Research Article

The Eyes Are Right When the Mouth Is Wrong

Zenzi M. Griffin

Georgia Institute of Technology

ABSTRACT—***When describing visual scenes, speakers typically gaze at objects while preparing their names. In a study of the relation between eye movements and speech, a corpus of self-corrected speech errors was analyzed. If errors result from rushed word preparation, insufficient visual information, or failure to check prepared names against objects, speakers should spend less time gazing at referents before uttering errors than before uttering correct names. Counter to predictions, gazes to referents before errors (e.g., gazes to an axe before saying "ham–" [hammer]) highly resembled gazes to referents before correct names (e.g., gazes to an axe before saying "axe"). However, speakers gazed at referents for more time after initiating erroneous compared with correct names, apparently while they prepared corrections. Assuming that gaze nonetheless reflects word preparation, errors were not associated with insufficient preparation. Nor were errors systematically associated with decreased inspection of objects. Like gesture, gaze may accurately reflect a speaker's intentions even when the accompanying speech does not.***

People usually look at the things they think about. Often they talk about the things they think about, too. Despite their many differences, gaze and speech both provide a wealth of information about thought content. When people describe visual scenes, their gazes and their words are synchronized. Words begin roughly a second after speakers gaze at their referents (Griffin & Bock, 2000). These word-related gazes are correlated with the difficulty of word preparation. For example, speakers gaze for more time at objects that have multiple context-appropriate names (Griffin, 2001) or uncommon names (Meyer, Sleiderink, & Levelt, 1998) than at objects that have a single name or a common name. Priming the sounds of objects' names or repeating names reduces the time speakers spend gazing at the objects (Meyer & Van Der Meulen, 2000; Van Der Meulen, Meyer, & Levelt, 2001). Although speakers look at objects while preparing their names, they typically do not look at them while articulating their names. Instead, while articulating the name of one object, speakers gaze at the next object to be mentioned (Meyer et al., 1998; for review, see Griffin, 2004, and Meyer & Lethaus, 2004).

Speech errors reflect situations in which speakers are unsuccessful in expressing their intentions in words. Errors come in a variety of shapes and sizes, and many factors conspire to make speech errors more or less likely (see Dell, 1995). For example, faster speech rates are associated with higher error rates (e.g., Oomen & Postma, 2001b). Some models of word production capture this relation by allowing less time for representations of the intended word to accrue activation as speech rate increases (Dell, 1986; MacKay, 1982). As gazes reflect cognitive processes and word preparation, the eye movements that precede speech errors may provide clues about the origins of the errors. Specifically, errors due to too little preparation time should be associated with decreased time gazing at referents.

Some speech errors appear to be due to speakers gazing at unintended objects in the visual environment. Harley's (1990) corpus includes an example in which a speaker attempted to say, "My mind's gone," while looking at a college bill, and instead produced, "My bill's gone." Perhaps to reduce the amount of unrelated visual information they perceive during the most effortful aspects of formulating speech, speakers often avert their gaze from an interlocutor at the start of an utterance and during disfluencies (e.g., Exline & Fehr, 1978; Glenberg, Schroeder, & Robertson, 1998). However, when producing single words, unimpaired speakers seem immune to interference from viewing related or unrelated objects (Bloem & La Heij, 2003; Damian & Bowers, 2003). In contrast, one speaker with brain damage required undegraded perceptual information to successfully retrieve an object's name even when semantic information was accessible (Humphreys, Price, & Riddoch, 2000). Such observations suggest that speech errors may be

Address correspondence to Zenzi M. Griffin, 654 Cherry St., School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332-0170; e-mail: zenzi.griffin@psych.gatech.edu.

associated with gazing too briefly at intended objects or too long at other objects while preparing speech.

Gazes to objects before producing their names may reflect more than word preparation. Speakers may continue to gaze at an object after preparing its name, while checking whether the prepared name fits the object (Van Der Meulen, 2001). Indeed, speakers can edit their prepared words prior to speaking in order to avoid speech errors or embarrassing utterances (Levelt, 1983; Motley, Camden, & Baars, 1982). All else being equal, the time spent gazing at an object while preparing and checking its name should be longer than the time spent when a name check is omitted. In other words, speakers may gaze at referents for more time prior to uttering correct names that have been checked than prior to uttering erroneous names that have not been checked.

In the present study, I examined cases in which speech errors were produced during eye-tracking experiments. The primary question was whether speakers' gazes prior to errors were distinct from their gazes when correct words were produced. If errors result from rushed word preparation, insufficient visual information, or failure to check prepared names against objects, speakers should spend less time gazing at referents before uttering errors than before uttering correct names.

DATA SELECTION

To test this prediction, I identified 41 full or partial speech errors uttered by 33 participants, including university students (18–30 years old) and older adults (60–80 years old), during eye-tracking experiments (reported in Griffin, 2003a, 2003b; Griffin & Spieler, 2004; and Spieler & Griffin, in press). In these experiments, speakers produced utterances naming two or three unrelated line-drawn objects or describing actions in line-drawn scenes. Typically, every object displayed was meant to be named, so environmental intrusions from other objects were unlikely and not observed in this set of errors.

