed that while reading time on the middle word was related to ease of lexical encoding, it provided no evidence of distance priming as would be expected if there were encoding-based activation of semantic size for object and animal names, as has been found for numerals (Brysbaert, 1995).
Figure 4. Multiple regression of single fixations duration in Experiment 2, using word frequency (SUBTLEXus, Brysbaert & New, 2009) and word length as predictors. Both word frequency and length are known as robust predictors of encoding time of written words (Rayner, 1998).
Results of Experiment 2 with all 6 blocks: Accuracy across participants was 97% with no participant scoring below 92% accuracy. Encoding times for the middle words did not vary significantly between triplets with a small (244 ms), medium (247 ms) and large (245 ms) size difference between the first and middle word. $F_1(1,23) = .336, p = .568; F_2(1,41) = .214, p = .646$. Encoding times also did not vary as a function of the direction of the difference (ascending: 246 ms or descending: 245 ms) between the first and middle words, $F_1(1,23) = .261, p = .614; F_2(1,41) = .96, p = .759$, nor was there a significant interaction between size difference and the direction of the difference, $F_1(1,23) = .015, p = .904; F_2(1,41) = .006, p = .938$. Multiple regression showed that middle-word, single-fixation duration was strongly related to word length and frequency, $r = .771, R^2 = .60, F(2,39) = 28.64, p < .001$; word length significantly predicted single fixation duration, $\beta = .535, p < .001$, as did the log of word frequency per million, $\beta = -.326, p = .013$. Neither word length nor word frequency interacted with size difference.
Experiment 3

Neither Experiment 1 nor Experiment 2 provided any evidence of distance priming in the processing of animal and object names, though Experiment 1 showed very robust symbolic distance effects for these words. This pattern supports the conclusion that representations of semantic size are not automatically activated as part of encoding these items but that activation follows instead from the processes required to achieve task goals. In contrast, research on numerals has consistently shown distance priming (Brysbaert, 1995; Den Heyer & Briand, 1986; Marcel & Forrin, 1974; Reynvoet et al., 2002) and symbolic distance effects (Brysbaert, 1995; Henik & Tzelgov, 1982; Koechlin et al., 1999; Moyer & Landauer, 1967; Van Opstal et al., 2008), a pattern indicating that representations of semantic size for numerals are automatically activated during encoding and that the use of those representations is influenced by task goals. In Experiment 3 the size-order judgment task of Experiment 1 (Brysbaert, 1995, Experiment 1) is used with both word and number trials in order to determine whether this difference in how semantic size is activated for words and numerals can be demonstrated in the same experiment.

Method

Participants. Twenty-four undergraduates at the University of North Carolina at Chapel Hill participated for course credit. All were native English speakers with normal or corrected-to-normal vision and were naïve about the research goals. None had participated in Experiment 1 or 2.
Stimuli. The animal and object triplets were identical to those in Experiment 1. However, to allow for the addition of the number block while keeping the testing session under forty-five minutes each participant was presented with each target three times, once in each size condition, rather than six times as in Experiment 1. Thus, a participant saw each item-by-size combination either in an ascending or a descending triplet, but not in both. This resulted in a total of 126 animal and object trials per participant. This is equivalent to the first 3 blocks of data which were reported for Experiment 1. The trials were presented in random order.

The number triplets were constructed by having each number between 14 and 97 presented as a target (middle) number once, resulting in 76 trials per participant. The first and last numbers were selected so that the numerical difference between the target and the flankers was between 1 and 19. The direction of the difference between the first and middle number was either ascending or descending, and the third number was chosen so that the triplet was either consistent (e.g. 24 – 26 – 31) or inconsistent (e.g. 28 – 26 – 31), counterbalanced across four lists. Each list contained each target number once, with consistent/inconsistent and ascending/descending trials appearing equally often. Similarly, each numerical difference between 1 and 19 appeared approximately equally often per list. Each subject saw one list, preceded by eight warm-up trials which were excluded from all analyses. The number trials were presented in a separate block from the word trials and the order of the word- and number block was counterbalanced across participants.

