Title
Effect of Global Context on Homophone Ambiguity Resolution

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 30(30)

ISSN
1069-7977

Authors
Mirman, Daniel
Magnuson, James S.
Strauss, Ted J.
et al.

Publication Date
2008

Peer reviewed

Effect of Global Context on Homophone Ambiguity Resolution

Daniel Mirman (daniel.mirman@uconn.edu)
James S. Magnuson (james.magnuson@uconn.edu)
Ted J. Strauss (ted.strauss@uconn.edu)
James A. Dixon (james.dixon@uconn.edu)
Department of Psychology, University of Connecticut
406 Babbidge Rd., Storrs, CT 06269-1020 USA

Abstract

Effects of context are pervasive throughout perceptual and cognitive processing domains. Many studies have shown context effects in language processing, but these studies have mostly focused on local, linguistic contexts. As a step toward situating language processing in the broad scope of cognitive processing, we investigated the effect of a global, non-linguistic context on homophone ambiguity resolution. The context was implicitly induced by using only highly imageable target words. This context was predicted to shift attention away from non-imageable meanings, thus reducing activation of non-imageable meanings and consequently reducing ambiguity between meanings. We tracked eye movements as subjects heard spoken words and selected a matching picture from four displayed items. Results were consistent with the predictions: response times were faster for homophones with only one contextually appropriate meaning than homophones with two contextually appropriate meanings (reflecting reduced ambiguity) and participants were less likely to fixate semantic associates of contextually inappropriate meanings than contextually appropriate meanings (reflecting reduced activation of non-imageable meanings). These results suggest that global, non-linguistic contexts influence language processing by shifting attention away from contextually inappropriate meanings.

Keywords: homophones; ambiguity resolution; context effects; attention; eye-tracking; interactive processing.

Introduction

Effects of context are ubiquitous throughout language processing, typically speeding recognition and ambiguity resolution. Two major cases are that recognition of speech sounds is influenced by lexical context (McClelland, Mirman, & Holt, 2006) and recognition of words is influenced by sentence context (Gorfein, 2001). In general, the contexts that have been studied have two features in common. First, the contexts are local in the sense that the context immediately surrounds the target sound or word and changes from trial to trial (i.e., there is a new word or sentence context on each trial). Second, the contexts are linguistic in the sense that the knowledge and mechanisms involved in processing the context words or sentences are the same as those involved in processing the target sounds or words and resolving ambiguities. This leaves open the possibility that language processing is functionally isolated from other cognitive processing. To test this possibility, we examined whether a global (experiment-level), non-linguistic context can influence a critical aspect of word recognition: resolving the ambiguity that arises when a word has multiple meanings.

A few studies have examined contexts that are less local. Paragraph context effects on homophone ambiguity resolution parallel the effects of sentence context (Kambe, Rayner, & Duffy, 2001) and passage titles help readers to resolve word-level ambiguities (Wiley & Rayner, 2000). These studies examine larger-scale contexts, but the contexts under investigation are, as in other studies, local and linguistic. That is, they can be described in terms of word-to-word semantic relations or syntactic constraints. Visual world paradigm eye-tracking studies have provided some evidence of non-linguistic context effects by showing that visual information can influence resolution of syntactic referential ambiguities (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). We extended this approach to study the effects of global (experiment-level rather than trial-level) contexts on resolving ambiguity between meanings of a single word. Examining global, non-linguistic context effects on word processing is a step toward situating language processing in the broad scope of cognitive processing.

Context effects are a natural case for examining interactions between language and other cognitive processes because context effects are common in many non-linguistic domains, from vision (e.g., Bar, 2004) to motor action planning (Cooper, Schwartz, Yule, & Shallice, 2005). Domain-general views of context effects such as graded interactive constraint satisfaction (McClelland, 1993) and Bayesian inference (Geisler & Diehl, 2003) predict that non-linguistic contexts should influence word processing.

One way that global non-linguistic context can influence language processing is through attention. Previous work on attention and language processing suggested that attention works by damping activation of dis-attended representations (Mirman, McClelland, Holt, & Magnuson, 2008). Following this view of attention, we predict that global context will damp activation of the contextually inappropriate meanings of ambiguous words. To test this prediction we used the visual world eye-tracking paradigm (Allopenna, Magnuson, & Tanenhaus, 1998; Tanenhaus et al., 1995), which provides a very sensitive measure of activation of individual words and concepts during spoken word recognition.

