Psychonomic Bulletin & Review
1999, 6 (3), 511–517

Notes and Comment

Strength of context *does* modulate the subordinate bias effect: A reply to Binder and Rayner

GEORGE KELLAS and HOANG VU
University of Kansas, Lawrence, Kansas

Using a self-paced reading task, Kellas, Martin, Yehling, Herman, and Vu (1995) demonstrated that strength of context can modulate the effects of meaning frequency. Binder and Rayner (1998) initially replicated the results, using eye-tracking methodology. On further examination of the stimuli, Binder and Rayner eliminated 43% of the stimulus set and found that context strength failed to modulate meaning frequency. Binder and Rayner's initial replication of Kellas et al. and the convergence of results between their two main experiments established the validity of self-paced reading as a measure of on-line reading, when compared with eye-tracking methodology. However, their central conclusion, that context strength cannot modulate the subordinate bias effect, is open to question. In this commentary, we examine the criteria adopted to exclude items from our homonym set and discuss the issue of local versus published norms. We also discuss the issue of context strength, as related to the specific rating procedures employed. Finally, we conclude that strong context can, in fact, eliminate the subordinate bias effect and that the context-sensitive model can more fully account for the available data on lexical ambiguity resolution.

Research on lexical ambiguity resolution has demonstrated that the activation of word meanings can be markedly influenced by the variables of meaning frequency (cf. Hogaboam & Perfetti, 1975) and biasing context (e.g., Paul, Kellas, Martin, & Clark, 1992; Tabossi, Colombo, & Job, 1987; Van Petten & Kutas, 1987). This has led to an investigation of the relationship between meaning frequency and contextual bias (e.g., Duffy, Morris, & Rayner, 1988; Rayner, Pacht, & Duffy, 1994; Simpson, 1981; Simpson & Krueger, 1991; Vu, Kellas, & Paul, 1998) and to the proposal of the reordered access and the context-sensitive models of lexical ambiguity resolution.

Introduced by Duffy et al. (1988), the reordered access model assumes an exhaustive retrieval of the meanings of an ambiguous word (e.g., a homonym) in all contexts. The meanings are accessed in the order of frequency with which each sense is given as an associative response in norming studies. Biasing context can reorder the availability of the meanings by boosting activation of the contextually appropriate sense, but it cannot preclude the inappropriate meaning(s) from being accessed. Results from eye-tracking measures have demonstrated two consistent findings. First, when biasing context precedes a balanced homonym (equal frequency of dominant and subordinate meanings), or when the preceding context biases the dominant sense of a polarized homonym (the most frequent meaning), reading time on the ambiguous word is comparable with that on an unambiguous control word. Although all the meanings are initially available, the appropriate meaning is quickly integrated with the context, because there is no competition with the less activated inappropriate meanings. Second, when the preceding context is biased toward the subordinate meaning of a polarized homonym, reading time is longer, relative to the control word. This outcome is referred to as the subordinate bias effect (SBE) and is assumed to reflect a time-consuming competition for text integration between two equally available meanings (the dominant sense, because of meaning frequency, and the subordinate sense, because of context).

The context-sensitive model recognizes the importance of meaning frequency and biasing context, but it is the parameter of context strength that will determine the pattern of meaning activation (cf. Vu et al., 1998). According to the context-sensitive position, a balanced homonym preceded by biasing context or a polarized homonym preceded by a dominant-biased context will lead to the activation of only the dominant meaning of the ambiguous word. However, when a polarized homonym is preceded by a subordinate-biased context, there are two possible outcomes. The SBE may emerge, or only the subordinate sense of the homonym will be activated, *contingent on the strength of the subordinate bias context.*

In this model, the so-called access and integration processes are simultaneously achieved with the incremental processing of words in text. The computations of interest begin at the point at which an ambiguity is encountered. If the context is strongly biased, multiple constraints (e.g., syntax, semantics, pragmatics) will converge and result in rapid activation of the contextually appropriate meaning. This meaning will be quickly integrated with the context as part of the computational process. However, if the context is weak, more time will be required to activate the contextually appropriate meaning. The additional time will enable computation of the contextually inappropriate meaning, because of the initial influence of meaning frequency information, and lead to a competition-for-activation between the two alternative senses. Accordingly, competitive activation culminates in slower integration

We thank Greg Simpson for his help on an earlier draft of this manuscript. Correspondence concerning this article should be addressed to G. Kellas, Department of Psychology, University of Kansas, Lawrence, KS 66045 (e-mail: gkellas@lark.cc.ukans.edu).

processes, and the delay is manifested in a longer reading time on the homonym. In sum, the emergence of the SBE in weak contexts and its elimination in strong contexts are predicted by the context-sensitive model, since these outcomes represent particular points on the continuum of contextual strength.