Procedure. The stimulus presentation and eye-tracking procedures were the same as in Experiment 1 with one exception. Whereas in Experiment 1 the first word of the triplet appeared in place of the fixation point, in Experiment 3 the fixation point and the three words
were each placed on a separate vertical quarter of the screen. This allowed for more reliable measurement of the reading time on the first word, as the appearance of all three words in the triplet was triggered in the same way. A gaze-contingent invisible boundary was placed to the right of the fixation point and 80 pixels from the left boundary of the first word. Once the fixation point was fixated and the trial started, all three words appeared to the right of the fixation point masked by hash marks, with 240 pixels between the left boundaries of the masks. Gaze-contingent invisible boundaries were placed 80 pixels from the left boundary of each mask in the triplet, preventing preview and rereading of any of the words before entering and after exiting the word’s region.

As in Experiment 1, participants were asked to indicate whether the size of the middle animal or object was in between the size of the two outer items. For the number trials, the task was to indicate whether the middle number was in between the two outer numbers. All responses were made by pressing ‘yes’ or ‘no’ on a hand-held console.

There was a short break between the number and word blocks, and the eye tracker was recalibrated for all participants between the two blocks.

*Analysis of eye movements.* Eye movements were analyzed as in Experiment 1.

**Results**

Two participants whose accuracy score was below 57% on either portion of the experiment were excluded from both sets of analyses and replaced. Comparing accuracy and total trial time for the number and object-animal portion of the experiment indicates that the task was easier to perform with numbers than with object and animal names. Overall accuracy was higher for the number (93%) than for the word portion of the experiment (89%): $F(1,23) = 12.04, p = .002$, and mean total trial time was faster for the numbers ($M = \ldots$)
2348 ms) than the object-animal names (M = 3590 ms): F(1,23) = 58.0, p < .001. For both the word and number tasks, inaccurate trials were excluded from all reading time analyses involving response time measures. Reading time results for the word and number blocks are discussed separately below.

Object and animal names. Mean decision times are shown in the top, left panel of Figure 5. As in Experiment 1, they showed a strong symbolic distance effect, with times decreasing as the difference in semantic size between the first and middle word increased, $F_1(1,23) = 31.10, p < .001; F_2(1,41) = 25.52, p < .001$. There was no significant difference in decision times between ascending and descending trials, $F_1(1,23) = 3.87, p = .061; F_2(1,41) = 3.40, p = .072$, though there was a tendency for the effect of size difference to be stronger for descending than ascending trials: $F_1(1,23) = 8.106, p = .009; F_2(1,41) = 5.00, p = .031$. Decision times were faster for consistent than inconsistent trials, $F_1(1,23) = 9.91, p = .005; F_2(1,41) = 14.07, p = .001$, but the interaction between consistency and size difference was not significant, $F_1(1,23) = 1.127, p = .299; F_2(1,41) = 0.850, p = .362$.

As shown in the top-right panel of Figure 6, encoding time on the middle word also showed a symbolic distance effect, such that mean gaze durations on the middle word decreased as the difference between the middle and first word increased, $F_1(1,23) = 9.51, p = .005; F_2(1,41) = 7.39, p = .01$. Middle-word encoding time did not differ significantly for ascending and descending trials, $F_1(1,23) = .37, p = .850; F_2(1,41) = .107, p = .745$, nor was there an interaction between size difference and direction of the difference, $F_1(1,23) = .063; p = .804, F_2(1,41) = .352, p = .556$.