In a typical visual world experiment, participants see several objects on a display, and hear an instruction to click on one of the objects. An eye-tracker records eye movements, which are closely time-locked to fine-grained details of the spoken instruction. Recent studies using this paradigm have shown that listeners fixate images corresponding to concepts that are semantically associated with the target word more than unrelated distractor images (Huettig & Altmann, 2005; Yee & Sedivy, 2006). This pattern is similar to semantic priming, but the paradigm provides much greater sensitivity and time course information (see Allopenna et al., 1998, for discussion).

The visual world paradigm also provides a global context that may influence semantic ambiguity resolution. In a typical visual world study, all targets are highly imageable nouns (though see e.g., Griffin & Bock, 2000, for studies using scenes that allow verbs to be used). Thus, on a visual world trial, a word such as *deck* is ambiguous because it could refer to two meanings that could appear in the study (a deck of cards or the deck of a boat), but a word such as *bark* is, in the extreme, unambiguous because it refers to only one possible target (the tree part, since the dog behavior meaning is not sufficiently imageable to be a possible target). Thus, the paradigm naturally induces an implicit expectation that only imageable nouns will be encountered. This manipulation is similar to attentional manipulations based on proportion of nonwords, which influence listeners' expectations that spoken items will be words or nonwords in an experiment-wise rather than trial-wise fashion (Mirman, McClelland et al., 2008; Monsell, Patterson, Graham, Hughes, & Milroy, 1992).

The present experiment tests the hypothesis that global context affects homophone ambiguity resolution by shifting attention away from contextually inappropriate meanings and thus reducing their activation. To test this hypothesis we compare processing of homophones for which both meanings are contextually appropriate to homophones that have only one contextually appropriate meaning. The hypothesis makes two specific predictions regarding this comparison. First, reduced activation of contextually inappropriate meanings will reduce competition between meanings for one-meaning-appropriate homophones, thus target meaning activation will be faster for these homophones than for those with two contextually appropriate meanings (reflected in response times and target fixation curves). Second, reduced activation of contextually inappropriate meanings will reduce activation of their semantic associates, thus there will be reduced fixation of semantic associates of non-target meanings of one-meaning-appropriate homophones compared to fixation of semantic associates of non-target meanings of both-meanings-appropriate homophones.

Experiment

Methods

Materials. There were three types of critical words (16 words in each condition): both-meanings-appropriate homophones, one-meaning-appropriate homophones, and unambiguous words. For each target homophone word, a semantic associate (Nelson, McEvoy, & Schreiber, 2004) of the non-target meaning was selected to serve as a semantic competitor (that is, to compete for fixation on the basis of semantic similarity to the non-target meaning of the stimulus word). For example, for the homophone *deck* (both-meanings-appropriate), the target image was a deck of cards and the competitor image was a boat; for the homophone *bark* (one-meaning-appropriate), the target image was tree bark and the competitor image was a dog. The unambiguous condition was used to provide a baseline of semantic competition effects in which the associate is related to a word that is favored both contextually (i.e., it is an imageable noun) and visually (i.e., it is in the display), so the competitor image was a semantic associate of the target word (e.g., for *acorn*, the target image was an acorn and the competitor image was a squirrel). The homophones were all balanced (i.e., both meanings were approximately equally frequent) and the words in all three conditions were matched on length, frequency, and association strength to the competitor image (all $t(15)<1.0$, all $p>0.3$, see Table 1). In addition to the 48 critical trials there were 28 filler trials on which an unambiguous word was presented and there were no related images on the screen.

All stimuli were produced by a female native speaker of American English in a sound-attenuated room and digitized at 44 kHz. The individual words were edited to eliminate silence at the beginning and end of the sound file.

Table 1. Mean (standard deviations in parentheses) properties of stimuli. *Both* refers to both-meanings-appropriate homophones (e.g., *deck*), *One* refers to one-meaning-appropriate homophones (e.g., *bark*), *Unambig.* refers to unambiguous words (e.g., *acorn*).