Another difference in the research testing the models is how meaning frequency is examined. Research investigating the reordered access model has typically treated meaning frequency as a dichotomous variable (balanced vs. polarized) and simply predicts an SBE for contexts biased toward the subordinate sense of a polarized homonym. In contrast, the context-sensitive model views meaning frequency as a continuous variable and predicts a graded outcome for the SBE. Critically, the model predicts a monotonic relationship between meaning frequency and the magnitude of the SBE. As the degree of polarity increases, the longer it takes to process the homonym in a weakly biased subordinate context, because more computational cycles are required to increase the activation level of the subordinate meaning to threshold. The result is that differential patterns of multiple meanings are activated for homonyms with different meaning frequencies, in which case SBEs of differing magnitudes become evident.

Using self-paced reading as an analog to eye-tracking measures, Kellas et al. (1995) examined the relationship between meaning frequency and strength of context in two experiments. The research tested two general predictions. First, not only was the SBE expected to emerge for a polarized homonym preceded by a weakly biased subordinate context, it was predicted that there would be a graded effect of meaning frequency. That is, the more polarized the homonym, the larger the magnitude of the SBE. The second prediction was that the SBE would be eliminated in strongly biased subordinate context, regardless of the polarity of the homonym. In order to examine meaning frequency as a continuous variable and to test the prediction of graded effects, Kellas et al. deliberately selected a wide range of meaning frequencies. The expected systematic increase in the magnitude of the SBE with increasing frequency of the dominant meaning was observed in both Experiments 1 and 2 [$r = +.64$ and $r = +.74$, $p < .05$]. In contrast, when the subordinate context was strongly biased, the SBE was eliminated, with no relationship being observed between meaning frequency and the SBE [$r = +.19$ and $r = -.10$, n.s., for Experiments 1 and 2, respectively]. The results demonstrate a graded effect of meaning frequency and underscore that a simple dichotomy between balanced and polarized ambiguous words is not fully informative of the relationship between the SBE and meaning frequency.

Furthermore, Kellas et al. (1995, Experiment 2) employed a naming task concurrent with self-paced reading and found that naming latencies for probe words supported what had essentially been inferred from reading times collected through self-paced and eye-tracking measures. That is, it has always been assumed that the presence of an SBE reflects the activation of multiple meanings and its absence implies selective activation. Self-paced reading and eye movement measures cannot provide a more direct test of these assumptions, but the naming task can. Indeed, Kellas et al. found that naming latencies for probe words related to both the dominant and the subordinate meanings of a polarized homonym were facilitated when the subordinate context was weakly biased (i.e., the presence of an SBE). However, when the context was strongly biased, only the contextually appropriate probe word was facilitated (i.e., the absence of an SBE).

Binder and Rayner (1998) used our stimuli and replicated our research, using eye-tracking measures. The purpose was to investigate the possibility that differences in reading methodologies may have contributed to our outcome. Their results replicated Kellas et al. (1995), thus providing evidence that differences in reading measures did not account for our findings. In fact, we now have converging evidence from three paradigms (self-paced reading, eye tracking, and naming) that the SBE and the activation of multiple meanings occur only when the subordinate context is weakly biased. However, on further examination of our stimuli, Binder and Rayner excluded 43% of the homonyms and conducted two additional experiments (one with eye tracking and the other with self-paced reading), using a modified and reduced stimulus set. They found that the SBE could not be eliminated by strong contexts. Rather, the results supported the reordered access model. Binder and Rayner concluded that context strength cannot modulate meaning frequency and that the Kellas et al. results were due to problematic stimuli.

In this commentary, we will first focus on the stimulus issue. We address Binder and Rayner's (1998) criteria for eliminating 43% of the homonym corpus. We subsequently discuss strength of context in detail, because we believe this to be the variable separating our research from that supporting the reordered access model.

STIMULI

Binder and Rayner (1998) reported three classes of stimuli that led to the exclusion of 24 homonyms from the original 56. These were (1) 3 homonyms in which the dominant- and subordinate-biasing contexts ostensibly biased the same meaning, (2) 7 homonyms in which the dominant and subordinate meanings were reversed, and (3) 14 homonyms that were balanced.