These results for the word portion of Experiment 3 replicate Experiment 1 on all measures of interest.
Variation in individual strategies for animal and object names. Analysis of variation in participants’ task strategies was assessed as in Experiment 1 by examining the relationship between middle-last ratio and symbolic distance effect slopes on encoding time. Individuals’ mean middle-last ratios ranged from .12 to .93. Figure 5 shows that there was a strong negative correlation between participants’ middle-last ratio and the slopes of their size-distance effects, $r = -.654, p = .001, R^2 = .43$, just as there was in Experiment 1. Individuals with higher ratios showed more negative encoding time slopes, indicating a stronger effect of size difference. Consistent with Experiment 1, judgment strategy, as measured by middle-last ratio, was not related to overall speed as measured by total time for trial completion, $r = -.108, p = .616$. In contrast to Experiment 1, where higher middle-last ratios were significantly related to overall accuracy, this relationship was not significant in Experiment 2, $r = .325, p = .122$. However, the results of one participant were very different from the general pattern, as this person had both the lowest decision ratio (.12) and the highest accuracy rate (1.0). When this outlier was removed there was a significant correlation between middle-last ratio and accuracy that was numerically similar to the one found in Experiment 1, $r = .490, p = .018, R^2 = .240$. Removal of this participant’s data caused little change in the critical relationship between encoding time slope and middle-last ratio, $r = -.642, p = .001, R^2 = .412$.

As in Experiment 1, the 50% of participants with the lowest middle-last ratios (mean = .27) did not show a significant effect of size difference on encoding time, $F_1(1,11) = .129, p = .726; F_2(1,41) = .034, p = .855$, while the 50% of participants with the highest middle-last ratios (M = .63) showed a significant effect on this measure, $F_1(1,11) = 14.22, p = .003$;
$F_{2}(1,41) = 7.148 \ p = .011$. These results on strategy effects for the word portion of Experiment 3 replicate those from Experiment 1 on all measures of interest.
Figure 5. Individuals’ mean middle-last ratio by the slope of the size-based effect on the middle word in Experiment 3. These results show the same pattern as Experiment 1. There is a strong, negative relationship between the individuals’ mean middle-last ratio and mean encoding slope. Individuals with a high middle-last ratio (incremental judgment strategy) show a strong symbolic difference effect during middle word gaze durations. Individuals with a low middle-last ratio (final judgment strategy) show no systematic effect of size difference on middle word gaze durations.
Numbers. Whereas the design of the object-animal names study used three categories of size difference (small, medium, large), the difference in size for the numbers varied continuously from one to nineteen. Accordingly, the effects of size difference were analyzed using regression rather than ANOVA, with times regressed on the log of size difference because of evidence that the mental number line is logarithmic (Shepard, Kilpatric & Cunningham, 1975; Dehaene, 2003). Trials on which the first and last number were identical (6.2% of the trials) were excluded from all analyses involving decision time and accuracy measures as were trials where decision times were more than three standard deviations above the grand mean (2.4% of correct trials).

As seen in the bottom-left panel of Figure 6, average decision times decreased with the log of the difference between the middle and final numerals, $r = .880, R^2 = .774$, $F(1,17) = 58.13, p < .001$. The generality across participants of this symbolic distance effect on decision times was further assessed by fitting the regression model to each participant’s data individually. The slope parameters calculated this way were significantly less than zero ($M = -51.1), t(23) = -3.1, p = .005$. Decision times were faster for consistent triplets (1051 ms) than inconsistent triplets (1196 ms), $F(1,23) = 20.46, p < .001$ and the effect of size difference on decision time was greater for consistent sequences than for inconsistent sequences, $t(23) = 4.6, p < .001$.

As seen in the bottom-right panel of Figure 6, average encoding times for the middle number increased with the log of the difference between the first and middle numerals, $r = .63, R^2 = .40, F(1,17) = 11.35, p = .004$. The generality across participants of this distance priming effect on encoding times was significant in a test of the slope parameters obtained from fitting the regression model to participants data individually, ($M = 39.3), t(23) = 2.44, p
Encoding times were slightly faster for ascending (565 ms) than descending (605 ms), $F(1,23) = 5.5, p < .05$, but the direction of the difference did not interact with size difference: $t(23) = -1.95, p = .064$.