Property	Both	One	Unambig.
Word Frequency	16.1 (25.0)	18.7 (21.5)	17.9 (25.3)
No. Syllables	1.25 (0.4)	1.25 (0.6)	1.38 (0.5)
No. Phonemes	3.88 (0.81)	3.88 (1.1)	3.75 (1.2)
Duration (ms)	588.3 (106)	607.0 (81.7)	603.4 (78.0)
Association	0.109 (0.08)	0.137 (0.2)	0.158 (0.18)

Procedure. On each trial, participants saw four images on a 17" screen. Each image was presented near one of the screen corners, 15% of the screen size away from the horizontal and vertical edge of the screen; images had a maximum size of 200 x 200 pixels and screen resolution was set to 1024 x 768. Gaze position and duration were recorded using an ASL 6000 remote eye-tracker. Each trial began with a 500 ms preview of the four images. The preview was intended to reveal visual salience differences between images and diminish their effects on the data of interest. At the end of the preview period the target word was presented through headphones and participants had to click on the image corresponding to the target word. The experiment began with 12 practice trials on which feedback was presented. On each critical trial, the display of four images contained a target (e.g., a deck or a piece of bark), a semantic competitor (e.g., a boat or a dog) and two distractors, which were phonologically and semantically unrelated to the target and each other (e.g., a head of lettuce and a bowl of soup). Three items were removed from each of the ambiguous conditions due to low accuracy and/or unequal visual salience between target, competitor, and distractor images.

To control for possible differences in visual salience of the images, a control version was conducted in which the same pictures were presented in the same conditions but target and distractor images were pseudo-randomly re-assigned such that there would be no predicted semantic competition. That is, each picture occurred in the same condition, but the target-distractor pairs were scrambled so that on each trial there were no semantic associates present in the display. In this case the term *competitor* refers to visual salience competition rather than semantic competition.

Participants. Forty undergraduate students at the University of Connecticut participated for course credit, 20 in the semantic associate matched version and 20 in semantic associate scrambled control version. Six participants were excluded due to problems with eye-tracking (4 associate matched, 2 control).

Results

Response time. The response time and accuracy data are shown in Figure 1 (only correct response trials were included in the response time and eye-tracking analyses). The top panel of Figure 1 shows that there were clear differences in response time across the word conditions (Associate matched: $F1(2,30)=30.2$, $p<0.001$, $F2(2,39)=5.6$, $p<0.01$; Associate scrambled: $F1(2,34)=73.7$, $p<0.001$, $F2(2,39)=8.6$, $p<0.01$). Pairwise comparisons revealed that response times for each condition were different from each other condition (Associate matched: all $t>2.2$, all $p<0.05$, except the contrast between the two types of ambiguous words; Associate scrambled: all $t>2.5$, all $p<0.05$, except the contrast between the one appropriate meaning homophones and the unambiguous words).

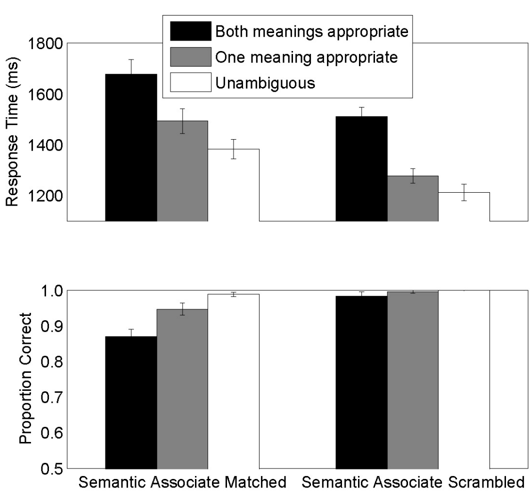

Figure 1. Behavioral data. Top panel shows response times, bottom panel shows accuracy. Error bars reflect 1SE.

