Journal of Experimental Psychology:
Learning, Memory, and Cognition
2006, Vol. 32, No. 4, 943–948

0278-7393/06/$12.00 DOI: 10.1037/0278-7393.32.4.943

Speakers Gaze at Objects While Preparing Intentionally Inaccurate Labels for Them

Zenzi M. Griffin
Georgia Institute of Technology and Stanford University

Daniel M. Oppenheimer
Princeton University and Stanford University

When describing scenes, speakers gaze at objects while preparing their names (Z. M. Griffin & K. Bock, 2000). In this study, the authors investigated whether gazes to referents occurred in the absence of a correspondence between visual features and word meaning. Speakers gazed significantly longer at objects before intentionally labeling them inaccurately with the names of similar things (e.g., calling a horse a *dog*) than when labeling them accurately. This held for grammatical subjects and objects as well as agents and patients. Moreover, the time spent gazing at a referent before labeling it with a novel word or accurate name was similar and decreased as speakers gained experience using the novel word. These results suggest that visual attention in speaking may be directed toward referents in the absence of any association between their visual forms and the words used to talk about them.

Keywords: language production, picture description, eye movement monitoring, object naming, deception

In communicating about the world around them, people tend to look at an object immediately before mentioning it (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; Richardson & Dale, 2005). Gazes to objects anticipate order of mention regardless of who did what to whom (e.g., *A girl is kicking a boy* vs. *A boy is kicking a girl*), voice (e.g., *A mailman is chasing a dog* vs. *A mailman is being chased by a dog*; Griffin & Bock, 2000), and for different ways of expressing time (e.g., *Fifteen past 10* and *Ten 15*; Bock, Irwin, Davidson, & Levelt, 2003). The link between gaze and speech is so reliable that infants make use of speakers' eye gaze in establishing the referents of new words (Baldwin, 1991, 1993).

Of course, a speaker must identify an object to name it accurately, and so the need for object identification could motivate speech-related gazes. However, the gazes that precede object names often occur after an object has already been fixated once or after it has been identified with parafoveal information (Griffin & Bock, 2000; Morgan & Meyer, 2005; see Griffin, 2004b, for review). Moreover, object recognition for such stimuli usually takes less than 150 ms (e.g., Biederman, Mezzanotte, & Rabinowitz, 1982; for review, see Henderson & Ferreira, 2004). Yet, gazes preceding object names persist for much longer, circa 500–1,500 ms (Griffin, 2001; Meyer et al., 1998). If object identification cannot explain the duration of gazes, then why are gazes and word preparation so tightly linked?

One reason, henceforth referred to as the *content hypothesis,* is that speakers may benefit from continued processing of visual information from the object. The visual features of an object lead to its initial identification by activating stored structural representations that in turn activate semantic/conceptual representations (e.g., Humphreys & Forde, 2001). After an object is recognized, its visual features may continue to activate its structural and conceptual representations, activating potential names for the object. Indeed, several researchers have suggested that gazing at an object may maintain or increase the activation of its mental representations, facilitating production of its name via cascaded activation (e.g., Humphreys & Forde, 2001; Humphreys, Riddoch, & Price, 1997; Roelofs, 1992; Van Der Meulen, 2001). Gazes might also aid in producing the semantic and pragmatic content of an utterance and in retrieving the sounds of words. Also, continuing to gaze at a referent after identifying it may prevent other objects from being viewed and interfering with these word production processes (cf. Glaser & Glaser, 1989; Harley, 1990; Lupker, 1988; Meyer & Van Der Meulen, 2000).

Despite the intuitive appeal of the content hypothesis, there is reason to doubt its validity. For example, several studies have indicated that speakers are untroubled by viewing semantically related or unrelated objects while generating single words (Bloem & La Heij, 2003; Damian & Bowers, 2003; La Heij, Heikoop, Akerboom, & Bloem, 2003). Other studies have suggested that speakers do not benefit from having objects available to gaze at continuously while they prepare to name them (Bock et al., 2003; Griffin, 2004a; La Heij, Van Der Heijden, & Plooij, 2001).