These results support both encoding- and goal-based activation of numerical magnitude. At the number-meaning level of comprehension, encoding-based activation results in a numerical distance priming effect on encoding times. At the task-level of comprehension, goal-based activation of numerical magnitude results in a symbolic distance effect on the comparison decision times.
Figure 6. Encoding and decision times for words and numbers in Experiment 3. The decision times (A) and middle word encoding times (B) for the words both show a symbolic distance effect. Decision as well as encoding times are shorter when the difference in semantic size between the middle and first item is larger. The number decision times (C) follow this pattern as well. Decision time decreases as the log difference between the last and middle number increases. The encoding times of the middle number (D) show the opposite pattern in the form of a numerical distance priming effect. Middle number encoding times increase as the numerical difference between the middle and first number increases.
Variation in individual strategies for numbers. Individual participant strategies were assessed the same as in Experiment 1 and the word portion of Experiment 3, using middle-last ratios computed by taking the encoding time on the middle word over the decision time. Participants’ mean middle-last ratios on the number portion ranged from .20 to 1.10. In addition, the strength of the numerical distance priming effect for each individual was calculated by taking the slope of encoding time as a function of the numerical distance between the middle and first number, so that more positive slopes indicated stronger priming effects. Whereas individual strategies were related to the strength of the symbolic distance effect on encoding times for the object and animal names, individual participant strategies did not uniformly affect the strength of the numerical distance priming effect as there was no relationship between the subject’s mean middle-last ratio and their encoding time slope: $r = .264, p = .213$.

Judgment strategy, as measured by middle-last ratio, was related to overall speed, with use of an incremental judgment strategy showing a positive correlation with average time needed to complete each trial: $r = .602, p = .002, R^2 = .362$. This is in contrast with results on object and animal names in Experiment 1 and the word portion of Experiment 3, where decision ratio was not related to overall speed. This suggests that individuals with an incremental judgment strategy for the numbers generally processed numerical stimuli more slowly than individuals with a final judgment strategy.

Middle-last ratio was not related to accuracy: $r = .228, p = .283$. Again, this result is in contrast with Experiment 1 and the word portion of Experiment 3 once we removed one participant outlier, in which higher middle-last ratios were related to better task performance.
Discussion

The results of Experiment 3 support our earlier results showing that representations of the size of objects and animals are not automatically activated during lexical encoding of their names, but rather that size information is activated for the comparison processes necessary to achieve task goals. Further, the results show that this pattern contrasts with that found for numerals, where representations of semantic size are automatically activated during encoding and where the use of those representations is influenced by task goals. These conclusions are supported by the finding that object and animal names do not show a distance priming effect but do show a symbolic distance effect, while numerals show distance priming early in processing and a symbolic distance effect later in processing as previously demonstrated by Brysbaert (1995).

Analyses of individual differences in task strategies provide further evidence that semantic size is activated in different ways for object-animal names and for numerals. The two types of stimuli lead to similar variation in middle-last ratio, a measure of the time spent on the middle word or number relative to the time spent on making the decision after seeing the final word or number. However, the characteristics of performance associated with this variation in task strategies differ for the two types of stimuli. For object and animal names, a high middle-last ratio was associated with showing a symbolic distance effect on encoding time for the middle word, with higher accuracy and with no increase in overall trial time. This suggests that participants showing high middle-last ratios made use of an incremental-judgment strategy that involved comparing the semantic size of the middle and first items before encoding and comparing the final item. This strategy is efficient in that it increased accuracy without increasing overall time to complete the trial. For numerals, variation in
middle-last ratio was not related to the magnitude of distance priming during encoding of the middle number, nor was it related to performance accuracy. However, high middle-last ratios were associated with slower trial times.

Differences in the meanings associated with animal-object names and numerals suggest a reason why there are differences in the use of semantic size information for these two types of symbols. Many different types of semantic properties are associated with object and animal names. Automatically activating all of those meanings upon encoding a word may be impossible or inefficient. Accordingly, semantic size information is only activated by processes that are specifically related to task goals, in this case comparison of the size of successive items. Use of an incremental-judgment strategy is efficient because activation of the size information is difficult, making it problematic to engage in a final-judgment strategy where activation and comparison of size information is postponed until all three items have been encoded. In contrast, because little meaning beyond semantic size is inherently associated with numerals, semantic size information may be automatically activated upon encoding. Such easy, automatic activation eliminates any processing advantage of an incremental-judgment strategy relative to a final-judgment strategy where the activated size representations for the three items can be evaluated with respect to the task demand of determining whether the middle number is in between the outer ones. For this reason greater time spent looking at the middle item is not efficient and is associated with longer overall trial times.