Participants were slowest to find the matching picture for ambiguous words with two contextually appropriate words, fastest for unambiguous words, and intermediate for ambiguous words with only one contextually appropriate meaning. This pattern held regardless of whether a semantic associate of the non-target meaning was present in the display. These results suggest a high degree of competition between meanings for the both-meanings-appropriate homophones and reduced, but not eliminated, competition between meanings for the one-meaning-appropriate homophones. Of particular interest is the difference between the two types of homophones, which suggests that there was more competition when both meanings were consistent with the context than when only one meaning was consistent with the context. In addition, the finding that responses to one-meaning-appropriate homophones were slower than to unambiguous words suggests that contextually inappropriate meanings became active and competed with the target meaning, indicating a graded effect of context. That is, the context of the visual world paradigm damped activation of non-imageable meanings relative to imageable meanings, but did not completely eliminate contextually inappropriate meanings from becoming active. The eye-tracking data provide converging evidence for this interpretation.

Eye-tracking data analysis. To analyze the time course of fixations we used growth curve analysis using orthogonal polynomials (Mirman, Dixon, & Magnuson, 2008). Fixation curves were modeled by fourth-order polynomials in order to capture the 3 inflection points of the curve and included individual subject effects on each of the polynomial time factors. Growth curve analysis is part of a family of multilevel modeling techniques in which condition effects can be evaluated based on their effects on parameters of the fixation curve. Orthogonal polynomials produce independent time terms, so effects on different aspects of curve shape can be considered independently. In particular, the intercept term captures overall curve height, the linear term captures the overall slope of the curve, and the quadratic term captures the steepness of the rise and fall of the curve around a central inflection point. Because fixation proportions tend to rise and fall more or less symmetrically in a typical visual world paradigm experiment, differences in fixation time course have the biggest influence on the quadratic component of the fixation curve (see Mirman, Dixon, & Magnuson, 2008, for further discussion of interpretation of model terms).

In the interests of brevity, we discuss only the effect of critical terms on the quadratic component. Condition effects were evaluated based on the extent to which they improve model fit; that is, the change in deviance (ΔD) due to adding the condition parameter to the model. One standard measure of deviance is -2LL (minus 2 times the log-likelihood). This measure is distributed as chi-square, with degrees of freedom equal to the number of parameters added. When there are more than 2 categorical conditions, one condition serves as the baseline and an individual term is estimated for each of the other conditions (thus the degrees of freedom are one less than the number of conditions). Individual conditions can be compared to the baseline by standard significance tests on the parameter estimates.

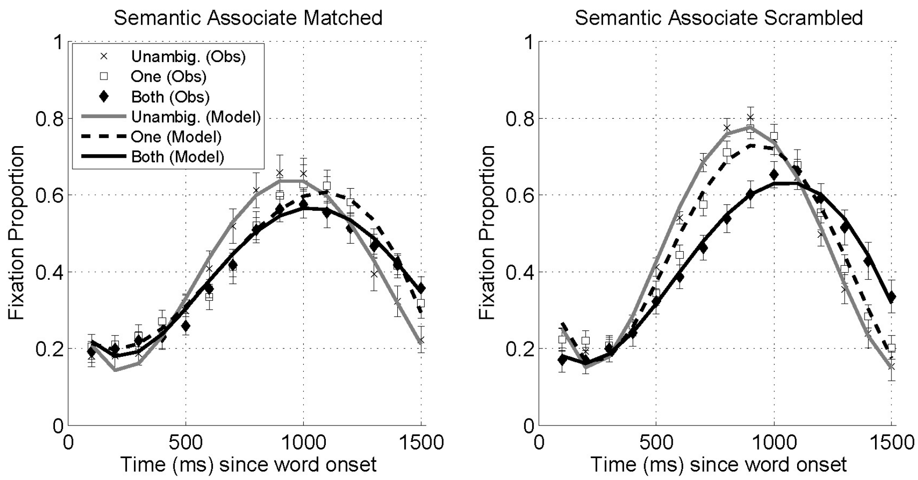

Figure 2. Observed and model fit proportion of fixations to target images for unambiguous words (*Unambig*.; x's, gray line), homophones with only one contextually appropriate meaning (*One*; white squares, dashed line), and homophones with two contextually appropriate meanings (*Both*; black diamonds, solid line). Left panel shows data from semantic associate matched version, right panel shows data from semantic associate scrambled control version. Error bars reflect 1SE.