With respect to point (1), we assert that the dominant and subordinate contexts did *not* bias the same meaning of the three homonyms in question (*date*, *trade*, *field*). The contexts were constructed and, critically, empirically *rated* to be biased toward meanings that had been *operationally defined* and classified as *distinct* categories according to several published norms, including those of Nelson, McEvoy, Walling, and Wheeler (1980) and Twilley, Dixon, Taylor, and Clark (1994). Not only did ratings by subjects separate the contexts distinctly into dominant and subordinate categories, but our criteria for selecting homonyms were very stringent (see Martin, Vu, Kellas, & Metcalf, in press, for details). A point

of confusion may have occurred because of *one* of the example items reported. For the homonym *date*, we constructed contexts to bias the *month* and *girl* meanings of *date*, respectively. These contexts are shown below, with prepositional phrase continuances in parentheses:

The student made an error. He forgot the *date*... (on the final exam). (1a)

The boyfriend was thoughtless. He forgot the *date*... (*of* the birthday party). (1b)

It must be emphasized that the subjects rated our contexts *only up to the point of the ambiguous word*. For this item, the ratings separated nicely into the *month* and the *girl* meanings, *not* the same meaning. The same sense variations of the calendar meaning arise solely from the fact that the preposition *of* in 1b was inadvertently substituted for the correct preposition *at*. If one uses the contextual information *following* the ambiguous word, the same meaning is being biased. However, ratings were collected without the prepositional phrase continuances, and reading times were measured on the ambiguous word. Consequently, the inadvertent substitution of the incorrect preposition is inconsequential, because subjects never saw the continuances during stimulus rating or when reading the homonyms, since the task was self-paced and subsequent stimulus information was not present when the homonym occurred in the text. Biasing of the "same" meaning *did not* occur for the other two homonyms according to published norms and our local norms (discussed below).

As for the seven homonyms that were *reversed* in meanings, these items were *not* reversed and do meet the criteria for being polarized according to published norms. Because of an error in coding, there was, in fact, *one* item (*bar*) that did have its dominant and subordinate meanings unintentionally reversed. However, exclusion of this item from Kellas et al. (1995) had no effect on their reported results.

The elimination of the 14 *balanced* ambiguous words constituted the largest portion (58%) of the rejected stimuli and was motivated by frequency information obtained by Binder and Rayner's (1998) local norms *and* by those of Twilley et al. (1994). The criterion that was adopted to define the upper limit of balanced items was .70/.30 relative frequencies for the dominant and subordinate meanings, a criterion that has never been previously reported. However, the use of this criterion for excluding stimuli prevents a more rigorous examination of graded effects using meaning frequency as a continuous variable, as investigated by Kellas et al. (1995). Nor does it appear psychologically valid to define meaning frequencies of .65/.35 (and greater) as being balanced, in the absence of empirical data showing equivalent reading time effects across the range of frequencies excluded. On the other hand, Twilley et al. suggested that balanced ambiguous words should be those items that have a relative frequency difference of .10. Intuitively, this latter criterion makes sense to us.

For the 14 balanced items in question, the average meaning frequency for the set was .79 according to Nelson et al., (1980), .61 for Twilley et al. (1994), and .58 by the local norms of Binder and Rayner (1998). On the basis of the similarity between Twilley et al. and their local norms on these 14 items, Binder and Rayner rejected our stimuli. This comparison is misleading, however, resting as it does on the *average meaning frequency* of a small number of items, and does not provide information on individual homonyms. For example, *board* has a .91/.09, *fan* a .80/.20, and *bark* a .72/.28 split between dominant and subordinate meanings from the Twilley et al. norms but were still rejected by Binder and Rayner.[1] Second, and more important, we compared our local norms with those of Nelson et al. (our main source of items in the Kellas et al. [1995] study) to determine the representativeness of our local norms. We had a corpus of 150 homonyms, 100 of which overlapped with Nelson et al., and a correlation of +.87 was obtained between the two homonym corpora. The similarity of our local norms, as compared with those of Nelson et al., is like that reported between Twilley et al. and Nelson et al. The norms of Twilley et al. contained 100% of the items (320) examined by Nelson et al., and 84% of the variance was reported in common. Thus, we have strong evidence for the representativeness of our local norms. On the other hand, comparable information is not available for the local norms of Binder and Rayner.