Comparison of overall performance in the numeral and animal-object tasks provides further support for the notion that activation of size information is much easier for numerals than for animals and objects. As noted above, average completion times and error rates were
substantially lower for numeral stimuli than for animal and object stimuli (2348 ms vs. 3590 ms, and 7% vs. 11%). It seems that semantic size information rides for free with numerals but not with words referring to concrete objects.
General Discussion

The three experiments reported here examined the way in which semantic size is activated and used to make judgments. It did so by measuring looking times and decision times as participants made judgments about three symbols (words or numbers) displayed on a screen. A gaze-contingent display technique was used so that participants only received visual information about each of the symbols during first-pass fixations on the region of the screen containing the symbol. This technique tightens the association between looking time and encoding time by preventing parafoveal preview and rereading (Brysbaert, 1995).

Experiment 1 showed that decision times in a three-item, size-order judgment task increased as the distance in the semantic size of the referents of successive animal or object names decreased. This symbolic distance effect was also observed on encoding time for participants who adopted an incremental-judgment strategy of comparing the size of the first two items before proceeding to encoding of the third item. This pattern indicates that semantic size of the referent was activated in a goal-driven manner by the mental processes for comparing the size of the referents. Experiment 1 provided no evidence of distance priming (reduction of encoding time when the preceding referent was close in size); an effect that would have indicated that semantic size was activated automatically during encoding. Experiment 2 used a memory-probe task rather than an order-judgment task but also found no evidence of distance priming within the three-item sequences. If the task goal of size comparison had masked distance priming due to encoding-based activation in Experiment 1, then the absence
of any size-related task goal should have unmasked the distance-priming effect. Experiment 3 replicated the pattern seen in Experiment 1 for animal and object names, indicating goal-driven activation of size representations but no automatic activation of size representations during lexical encoding. In addition, it showed that size representations were activated both in a goal-driven manner and automatically during encoding when the items in the sequence were numerals rather than words referring to concrete objects. This contrast between words and numerals suggests that size information is automatically activated during encoding when size is the dominant if not sole meaning of a symbol but that it is not automatically activated upon encoding of symbols, such as words, that have many associated meanings.

The pattern of results in these experiments is not consistent with predictions derived from accounts of language comprehension as an embodied process that is fundamentally based on perceptual simulation. Models of this sort, such as perceptual symbol systems (Barsalou, 1999) or the immersed experiencer framework (Zwaan, 2004), propose that language comprehension entails reactivation of the information associated with the perceptual experience of the word’s referent. If reactivation of perceptual experience is a necessary part of language comprehension, then distance-priming should have been observed for the encoding of words as it was for the encoding of numbers. However, our results suggest that words are encoded without such perceptual reactivation and that semantic size representations are activated in a goal-driven manner by processes necessary for successful task performance.

Of course it is possible to counter that encoding a word leads to reactivation of some of the perceptual properties of its referent but that priming was not observed for semantic size because it was not a sufficiently important or salient property of the object or animal
names. Priming might have been observed for other perceptual dimensions, such as color or texture, that are more salient than size in general or are particularly salient for a subset of the animals and objects. While this line of reasoning cannot be discounted completely, we believe that it is not compelling. For the objects and animals that we studied, semantic size is a very stable property of meaning, as demonstrated by the consistency of participants’ size ratings across studies (Dean et al., 2005; Holyoak et al., 1979). Further, abundant demonstrations of the symbolic distance effect (Banks & Flora, 1977; Dean et al., 2005; Holyoak, et al., 1979; Moyer & Bayer, 1976) show that size comparisons for these words are made using the same kind of analog representations that are used for numbers (Henik & Tzelgov, 1982; Koechlin et al., 1999; Moyer & Landauer, 1967; Van Opstal et al., 2008) and for actual perceptual stimuli (Johnson, 1939; Moyer & Bayer, 1976). The symbolic distance effect has long been used to demonstrate an “internal psychophysics” (Moyer, 1973 p.183) where the relative size of the referents is processed using representations that are akin to perceptual representations. Accordingly, the characteristics of semantic size make it an excellent property in which to look for reactivation of perceptual experience as a result of encoding words.