What else could lead speakers to gaze at already identified objects while preparing their names? When people attend to ob-

Zenzi M. Griffin, School of Psychology, Georgia Institute of Technology, and Department of Psychology, Stanford University; Daniel M. Oppenheimer, Department of Psychology, Princeton University, and Department of Psychology, Stanford University.

This research was supported by National Institutes of Health Grant R03 MH61318-01. Daniel M. Oppenheimer was supported under a National Science Foundation Research Fellowship. We are grateful to Demetria Mitchell and Azucena Rangel for their help in running participants and analyzing the eye-movement data. We also thank the following programmers who were responsible for the experimental and data-treatment software: Prateek Bansal, Madan Jampani, Jini Khetan, and Amit Mookerjee.

Correspondence concerning this article should be addressed to Zenzi M. Griffin, School of Psychology, Georgia Institute of Technology, 654 Cherry Street, Atlanta, GA 30332-0170. E-mail: zenzi.griffin@psych.gatech.edu

jects, they create spatial indices for their locations (e.g., Hommel, 2002; Kahneman, Treisman, & Gibbs, 1992; Pylyshyn, 2001). If an object (or information associated with an object) again becomes relevant to cognitive processing, then its spatial index is activated, and gaze tends to return to the object's current (see Just & Carpenter, 1976), earlier (see Spivey, Richardson, & Fitneva, 2004), or imagined (see Altmann & Kamide, 2004) location. Thus, an alternative to the content hypothesis—the *referential hypothesis*—is that speech-related gazes follow from speakers' intentions to refer to particular objects. Rather than being mediated by the visual–conceptual features of the referent object, gaze is mediated by the spatial index of the referent (see Griffin, 2004b). In other words, the intention to talk about an object orients a speaker's visual attention to the object's spatial location while processes related to the object are carried out.

As long as speakers accurately refer to visible objects, the content and referential hypotheses make similar predictions and can account for the same observations. According to both hypotheses, speakers will gaze at accurately named objects until their names are retrieved. However, the hypotheses differ in predicting how speakers will gaze at objects that they intentionally label inaccurately. According to the content hypothesis, as the word used to refer to an object becomes less related to it, gazing at the object provides less facilitation and potentially more interference for word production processes. Speakers should therefore become less likely to fixate on it. Ideally, speakers should avert their gaze to blank or neutral content as observed when speakers generate words from memory rather than the current environment (see Glenberg, Schroeder, & Robertson, 1998). In contrast, the referential hypothesis suggests that speakers will gaze at objects while they engage in linguistic (or other cognitive) processes related to the objects, despite mismatches between visual and linguistic content. The fact that a label for an object changes does not alter the fact that it is that particular object that is referred to, and thus, gazes to the referent should persist. The contrast between the content and referential hypotheses roughly maps onto Frege's (1948/1997) famous distinction between sense and reference. We tested the hypotheses by asking speakers to describe one actor in a scene accurately or inaccurately, with either the name of a similar concept (e.g., *raft* for ship) or a novel word (*blick*).

Method

Participants

Sixteen undergraduates at the Georgia Institute of Technology participated for extra credit in an introductory psychology course. All participants considered English to be their first language and reported normal or corrected-to-normal vision.

Apparatus

Eye movements were monitored with a remote video-based pupil/corneal reflection system—an ISCAN (Burlington, Massachusetts) ETL-400 with a high-speed upgrade, sampling at 120 Hz. Stimuli were displayed on a 21-in. (53.34-cm) monitor. Speech was recorded at 12 kHz via a SoundBlaster (Milpitas, California) card and a headset microphone. Participants placed their foreheads against a rest to prevent movements in depth. Scenes subtended about 26° of visual angle horizontally.

Materials

Twenty-four line drawings of simple scenes composed the experimental items. Half had animate patients (e.g., the mouse in a scene of a turtle squirting a mouse with a squirt gun), and half had inanimate ones (e.g., the ship in a scene of a torpedo hitting a ship). The animacy of the agents also varied to ensure that inanimacy was not correlated with patienthood. Five additional scenes served as examples and practice items.

Procedure and Design

Participants were instructed to describe the action in each scene using a single sentence in the active voice. Patients were identified as things affected by actions or acted on. Half of the participants were asked initially to refer to "living" patients accurately but "nonliving" patients inaccurately. The other half of participants was asked to refer to inanimate patients accurately but animate patients inaccurately. Within each of these groups, half of the participants were told to use the novel word *blick* as the inaccurate reference, whereas the other half were told to use the name of a similar but different concept (e.g., *rabbit* for mouse or *raft* for ship).