The use of local norms is inherently problematic, because the task of collecting norms is notoriously inconsistent (Azuma, 1996). Subjects do not clearly provide one dominant or one subordinate response to a given homonym. Instead, several factors complicate the process of classifying responses into distinct categories. For example, ambiguous words may have two or more distinct meanings or share variations of the same sense, and/or their meanings may fall within the same or different lexical categories (e.g., *bank* has noun/noun associates, whereas *draw* has verb/noun associates). In addition, the criteria for the classification of meanings vary among investigators (Azuma, 1996). Published norms, although locally derived, have proven to be psychologically valid across laboratories. Published norms also include statistics showing the relationship to other available norms, number of subjects, stimuli used, and importantly, the criteria employed for classifying the associative responses into meaning categories. With regard to Binder and Rayner's (1998) local norms, how many meanings were given for a certain homonym? Were sense variations operationally defined and empirically categorized, or were they assumed to have the same meaning? Were classifications based on dictionary meanings, researchers' intuition, or independent judges? Were the most dominant and least subordinate meanings used in the experiment? If the associative meanings of the homonym were from different lexical categories, how were these classified? For example, the homonym *punch* can mean *a drink*, *a hit*, or *to hit*. These are but a few of the questions that could be raised if all experimental laboratories

were to use local norms. We believe that this would be a step backward in reading comprehension research that relies on the behavioral effects of stimulus characteristics, especially given the high level of correspondence among published norms that have proven to be psychologically valid across many laboratories.

We next turn our discussion to the stimuli that were used by Binder and Rayner (1998). Given that a large portion of our items were eliminated, the authors should have used the remaining items to conduct their research. This, however, was not the case. Through personal correspondence, we were able to obtain their list of included and excluded items and reconstructed the following.

Kellas et al. (1995) used 56 items and Binder and Rayner (1998) rejected 43% of these as being "balanced," "reversed meaning," or "same sense variations," citing a total of 24 excluded items. However, as it turns out, 2 of these items (*date* and *range*) were considered problematic by two different rejection criteria. So, only 22 unique items were excluded, with 34 remaining (20 in strong contexts and 14 in weak contexts). The excluded and remaining items are presented in Table 1.

The counterbalanced design required 16 stimuli in both strong and weak contexts. Notice that there are 20 stimuli remaining in the strong condition and 14 in the weak. The question is how were the stimuli selected to satisfy the design requirement of 16 stimuli per strength condition? The most straightforward way to implement the design would be to select 16 of the 20 remaining items for the strong context condition and to augment the 14 items in the weak context condition by adding 2 items from their local norms. Instead, for the strong contexts, 5 extra items were eliminated (*file, pen, present, shower, strike*), with the reinstatement of one item from the *excluded* list (*plot*). For the weak contexts, since there were only 14 items, 2 additional items were needed. The procedure taken was to eliminate 2 more items from this list (*set, switch*), thus leaving only 12 items. Subsequently, 4 items (*bar, screen, tip, trace*) from the *excluded* list were reinstated to obtain the 16 ambiguous words required for the design. It is not clear on what basis items were excluded and then reinstated, especially for those items that met original criteria according to their local norms. In conjunction with their criteria for item exclusion discussed earlier, especially for the balanced homonyms, we find their stimulus selection to be confusing.

Table 1
Items That Were Included/Excluded by Binder and Rayner (1998)

Originally excluded items	
Strong context	bark, board, bow, field, plot, pot, range, tag
Weak context	bar, court, date, draft, express, fan, marble, screen, speaker, stock, terminal, tip, trace, trade
Originally remaining items	
Strong context	bank, bulb, deck, dough, file, flights, lock, mole, note, pen, pitch, present, ring, shower, spare, spread, strike, toast, trunk, vessel
Weak context	bluff, break, coach, course, gin, hail, key, mine, object, perch, sage, scales, set, switch
Additionally excluded items	
Strong context	file, pen, present, shower, strike
Weak context	set, switch
Retracted items from originally excluded list	
Strong context	plot
Weak context	bar, screen, tip, trace