Target fixation. The average fixation proportions over time to the target image and model fits are plotted in Figure 2. Consistent with the response time data, the time course of fixation was fastest for the unambiguous targets, somewhat slower for the one-meaning-appropriate homophones, and slowest for both-meanings-appropriate homophones. There was a significant effect of word condition (e.g., differences between both-meanings-appropriate, one-meaning-appropriate, and unambiguous words) on the quadratic term when the competitor was a semantic associate ($\Delta D=46.3$, $p<0.0001$) and when the semantic associates were scrambled such that there was no semantic associate in the display ($\Delta D=138.4$, $p<0.0001$). Comparison of the parameter estimates confirmed that the fixation time course for unambiguous words was significantly different from the time course for both-meanings-appropriate homophones (Semantic associate matched: B = -0.229, $t(600)=6.13$, $p<0.0001$; Semantic associate scrambled: B = -0.370, $t(676)=12.05$, $p<0.0001$), and one-meaning-appropriate homophones (Semantic associate matched: B = -0.193, $t(600)=5.66$, $p<0.0001$; Semantic associate scrambled: B = -0.128, $t(676)=4.65$, $p<0.0001$). The difference between the types of homophones was reliable when the semantic associates were scrambled (B = -0.242, $t(676)=7.87$, $p<0.0001$), but did not reach significance when they were matched (B = -0.036, $t(600)$ = 0.96, n.s.). This general pattern is consistent with previous studies showing that ambiguity slows word recognition and converges with the response time data to suggest that participants were slower to activate the target meanings of homophones when both meanings were contextually appropriate compared to when only one was appropriate.

Semantic competition. Fixation proportions and model fits for the semantic competitor and unrelated images for the semantic associate matched version are plotted in the top row of Figure 3. As in priming experiments, we assume that activation of a semantic associate is an index of activation of a particular meaning. Non-negligible activation of a particular meaning should be reflected by a greater proportion of fixations to a semantic associate of that meaning than an unrelated image. Our unambiguous condition (Figure 3, top right panel) replicates previous findings of semantic competition (Huettig & Altmann, 2005, Yee & Sedivy, 2006): there are more fixations to a semantic associate (*Competitor*) image than an unrelated image. The both-meanings-appropriate homophones (top left panel) show the same semantic competition pattern, but the one-meaning-appropriate homophones (top middle panel) show weak, if any, semantic competition.

There was a significant word-by-image condition interaction effect on the quadratic term ($\Delta D=52.6$, $p<0.0001$), indicating that the competitor-vs-unrelated contrast differed across the word conditions. Evaluation of the parameter estimates revealed that, relative to unrelated images, participants were approximately equally likely to fixate semantic associates of unambiguous target words and semantic associates of non-target contextually appropriate meanings (B = -0.019, $t(1260)=0.605$, n.s.), but they were less likely to fixate semantic associates of non-target contextually inappropriate meanings (B = -0.208, $t(1260)$ = 6.47, $p<0.0001$). That is, there was little difference in semantic competition between the both-meanings-appropriate and unambiguous conditions and large differences in semantic competition between the both-meanings-appropriate and one-meaning-appropriate conditions. Put simply, semantic associates of contextually appropriate meanings were fixated more than semantic associates of contextually inappropriate meaning, indicating that contextually appropriate meanings were activated more than contextually inappropriate meanings.