After responding to all experimental items under initial instructions in a first block of trials, participants responded to them again but switched the animacy of the patients that they inaccurately described. For example, a speaker who inaccurately described animate patients in the first block of trials would inaccurately describe inanimate patients in the second. Thus, speakers who followed instructions perfectly would describe the patient in each scene once accurately and once inaccurately. Furthermore, participants switched the form of inaccurate references. Thus, speakers who followed instructions perfectly would use *blick* for exactly half of the inaccurate descriptions and similar object names for the other half of the inaccurate descriptions. Items appeared in one fixed pseudorandom order for the first block and a different pseudorandom order for the second block. Within the second block, half of the scenes were presented in the same left–right orientation as the first block, and half were mirror reversed. The order in which participants used the inaccurate reference forms and applied them to the animacy categories was completely counterbalanced across participants.

A validation point presented in the top center of the screen preceded each trial. After fixating within 2° of this point for 400 ms or recalibrating, a scene was presented black on white. Scenes remained in view until participants completed their descriptions and the experimenter pressed a button to proceed to the next trial.

Data Treatment

Responses were transcribed by a naïve research assistant. Trials in which participants referred to patients before the verb, inaccurately described agents, or produced both accurate and inaccurate labels for patients within the same trial were excluded, as were trials in which speakers produced the wrong type of inaccuracy (e.g., *blick* instead of a similar concept's name). However, if speakers only produced an accurate label when an inaccurate one was called for or vice versa (6.5% of trials), then the trial was still included. Of the 768 trials, 701 (91.1%) met criteria for inclusion.

Two speech-processing programs, Sphinx-2 (CMU Sphinx Group, 2001) and Fasttalk (Georgia Institute of Technology/Nexidia, 2004), produced forced alignments from sound recordings and transcriptions that were used for agent and patient noun onsets. Values from the two were highly related (R^2s > 0.988). Occasionally, Sphinx-2 failed to produce onsets for words, whereas Fasttalk sometimes yielded estimates with confidence levels of zero. As long as the estimated agent onsets were greater than amplitude-based measures of speech onsets (Bansal, Griffin, & Spieler, 2001), values used came from Sphinx-2 (84% of analyzed trials); otherwise, they came from Fasttalk (12%) or measurements made with an audio editor (4%). As long as estimated patient onsets were greater than selected agent onsets and less than amplitude-based measures of speech offsets, values came from

Table 1
Descriptive Statistics Based on Participant Analyses Collapsing Over Block Order

	Similar		Accurate		Blick		Accurate	
Measure	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Speech onset (ms)	1,918	170	1,824	139	1,759	122	1,648	70
Agent onset–speech onset (ms)	714	168	486	132	359	100	331	70
Patient onset–agent onset (ms)	2,735	276	1,363	100	1,280	68	1,343	107
Eye–voice span (ms)	2,060	201	1,347	122	1,134	95	1,180	87
End last gaze relative to name onset (ms)	363	128	252	86	99	51	184	58

Sphinx-2 (84%); otherwise, they came from Fasttalk (13%) or audio-editor measurements (3%).

Saccades were defined as sequences of sampled fixation coordinates spanning at least 0.5° of visual angle, for which the velocity of movement exceeded 40° per second, beginning and ending when velocities fell below 20° per second. Remaining samples constituted fixations if they had a minimum duration of 50 ms. For each scene, polygons surrounding the agents and patients with a margin of approximately 1° of visual angle were used to identify fixated regions. *Gazes* were defined as beginning at the onset of a fixation within a polygon and ending with offset of the last fixation within it. Thus, the duration of saccades within a polygon contributed to gaze durations (see Irwin, 2004). Time spent gazing at agents, patients, and other areas of scenes prior to the onsets of agent nouns and between agent and patient nouns was calculated for each trial and region. From these values, mean gaze times per time period, region, and condition were calculated for participants and items.