It is also possible to counter that people create simulations rich enough to extract the relevant aspects of the meaning of words without expending more cognitive effort than strictly necessary (Gibbs & Perlman, 2010), and that therefore activation of perceptual properties is selective. This view readily accounts for the absence of distance priming in Experiment 2, where the memory task made semantic size irrelevant, though it is less clear how it would account for the results of Experiments 1 and 3, which explicitly required use of semantic size. An optimal simulation of the word meanings for this task certainly involves
size information, yet there is no evidence of such a simulation during encoding. This suggests that the activation of analog, modal representations is not what constitutes language comprehension, or at least not exclusively. Moreover, the idea that perceptual information is an optional, goal-directed aspect of language meaning makes it difficult to contrast the embodied approach to language comprehension with other approaches in a way that is problematic for either.

Our goal-driven activation model proposes that, during encoding, comprehension may rely on activation of abstract, propositional or perceptually underspecified representations of word meaning, which in turn allow for activation of context- or task-relevant aspects of the word’s referent. As such, the goal-driven activation model can account for findings of perceptual information affecting language comprehension processes in a ‘disembodied’ system, by assuming flow of activation between perceptual and non-perceptual systems according to goals associated with comprehension and task performance.

**Sensorimotor Representations and Action-Oriented Language Comprehension**

Notwithstanding these conclusions, the growing body of research on embodied language processing has provided ample evidence that sensorimotor processes play a role during language comprehension. Although the current study adds additional weight to the argument that these findings to not support the conclusion that sensorimotor simulation is constitutive of comprehension, this type of information undeniably supports and enhances our understanding of linguistic input, specifically within the framework of task-oriented behavior. Therefore, any comprehensive model of language processing needs to incorporate a role for analogue, modality specific representations. As discussed in the introduction, sensorimotor representations have traditionally received ample attention in the study of
processes related to (language-based) reasoning and problem solving. In addition, several
models of comprehension explicitly include a role for both abstract and sensorimotor
representations of meaning, ranging from Paivio’s dual-coding theory (1986) to more recent
developments such as Mahon & Caramazza’s (2008) ‘grounding by interaction’ approach,
Dove’s (2009) representational pluralism, and the goal-driven activation model presented in
this paper. In all of these, sensorimotor representations are considered crucial to building
rich, contextually relevant representations of linguistic meaning that also incorporate the
reader’s background knowledge and task-goals, recognizing the importance of fast and
abundant cross-talk between the conceptual and modal system. Unlike the embodied account,
however, these models do not altogether reject the idea of an abstract, symbolic system
performing a key role in the representation of meaning.

Regardless of the important role mental simulation of perceptual information may
play during every day language use, the lack of an automatic connection between words as
abstract symbols and sensorimotor representations may be exactly what allows our linguistic
system to be such a powerful representational tool. As noted earlier, words referring to
congrete objects are part of a rich semantic space, which can include category information,
semantic associates and linguistic co-occurrence information as well as a multifaceted
perceptual representation. As such these models much more readily account for the current
findings, as well as several other key observations in the study of language comprehension,
including our ability to flexibly access different aspects of word’s meaning depending on the
(task) context (Barclay et al., 1974) and our ability to represent abstract concepts.

Weiskopf (2010) argues that the different interpretations of the role perceptual
information plays or should play during language comprehension as proposed by embodied
and disembodied accounts of language comprehension is based in a fundamental
disagreement about what language comprehension essentially is, and what it is for. On the
traditional assumption that cognition is first and foremost a symbol processing system,
integration with perceptual processing is by no means essential to its operation.
Understanding of the propositional content or truth conditions of an utterance can be
achieved through manipulation of abstract symbols, which are divorced from bodily
experience. Embodied accounts of cognition, however, are based on the idea that the function
of cognition is to guide actions in the environment in real time (Clark, 1997, Wilson, 2002).
To allow for efficient interaction with our environment, perceptual information is essential to
cognitive processing: the more closely integrated perceptual and motor processes are with
other cognitive processes, including language, the more efficiently a cognitive organism will
operate in the real world.