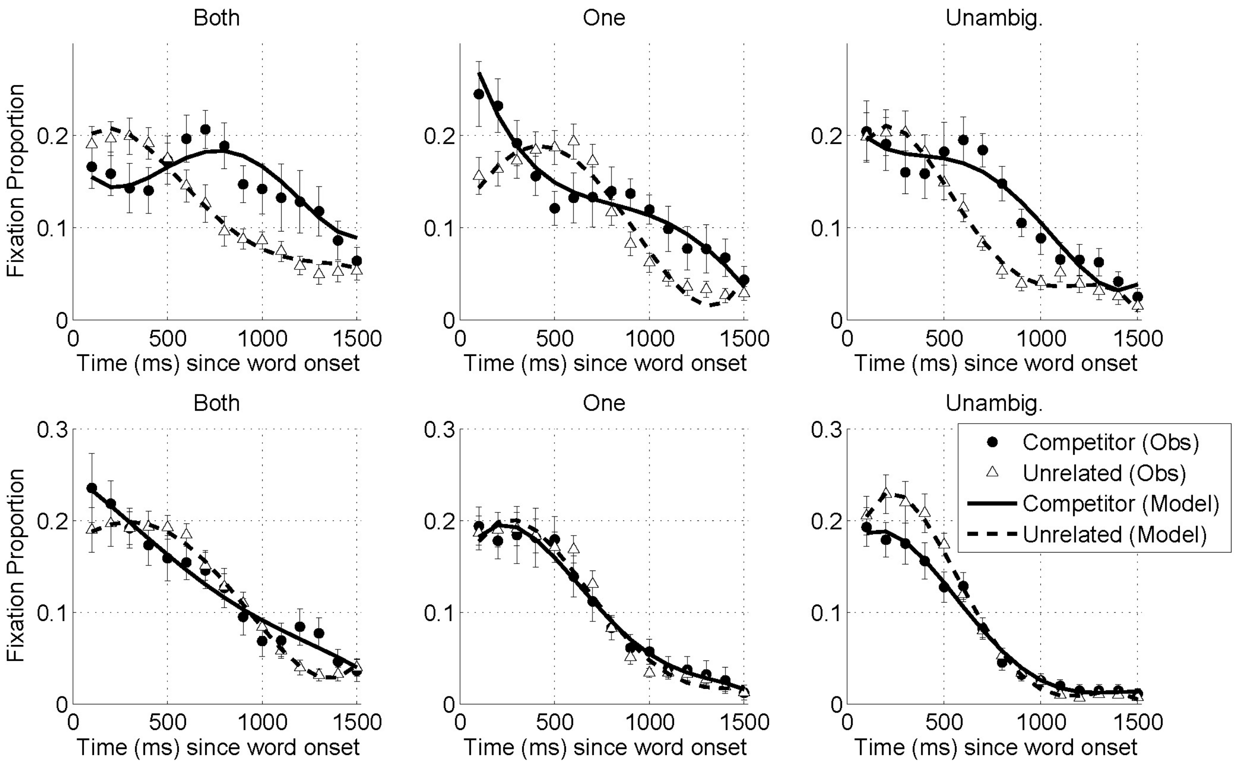

Figure 3. Observed and model fit proportion of fixations to competitor images (black circles, solid line) and unrelated images (white triangles, dashed line). *Both* refers to both-meanings-appropriate homophones, *One* refers to one-meaning-appropriate homophones, *Unambig*. refers to unambiguous words. Top row shows data from semantic associate matched version, bottom row shows data from semantic associate scrambled control version. Error bars reflect 1SE.

To examine whether contextually inappropriate meanings became active at all, we restricted the semantic competition analysis just to the one-meaning-appropriate homophones. This test addresses the question of whether the pragmatic constraint under investigation can completely eliminate contextually inappropriate meanings from activation. There was a significant effect of semantic relatedness on the quadratic term ($\Delta D=14.3$, $p<0.0001$), reflecting a difference in the time course of fixation of semantic associates relative to unrelated distractors. These results indicate that the non-target, contextually inappropriate meanings did become partially active, leading to a somewhat greater fixation of their semantic associates relative to the unrelated baseline.

Fixation proportions and model fits for the semantic competitor and unrelated images for the control version are plotted in the bottom row of Figure 3. Visually, it is clear that random re-assignment of images eliminated the competition effect (i.e., there is no difference between competitor and unrelated image fixation proportions) for all word conditions, indicating that the results of the semantic associate matched version cannot be attributed to visual salience differences. There was a small bias favoring the unrelated distractors ($\Delta D=8.3$, $p<0.01$; all B's>0, opposite to the pattern when the competitor was a semantic associate). That is, participants were slightly more likely to fixate the unrelated distractors than the visual salience (scrambled) competitors, particularly for the unambiguous words and both-meanings-appropriate homophones. This suggests that the semantic competition effects found in the semantic associate matched version had to overcome a slight visual bias to look at unrelated distractor images.