All statistical tests were performed on means of participants and items (see Clark, 1973). Unless otherwise indicated, reported inferential statistics reflect significant effects at $p < .05$. Scenes were repeated in the second block of trials, and this repetition often facilitated responses. Because the effect of repetition was not of interest, the block factor is only reported below when it interacts with other factors. Repeated measures analyses of variance were carried out with "Name Accuracy" as a within-participants and within-items factor, and "Presentation Block" (1st or 2nd) as a between-participants and within-items factor. For analyses of eye-movement data, "Region" was a within-participants and within-items factor.

Results

Similar-Concept Names

Mean values are reported in Table 1. Speakers took only a nonsignificant 94 ms longer to begin utterances in which patients received the names of similar concepts as opposed to accurate names, Fs < 1.3. The latency between speech onset and agent nouns was a nonsignificant 228 ms greater when patient names were inaccurate, Fs < 2.6. However, the latency between the onset of the agent noun and the patient noun was twice as large when patients were named inaccurately rather than accurately. This 1,372-ms difference in latencies was significant, $F_1(1, 14) = 24$, $MSE = 609{,}575$; $F_2(1, 23) = 41$, $MSE = 995{,}818$. Analyses of disfluency rates for subject noun phrases, verbs, and object noun phrases showed a similar pattern. That name accuracy only affected timing after speech began suggests that speakers spent no extra time selecting similar concepts or preparing inaccurate patient names before starting to articulate their sentences.

Before articulating patient names, speakers gazed at patients[1] ($M = 3{,}216$ ms) for more time than at agents ($M = 1{,}738$ ms) or other regions ($M = 957$ ms), resulting in a significant main effect of region: $F_1(2, 28) = 51$, $MSE = 822{,}200$; $F_2(2, 46) = 35$, $MSE = 3{,}781{,}808$. Collapsing across regions, mean gaze times were 1,127 ms longer when patients were named inaccurately rather than accurately, so there was a main effect of accuracy: $F_1(1, 14) = 7.46$, $MSE = 4{,}087{,}626$; $F_2(1, 23) = 22$, $MSE = 2{,}811{,}333$. Most of this increase in gaze time was spent on patients (the object inaccurately named), which resulted in a significant interaction between region and accuracy: $F_1(2, 28) = 5.84$, $MSE = 377{,}602$; $F_2(2, 46) = 5.27$, $MSE = 1{,}539{,}599$. Mean gaze times are shown on the left side of Figure 1. No other effects or interactions approached significance by participants and items. Relative timing data further supported the conclusion that speakers gazed at patients while preparing inaccurate names. Eye–voice spans were a significant 713 ms greater before the names of similar objects: $F_1(1, 14) = 9.34$, $MSE = 435{,}685$; $F_2(1, 23) = 20$, $MSE = 1{,}176{,}503$. Speakers did not shift their eyes away from objects until well after they began referring to them. The gaze-offset times did not differ significantly, Fs < 1. The same pattern of effects was found when analyses were limited to fluent utterances.

Novel Word Labels

Speakers initiated wholly accurate utterances after pictures were displayed with mean latencies of 1,648 ms. When speakers used the novel word *blick*, utterances began a nonsignificant 111 ms later, Fs < 2.5. The smaller 28-ms difference in when speakers began to articulate agent nouns was not significant either, Fs < 1. Surprisingly, the onsets of patient nouns began a nonsignificant 63 ms earlier for utterances that contained *blick* rather than accurate names, Fs < 1. So, the timing of speech alone shows no significant difference between a speaker using a novel word as opposed to an accurate name. Analysis of disfluency rates for the constituents of the sentence showed that speech tended to be more fluent before novel than accurate names.

The critical question was whether speakers gazed at patients even when preparing this repeated novel word. Total times spent gazing at each region were calculated for the period prior to the onset of the patient noun. The right side of Figure 1 shows these mean gaze times. Speakers spent more time gazing at patients and agents than elsewhere, which resulted in a significant main effect

[1] Short movies illustrating the timing of speech and eye movements in these experiments are available at http://oak.psych.gatech.edu/~zgriffin/