For example, Glenberg and Robertson (2000) argue that non-afforded sentence such
as ‘Adam used a ham sandwich to chisel the ice of his windshield’ are difficult to understand
because it is difficult to represent the described situation based on the affordances of the
objects in the sentences. Indeed, they found that these sentences are rated low in sensibility
and more difficult to envision. However, Weiskopf (2010) points out that such difficulty to
construct a simulated perceptual experience does not inhibit linguistic understanding in a
more minimal sense. We are still able to assess the truth conditions or propositional content
of the sentence: Adam used something to chisel ice of his windshield, and the object he used
was a ham sandwich. Thus, it is possible to understand the syntactic and semantic features of
the sentence without being able to readily simulate the situation. However, in order to apply
cognitive processes towards successful operation in the environment, the system needs the
type of interface provided by perceptual and motor processing. For example, in order to assess whether chiseling ice off your windshield with ham sandwich is plausible, or to make inferences about the described situation (e.g. ‘maybe the ham sandwich was frozen as well’), a sensorimotor simulation based on (perceptual) background knowledge about the object’s affordances will be an effective cognitive tool. Therefore, ability to access perceptual properties, or whether something is easy or difficult to simulate perceptually, does not affect our understanding of the truth conditions or propositional content of an utterance (Weiskopf, 2010). However, other comprehension goals, which may be as simple as an implicit plausibility check, often rely on the goal-driven activation of perceptual properties for efficient task performance. Similarly, a person with aphasia may be able to name a picture of a hammer, but unable to demonstrate how to use it. Depending on one’s adopted theory of what language is for, it may be argued that this person does or does not ‘understand’ the meaning of HAMMER (Caramazza & Mahon, 2006). In other words, the importance of sensorimotor representations during comprehension of language depends on how much importance is granted to the ability to perform actions in the environment.

Future directions

Instead of presenting the current debate on the nature of language comprehension as a choice between two representational systems, one abstract and one perceptual, focusing on the functional role of either aspect of meaning may prove to be a more fruitful approach. Acknowledging that symbolic and perceptual representations of meaning need not be mutually exclusive, such an approach would be able to further explore the ways in which abstract and sensorimotor information contribute to comprehension in a complementary way. According to models integrating both types of representation, both symbolic and perceptual
components of a word’s complex meaning space can be activated flexibly, based on the reader’s specific context or task-related goals. While comprehension in the current study explicitly required the assessment of a spatial property, the goals that guide comprehension during everyday language use are usually implicit. The demands of comprehension or conversation may imply some of these goals, while others may be aimed at guiding action in the environment in real time, as discussed by proponents of embodied cognition (Wilson, 2002). Future studies may manipulate reader goals by varying the explicit and implicit demands of the task, or by presenting words in a sentential context that makes the properties in a specific modality more or less salient. Such research will allow will for further distinction between those comprehension processes that can be performed based exclusively on the abstract information that is retrieved during of encoding-based activation, and those task-related processes that require goal-based activation of the sensorimotor aspects of word meaning. Although there is no evidence that motor activation is constitutive of concepts, there may be some cognitive operations that are extremely difficult to perform without recruiting at least some sensorimotor processes, mental comparison of referent size being just one example, while others depend mainly on abstract aspects of meaning.

Close integration of perceptual and motor processes with other cognitive activity, including language processing, increases the efficiency of action in the environment, suggesting a substantial flow of goal-directed activation between linguistic, perceptual and motor processes. A goal-based approach to characterizing the nature of linguistic processing will provide a way of reconciling embodied and disembodied accounts of comprehension, opening the door towards a detailed understanding of how the comprehension system recruits
both abstract and modal representations of word meaning in order to optimize cognitive performance.
References