CONTEXT STRENGTH

Strength of context is a pivotal variable for differentiating between the reordered access and the context-sensitive models. In our view, when a polarized homonym is weakly biased by a subordinate context, the SBE will emerge and increase in magnitude as homonym polarity increases. When the subordinate context is strongly biased, the SBE will be eliminated. In our laboratory, we use contexts rated by subjects as being strongly biased and have consistently demonstrated selective activation of the subordinate meaning of polarized homonyms (e.g., Vu et al., 1998). On the other hand, Rayner and his colleagues have not empirically established the strength of contexts in their published research. In Kellas et al. (1995), we speculated that it was the use of insufficiently constraining contexts that produced the SBE in their laboratory, because meaning frequency in weak contexts will dominate the activation process. Binder and Rayner (1998) addressed the issue of context strength by conducting two norming studies that examined the strength of contexts employed by Duffy et al. (1988) and Kellas et al. and reported that the Duffy et al. contexts were rated as being *numerically* stronger (5.9) than the strong contexts used by Kellas et al. (5.1). However, the supporting evidence from the norming studies is open to question, inasmuch as the Binder and Rayner rating procedure differed markedly from that reported by Kellas et al. in a manner that may have influenced the outcomes. Let us first examine an example from Kellas et al. and Duffy et al:

Kellas et al.

The gardener dug a hole. She inserted a *bulb* ...

Duffy et al.

If you are concerned about having made an error, the *table* ...

Note that the Kellas et al. (1995) example is a short passage in which, in isolation, the second sentence, up to and including the ambiguous word (i.e., She inserted the *bulb*), had been rated by subjects as being ambiguous (cf. Vu et al., in press). However, by assigning the pronoun to its referent, the passage as a whole was rated as being strongly biased toward the subordinate meaning of *bulb* (*flower*). The situation conveyed in the first sentence is about a gardener digging a hole. When *she* is encountered, the pronoun reinstates the gardener, along with the associated information from the first sentence. This information, along with the transitive verb *inserted*,

constrains the interpretation of *bulb*. The only *bulb* that would be logically inserted, is a flower. Although plausible, it would be improbable or incoherent for subjects to invoke a *light* bulb.

Now consider the example from Duffy et al. (1988). Here, we have a fragment of a complex clause beginning with a subordinate clause followed by the main clause containing the ambiguous word. At the point at which the ambiguity is encountered, the main clause contains no constraining information relevant to the meanings of the homonym. Any biasing information occurs in the subordinate clause. Whereas there is sufficient information to resolve the homonym in the Kellas et al. (1995) stimuli, there is not in the Duffy et al. stimuli. Because the ambiguous word is the lead subject noun of the main clause, it is possible for the homonym to assume any of its alternative meanings. In the Duffy et al. example, although the subordinate meaning (e.g., table of *data*) is intended, the dominant *furniture* sense of *table* is not precluded at the point the ambiguity is encountered. Instead, the dominant sense is completely plausible for the sentence (e.g., If you are concerned about having made an error ... [the *table* in the corner has a calculator on it]).

Next, let us illustrate how the rating procedure employed by Binder and Rayner (1998) could have inflated the strength rating for the stimuli of Duffy et al. (1988). Binder and Rayner used a 7-point scale, ranging from *weakly* to *strongly biasing*, and the subjects were initially provided with an ambiguous word and *given* a definition of the intended meaning of the homonym for the upcoming context. In comparison, Kellas et al. (1995) simply gave subjects experimental passages, along with a 9-point scale anchored on each end by a word related to the dominant or subordinate meaning of the ambiguity. The subjects were instructed to rate the bias of each passage *according to their own judgments*. Importantly, no prior exposure of the ambiguous word occurred, nor was a definition of the intended meaning provided to the subjects at any time. These bias ratings were subsequently converted to a *strength* scale by calculating the deviation from center (5). In the converted scale, 0 represented *ambiguous contexts*, 1–2 *weak contexts*, and 3–4 *strong contexts* across the five scalar units.[2] We use the following examples, along with the different procedures, to illustrate our concern.

Binder and Rayner's Procedure

Table: an orderly display of data.

If you are concerned about having made an error, the *table*...

Weak 1 2 3 4 5 6 7 Strong

Kellas et al.'s Procedure

The gardener dug a hole. She inserted a *bulb*...