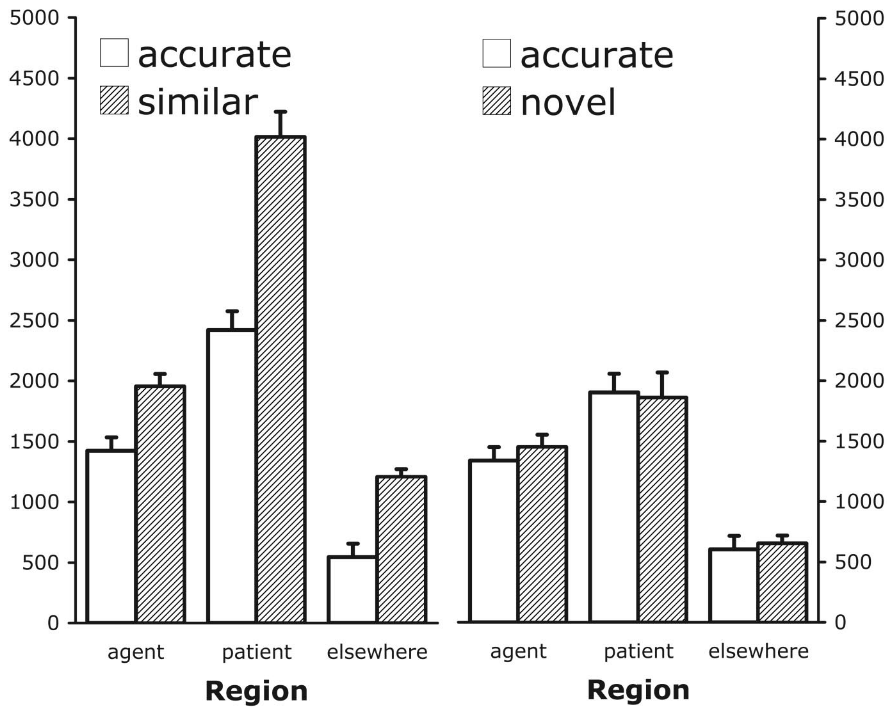

Figure 1. Mean time spent (in milliseconds) gazing at regions prior to the onset of a reference to the patient for sentences containing accurate and inaccurate similar object labels for patients (left) and for sentences containing accurate and novel word labels for patients (right). Error bars show point estimates for standard errors of the means.

of region: $F_1(2, 28) = 61$, $MSE = 208{,}874$; $F_2(2, 46) = 18$, $MSE = 1{,}966{,}739$. Speakers spent a nonsignificant 43 fewer ms gazing at patients before uttering *blick* compared with naming them accurately. Therefore, the mean time spent gazing at an object before referring to it as *blick* was 97.7% of the mean time gazing at it before naming it accurately. There was no main effect or interactions involving name accuracy, $F\text{s} < 1$. Similarly, mean eye–voice spans for accurate names and the novel word differed by only a nonsignificant 46 ms, $F\text{s} < 1$.

If gaze durations reflect the difficulty of name preparation, then the repeated use of *blick* should have made the novel word easier to produce across the experiment, especially relative to generating accurate names without the benefit of repetition. Indeed, previous research has demonstrated that speakers generally spend less time gazing at objects when repeating their (accurate) names (Van Der Meulen, Meyer, & Levelt, 2001). We tested whether this also held for novel names with no association to their referents. For every item in the first block of trials, mean gaze times on patients between noun onsets for *blick* trials were subtracted from mean gaze times for accurate name trials. This yielded a measure that controlled for item differences in the identifiability of patients. Early in the block, speakers gazed at referents longer when naming them with *blick* than with an accurate name; however, with repeated production of the novel word, gaze durations before *blick* became shorter than those before accurate names. The difference between these gaze durations on patients significantly decreased as speakers gained experience using *blick* across increasing trial positions, $r(n = 24) = -.459$, $p < .03$. However, much of this correlation may be accounted for by decreasing latencies between nouns that also correlated highly with trial position, $r(n = 24) = -.389$, $p < .07$.

All of the analyses above were based on gazes in which multiple fixations on an object may be collapsed together into a single gaze. If fixations were used for taking in useful information for word retrieval, then speakers might have made more fixations when preparing accurate rather than novel names. However, analysis of the number of fixations on patients prior to reference onset showed no significant difference between the two conditions and numerically slightly more fixations before novel words ($M = 2.42$ fixations) than accurate names ($M = 2.28$), $t\text{s} < 1.0$.