Discussion and Conclusions

We examined the impact of global, non-linguistic context on the activation of contextually appropriate and contextually inappropriate meanings. Participants heard frequency balanced homophones that either had two contextually appropriate meanings or only one contextually appropriate meaning. The prediction was that if the context influences ambiguity resolution, then the non-target meaning of both-meanings-appropriate homophones should be more active than the non-target meaning of one-meaning-appropriate homophones because the former is contextually favored but the latter is not. Consistent with this prediction, the response time and target fixation data suggest that there was much more competition for the both-meanings-appropriate homophones than one-meaning-appropriate homophones. Semantic competition data provided converging evidence that there was greater activation of non-target contextually appropriate meanings than non-target contextually inappropriate meanings. Evidence of modest semantic competition in the one-meaning-appropriate condition suggested that this contextual effect was graded – contextually inappropriate meanings did become partially active, but not as active as contextually appropriate meanings.

Previous studies of homophone ambiguity resolution have generally focused on local, linguistic contexts. This allowed the possibility that contextual constraints might be limited to cases where there are direct syntactic or semantic connections between context and targets. By using a global, non-linguistic context we have shown the domain-general nature of context effects on ambiguity resolution. Indeed, ambiguity resolution is a key issue in many perceptual and cognitive domains, from object recognition to speech perception, and context effects are pervasive in those domains as well (Bar, 2004; McClelland, et al., 2006). The present demonstration of effects of global, non-linguistic context on homophone ambiguity resolution suggests domain-general mechanisms of contextual influences.

Attention is one domain-general mechanism that is consistent with the present results. On this view, participants' attention was partially shifted away from non-imageable concepts, which did not occur as targets. This shift in attention would have the consequence of reducing activation of non-imageable concepts (Mirman, McClelland, et al., 2008). The reduced activation of non-target meanings reduced ambiguity, thus speeding the resolution process and yielding the observed faster response times for one-meaning-appropriate homophones than both-meanings-appropriate homophones. In addition, the reduced activation caused less activation of semantic associates, thus yielding the observed reduced fixation of semantic associates for one-meaning-appropriate homophones. In sum, the present results support a domain-general view of context effects and language processing in which all aspects of the context influence homophone ambiguity resolution and global contexts influence processing by shifting attention away from contextually inappropriate meanings.

Acknowledgements

This research was supported by National Institutes of Health grants DC005765 to JSM, F32HD052364 to DM, and HD001994 and HD40353 to Haskins Labs.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory & Language, 38*(4), 419-439.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629.

Cooper, R. P., Schwartz, M. F., Yule, P., & Shallice, T. (2005). The simulation of action disorganization in complex activities of daily living. *Cognitive Neuropsychology, 22*(8), 959-1004.

Geisler, W. S., & Diehl, R. L. (2003). A Bayesian approach to the evolution of perceptual and cognitive systems. *Cognitive Science: A Multidisciplinary Journal, 27*(3), 379-402.

Gorfein, D. S. (2001). *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity*. Washington, DC, US: American Psychological Association.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274-279.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition, 96*(1), B23-B32.

Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory & Cognition, 29*(2), 363-372.

McClelland, J. L. (1993). Toward a theory of information processing in graded, random, interactive networks. In D. E. Meyer & S. Kornblum (Eds.), *Attention & performance xiv: Synergies in experimental psychology, artificial intelligence and cognitive neuroscience* (pp. 655-688). Cambridge, MA: MIT Press.

McClelland, J. L., Mirman, D., & Holt, L. L. (2006). Are there interactive processes in speech perception? *Trends In Cognitive Sciences, 10*(8), 363-369.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008a). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, doi:10.1016/j.jml.2007.1011.1006.

Mirman, D., McClelland, J. L., Holt, L. L., & Magnuson, J. S. (2008b). Effects of attention on the strength of lexical influences on speech perception: Behavioral experiments and computational mechanisms. *Cognitive Science, 32*(2), 398-417.

Monsell, S., Patterson, K. E., Graham, A., Hughes, C. H., & Milroy, R. (1992). Lexical and sublexical translation of spelling to sound: Strategic anticipation of lexical status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 452-467.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The university of south florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments & Computers, 36*(3), 402-407.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 632-634.

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: Evidence from eye movements. *Memory & Cognition, 28*(6), 1011-1021.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 1-14.