LIGHT 1 2 3 4 5 6 7 8 9 FLOWER

With the Kellas et al. (1995) procedure, the subjects are free to decide which meaning(s) are appropriate, since the scale has anchor points representing each meaning and an ambiguous center. This allows for an evaluation of the appropriateness of our contexts. For example, if a context was constructed to bias the subordinate meaning but the subjects regard the bias to be toward the dominant meaning or to be equally biasing, the item can be either eliminated or the context rewritten. The Binder and Rayner (1998) method does not provide for this safeguard. By explicitly providing a definition of the intended meaning, this method virtually *directs* the subjects to rate a context as biasing toward the subordinate meaning, even though the subjects might not normally have perceived it so (all Duffy et al. [1988] stimuli were intended to bias the subordinate meaning). As we have previously noted, the contexts of Duffy et al. do not clearly eliminate the dominant meaning. It is possible that subjects would rate Duffy's example above as being equally biased or biased toward the *furniture* sense of *table* if the alternative anchor point was provided. Without the prior presentation of the ambiguous word and its intended definition, the subjects most likely would have rated the context as weakly biasing toward the intended meaning.

Earlier in our research program, we had considered a procedure similar to that of Binder and Rayner (1998) but rejected it because it appeared as though the definition of the ambiguous word was combining with the experimental contexts to yield a relatively strong bias rating, even for passages constructed to be weakly biased. We suspect this may have been the case for the strength ratings reported above. Consequently, the overall strength ratings may be biased toward the strong end of the scale. Consider the above example of *table*. When the intended meaning is not known, it can be seen that the subordinate clause leading up to the homonym contains little information to signify that the appropriate meaning being biased is a *table of data*. However, when the intended meaning is explicitly provided, there is a framework for interpreting the ambiguous word and constructing a strongly biased context. This process is not surprising, given that it has been known for some time that related information is integrated during comprehension. As a result of the procedure used, there would be a reduced likelihood of showing strength differences between the contexts of Duffy et al. (1988) and Kellas et al. (1995). To make their point regarding the strength of bias of Duffy et al. and our stimuli, Binder and Rayner must employ the procedures used in the research they challenge. Any conclusion based on such markedly different measurement instruments is open to question.

The impact of strength of context on lexical ambiguity resolution reported by Kellas et al. (1995) is not an isolated finding. Earlier studies have found that when contexts are weakly biasing (e.g., Simpson, 1981) or ambiguous (Vu et al., 1998), multiple meanings of an ambiguous word are available. Also, these investigations have found selective activation of only the contextually appropriate meaning when the contexts are strongly biased. Importantly, the research of Simpson (1981) and

Simpson and Krueger (1991) demonstrated the modifying influence of context strength, using homonyms selected from Cramer (1970) and Perfetti, Lindsey, and Garson (1971), indicating that the results from Kellas et al. were not norm specific.

CONCLUSIONS

The research by Binder and Rayner (1998) is potentially important because it established the validity of self-paced reading as a method for evaluating on-line reading comprehension. However, their central conclusion, that context strength cannot modulate the SBE, is open to question because of the stimulus selection and strength rating procedures that we have detailed. In summary, Binder and Rayner eliminated a large portion of our items according to what they perceived to be similarities between their local norms and the normative stimuli of Twilley et al. (1994), but not ours. However, as we have pointed out, the average meaning frequency criterion is misleading, because only a small subset of items was compared between corpora. This method does not show, in any informative way, the representativeness of our full data set, nor does it reveal the correspondence of Binder and Rayner's local norms with Twilley et al., except for the handful of items based on average meaning frequency. We demonstrated a strong relationship between our local norms and those of Nelson et al. (1980), but there is no assurance that the local norms of Binder and Rayner are correlated with Nelson et al. or Twilley et al. The largest portion of our items that were eliminated were the 14 balanced homonyms. As Kellas et al. (1995) reported, however, the graded effects of meaning frequency were examined in two experiments, and the results conformed with predictions derived from a multiple constraint architecture. Clearly, a simple dichotomous classification of meaning frequency is not fully informative regarding the relationship between meaning frequency and subordinate-biased contexts.

Finally, although the reordered access proposal is not strictly a modular model (cf. Fodor, 1983; Forster, 1979), it does assume exhaustive activation of the meanings of an ambiguous word in all contexts (Duffy et al., 1988). This is critical in that it suggests that even when reading time on a homonym is equal to an unambiguous control word, the alternative meaning(s) are also available. However, neither eye-tracking nor self-paced reading measures are analytic for examining activation of specific word meanings; the availability of individual meanings can only be inferred. On the other hand, the naming task can probe for specific meaning activation. Thus far, all the research employing a naming task and examining context strength (e.g., Kellas et al., 1995; Simpson, 1981; Simpson & Krueger, 1991; Vu et al., 1998) has demonstrated that only the contextually appropriate probe word is facilitated in a strong context, whether dominant or subordinate biased. This is problematic for the reordered access model, inasmuch as activation of the inappropriate meanings in strong contexts has not been clearly demonstrated.