Altogether, these analyses suggest that speakers gazed at patients for the time needed to prepare a referring expression—both when naming them accurately and when calling them *blick*. This result challenges the idea that speech-related gazes to objects is solely mediated by associations between objects' visual properties and the words used to refer to the objects.

Discussion

The current study has shown that speakers gaze at objects while preparing words to refer to them, regardless of the relationship between the labels and the objects. Before intentionally calling a horse a *dog*, speakers gazed at the horse for far more time than when calling it a *horse*. Speakers gazed at the referent immediately before labeling it even when using the novel word *blick*, which lacked any association with the object to be described and lacked a history of use for directing a listener to an object. Thus, the gazes that anticipate reference to an object do not require that the meaning of the uttered word correspond to the visual features of the viewed object.

One potential concern with these results may be that the object referred to was the last one mentioned in each sentence. Earlier studies have found that speakers gaze longer at the same object when it is named last in an utterance rather than earlier (e.g., Griffin, 2001). Although this tendency for gazes to linger on last-named objects probably inflated gaze times, this tendency cannot account for the important effects observed. For example, the results for speakers using similar-concept names have been replicated in another experiment (Griffin & Oppenheimer, 2003). Speakers were asked to name agents (which were named first in all sentences) accurately or with the name of a similar concept. For example, a speaker might say “a dog is kicking a cow” instead of “a horse is kicking a cow.” Speakers gazed longer at objects before labeling them with similar-concept names than naming them accurately, even when another object was always named subsequently in the same sentence. Therefore, the effect of generating a similar-concept name holds for agents and patients as well as pre- and postverbal nouns, suggesting that they are specific to referring to the objects rather than processing particular thematic or grammatical roles, encoding actions, or generating verbs.

Why did speakers spend so much more time gazing at objects before labeling them with names of similar concepts than when naming them accurately (e.g., calling a horse a *dog*)? It could be that viewing the visual form of the object interfered with retrieving the name of a similar concept, as semantically similar words interfere in the picture–word interference paradigm (e.g., Glaser & Glaser, 1989; Lupker, 1979). However, there is a different reason that is consistent with the facilitatory effects of semantically related pictures on word production (Bloem & La Heij, 2003; Damian & Bowers, 2003) as well as codability effects in naming (e.g., Griffin, 2001; Lachman, 1973). Specifically, there are far more similar concepts and names to choose among when naming an object inaccurately than accurately. A lack of constraint and a wealth of choices should result in slower word preparation times and therefore longer gaze times. This is also consistent with the observation that speakers quickly and fluently prepared a prespecified but inaccurate label *blick* to the same referents, with gazes almost indistinguishable from those before accurate labels. Therefore, the use of an inaccurate label does not inevitably slow preparation time, decrease fluency, or prolong gazes.

Although the content and referential hypotheses are contrasted in the current experiments, they are not necessarily mutually exclusive. The current experiments indicate that the content hypothesis is not sufficient to account for all speech-related gazes because such gazes were made when visual content did not correspond to lexical content. To establish whether a version of the content hypothesis is necessary in addition to the referential hypothesis, one must test situations in which the visual information that speakers may gaze at is semantically but not referentially related to the content of speech. An example would be whether a speaker talking about a car that was not present directed his or her gaze to a car that was visible. In such a situation, the referential hypothesis would predict no gaze to the visible car because it would not be referred to by the speaker. However, the content hypothesis would predict speech-related gazes in these situations because of the correspondence between the visual features of a visible car and the intention to talk about a car or cars. Although the current experiments indicate that the intention to refer (probably mediated by spatial indices) can direct speakers' eye movements, future studies may show that matches between visual and conceptual content independent of referential intentions also can direct eye movements.

Conclusion

In the reported experiments, we compared two accounts of the tight linkage between word preparation and gaze fixation in scene descriptions. The content hypothesis suggests that speakers gaze at objects because of a correspondence between their visual forms, conceptual representations, and lexical entries (Humphreys et al., 1997; Roelofs, 1992; Van Der Meulen, 2001), whereas the referential account proposes that speakers' eye gaze is a reflection of visuospatial attention to locations associated with current contents of thought (Griffin, 2004b; Spivey et al., 2004). By asking participants to inaccurately label objects, it was possible to derive contrasting predictions from these two hypotheses. The results support the referential hypothesis, which predicts that speakers gaze at what they refer to regardless of the accuracy of the referring expression (see also Griffin, 2004a). These results suggest that the intent to refer to an object in a particular location can direct speakers' gazes while they prepare a name rather than an association between the visual features of the object and the sense of the word to be prepared.

References

Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). New York: Psychology Press.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62,* 875–890.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29,* 832–843.

Bansal, P., Griffin, Z. M., & Spieler, D. H. (2001). *Epd: Matlab Wave File Parsing Software.* [On-line]. Retrieved July 15, 2001, from http://oak.psych.gatech.edu/~spieler/software.html

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14,* 143–177.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language, 48,* 468–488.

Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language, 48,* 653–685.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12,* 335–359.

CMU Sphinx Group. (2001). Sphinx-2 [Computer software]. Retrieved November 2, 2001, from http://sourceforge.net/projects/cmusphinx/

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture–word interference tasks. *Psychonomic Bulletin and Review, 10,* 111–117.

Frege, G. (1997). On sense and reference. In P. Ludlow (Ed.), *Readings in* *the philosophy of language* (pp. 563–583). Cambridge, MA: MIT Press. (Original work published 1948)

Georgia Institute of Technology/Nexidia. (2004). Fasttalk [Computer software]. Atlanta, GA: Author.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Glenberg, A. M., Schroeder, J. L., & Robertson, D. A. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition, 26,* 651–658.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82,* B1–B14.

Griffin, Z. M. (2004a). The eyes are right when the mouth is wrong. *Psychological Science, 15,* 814–821.

Griffin, Z. M. (2004b). Why look? Reasons for eye movements related to language production. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 213–247). New York: Psychology Press.

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

Griffin, Z. M., & Oppenheimer, D. M. (2003, August). *Looking and lying: Speakers' gazes reflect locus of attention rather than speech content.* Paper presented at the 12th European Conference on Eye Movements, Dundee, Scotland.

Harley, T. A. (1990). Environmental contamination of normal speech. *Applied Psycholinguistics, 11,* 45–72.

Henderson, J., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology Press.

Hommel, B. (2002). Responding to object files: Automatic integration of spatial information revealed by stimulus–response compatibility effects. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55,* 567–580.

Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral & Brain Sciences, 24,* 453–509.

Humphreys, G. W., Riddoch, M. J., & Price, C. J. (1997). Top–down processes in object identification: Evidence from experimental psychology, neuropsychology, and functional anatomy. *Philosophical Transactions: Biological Sciences, 352,* 1275–1282.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 105–134). New York: Psychology Press.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8,* 441–480.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology, 99,* 199–208.

La Heij, W., Heikoop, K. W., Akerboom, S., & Bloem, I. (2003). Picture naming in picture context: Semantic interference or semantic facilitation? *Psychology Science, 45,* 49–62.

La Heij, W., Van Der Heijden, A. H. C., & Plooij, P. (2001). A paradoxical exposure-duration effect in the Stroop task: Temporal segregation between stimulus attributes facilitates selection. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 622–632.

Lupker, S. J. (1979). The semantic nature of response competition in the picture–word interference task. *Memory & Cognition, 7,* 485–495.

Lupker, S. J. (1988). Picture naming: An investigation of the nature of categorical priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 444–455.

Meyer, A. S., Sleiderink, A., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25–B33.

Meyer, A. S., & Van Der Meulen, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review, 7,* 314–319.

Morgan, J. L., & Meyer, A. S. (2005). Processing of extrafoveal objects during multiple-object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 428–442.

Pylyshyn, Z. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80,* 127–158.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29,* 1045–1060.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42,* 107–142.

Spivey, M. J., Richardson, D. C., & Fitneva, S. A. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 161–190). New York: Psychology Press.

Van Der Meulen, F. F. (2001). *Moving eyes and naming objects.* Unpublished doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen, the Netherlands.

Van Der Meulen, F. F., Meyer, A. S., & Levelt, W. J. M. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition, 29,* 512–521.

Received September 8, 2004
Revision received January 18, 2006
Accepted February 6, 2006 ■