In sum, we stand by our original conclusion that strong context can override the subordinate bias effect. The context-sensitive model can account for selective activation and multiple activation, including the SBE, since these are empirical outcomes at particular points on the continuum of context strength. In addition, our research provides confirmation of the fundamental prediction of multiple constraint models, that there should be graded effects for continuous variables. In future research, we plan to examine the graded effects of contextual strength as well.

REFERENCES

AZUMA, T. (1996). Familiarity and relatedness of word meanings: Ratings for 110 homographs. *Behavior Research Methods, Instruments, & Computers*, **28**, 109-124.

BINDER, K. S., & RAYNER, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review*, **5**, 271-276.

CRAMER, P. (1970). A study of homographs. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 361-382). New York: Academic Press.

DOPKINS, S., MORRIS, R. K., & RAYNER, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory & Language*, **31**, 461-476.

DUFFY, S. A., MORRIS, R. K., & RAYNER, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory & Language*, **27**, 429-446.

FODOR, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

FORSTER, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett* (pp. 27-85). Hillsdale, NJ: Erlbaum.

HOGABOAM, T. W., & PERFETTI, C. A. (1975). Lexical ambiguity and sentence comprehension. *Journal of Verbal Learning & Verbal Behavior*, **14**, 265-274.

KELLAS, G., MARTIN, C., YEHLING, K., HERMAN, R., & VU, H. (1995, November). *Contextual strength as a determinant of the subordinate-bias effect*. Poster presented at the 36th Annual Meeting of the Psychonomic Society, Los Angeles.

MARTIN, C., VU, H., KELLAS, G., & METCALF, K. (in press). Strength of discourse context as a determinant of the subordinate bias effect. *Quarterly Journal of Experimental Psychology*.

NELSON, D. L., MCEVOY, C. L., WALLING, J. R., & WHEELER, J. W., JR. (1980). The University of South Florida homograph norms. *Behavior Research Methods & Instrumentation*, **12**, 16-37.

PAUL, S. T., KELLAS, G., MARTIN, M., & CLARK, M. B. (1992). Influence of contextual features on the activation of ambiguous word meanings. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 703-717.

PERFETTI, C. A., LINDSEY, R., & GARSON, B. (1971). *Association and uncertainty: Norms of association to ambiguous words*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center.

RAYNER, K., & DUFFY, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, **14**, 191-201.

RAYNER, K., PACHT, J. M., & DUFFY, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory & Language*, **33**, 527-544.

SIMPSON, G. B. (1981). Meaning dominance and semantic context in the processing of lexical ambiguity. *Journal of Verbal Learning & Verbal Behavior*, **20**, 120-136.

SIMPSON, G. B., & KRUEGER, M. A. (1991). Selective access of homograph meanings in sentence context. *Journal of Memory & Language*, **30**, 627-643.

TABOSSI, P., COLOMBO, L., & JOB, R. (1987). Accessing lexical ambiguity: Effects of context and dominance. *Psychological Research*, **49**, 161-167.

TWILLEY, L. C., DIXON, P., TAYLOR, D., & CLARK, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, **22**, 111-126.

VAN PETTEN, C., & KUTAS, M. (1987). Ambiguous words in context: An event-related potential analysis of the time course of meaning activation. *Journal of Memory & Language*, **26**, 188-208.

VU, H., KELLAS, G., & PAUL, S. T. (1998). Sources of sentence constraint on lexical ambiguity resolution. *Memory & Cognition*, **26**, 979-1001.

NOTES

1. We further note that there are at least three items (*bark*, *pen*, *speaker*) that were eliminated from our research as being problematic but have been consistently used in studies supporting the reordered access model (e.g., Dopkins, Morris, & Rayner, 1992; Duffy et al., 1988; Rayner & Duffy, 1986; Rayner et al., 1994).

2. Our procedures were explicitly explained in the original Kellas et al. (1995) and also in the Martin et al. (in press) paper under review. Thus, we are puzzled by footnote 2 from Binder and Rayner (1998), where it was mentioned that we used a 5-point rating scale, from *weakly biasing* to *strongly biasing*.

(Manuscript received January 15, 1998;
revision accepted for publication April 13, 1998.)