Identifying speech errors requires knowing that speakers intended to say something other than what they actually said. Speakers sometimes offered multiple names for an object in a way that could reflect uncertainty rather than a correction (e.g., "th:ee . . belt, or wristband" and "ma:ze labyrinth").[1] To minimize inferences about speakers' intentions, I limited the corpus to cases in which speakers either signaled the error lexically (e.g., "oops" or "no") or interrupted their own speech. Speakers are more likely to interrupt themselves in midword when they produce erroneous words rather than merely suboptimal ones (Brédart, 1991; Levelt, 1983). It seems reasonable to infer that the speakers considered incomplete words to be errors. Repeated word-initial sounds (e.g., "s–s–spider" and "a ma– a man") could be due to difficulty retrieving sounds rather than error detection (e.g., Oomen & Postma, 2001a), so such cases were not included. Therefore, errors that did not form complete words were restricted to cases in which speakers resumed speaking with different word-initial sounds (e.g., "ham– axe" when intending to name an axe). Errors that occurred last in a list of unrelated objects were excluded because gazes tend to dwell on such objects (Griffin, 2001; Van Der Meulen, 2001). The 41 utterances that met the criteria for inclusion are listed in Table 1. Thirty-three of the errors occurred for the first object mentioned in an utterance.

Eye movement measures associated with speech errors were compared with those associated with correct names (or at least, names that seemed appropriate for the objects and that were not corrected by the speakers). As in other psycholinguistic studies, analyses based on participant and item variability were performed (see Clark, 1973). In the analysis based on participant variability, measures for correct names came from other trials in which the same speakers named objects of similar naming difficulty[2] (total of 1,692 trials). With four exceptions,[3] these *same-speaker* data came from the first objects named in utterances. Twenty-five percent of the erroneous names were preceded by disfluencies such as filled pauses, silent pauses, or prolonged words. Speakers in similar studies spent more time gazing at objects that were named disfluently than at objects named fluently (Griffin, 2001). To avoid confounding disfluencies with overt errors, I calculated mean eye movement measures separately for fluent and disfluent correct names. When a participant uttered a speech error disfluently, its measures were compared with the mean measures for disfluent correct names. Likewise, fluent errors were compared with fluent correct names, and a combination of fluent and disfluent errors was compared with a weighted mean for correct names.[4]

When members of different age groups erred in naming the same object, the object was treated as two different items (the object named by college-aged speakers and the object named by older speakers). By this definition, errors were distributed over 40 objects. Measures for these objects were compared with measures taken when other college-aged or elderly speakers in the same experiment produced correct names for the same objects (280 trials total). Errors produced during scene descriptions ($n = 16$) were matched to correct names for the same objects regardless of utterance position. In all other cases, correct names in the *same-object* data matched errors in utterance position and intended object. Correct names for these

[1]Periods indicate pauses, commas intonation breaks, dashes halted speech, and colons prolonged sounds.

[2]In experiments with manipulations of object naming difficulty, correct names came from objects in the same condition as the one that elicited the speech error.

[3]Four errors were made in naming the second of three objects in a fixed order. For these speakers, second named objects of the same item type were used for comparison. The other four errors that were not utterance initial occurred for objects in scene descriptions.

[4]For 2 participants, mean measures used all similar items without weighting by fluency because of the small number of similar items.

TABLE 1
Transcriptions of the 41 Utterances Containing Errors

Source	Speaker	Utterance	Intended object
G2003a1	Y1	Shove:l–er rake barrel	rake
G2003a1	Y2	Booma– . . . bait . . . what is that thing . . . chain	slingshot
G2003a1	Y3	Ham– . . axe? . . juice mixer	axe
G2003a1	Y4	Sh– . . spoon	spoon
G2003a1	Y5	S–s–rake	rake
G2003a1	Y6	Ss– . . ha–axe . . . blender	axe
G2003a1	Y7	St– . . scr::ewdriver . . . rope	screwdriver
G2003a1	Y8	Rope pipe . . oops, scarf pipe	scarf
G2003a2	Y9	F– . uh tie hammock	tie
G2003a2	Y10	Ku– octopus knife	octopus
G2003a2	Y10	Shtrt safe next to fish	safe
G2003a2	Y10	Brush ne–, broom next to can	broom
G2003a2	Y11	Sss– raccoon balloon	raccoon
G2003b	Y12	Wind– . . . whatever something that does wind and carrot	windmill
G2003b	Y13	Uh lawn chair oh that's a sled, sled turtle	sled
G2003b	Y14	Win– win- um . . . helicopter bomb	helicopter
G2003b	Y15	Mowzun . maze	maze
G2003b	Y16	Frisur– refrigerator barrel	refrigerator
G2003b	Y17	Spider . . . . that's a that's an ant actually . . . peanut	spider
GS2004	O1	The gir– the woman i:s . offering a book to a boy	librarian
GS2004	O2	Looks like a la– avalanche is overcoming some skiers	avalanche
GS2004	O3	Looks like a: sh– . . man-eating shark's about ready to go after a swimmer	shark
GS2004	O4	The raid– uh . the roadie is accepting uh . guitar from a young man	roadie
GS2004	O4	[laughs] hehe a sad bu– a sad man is accepting a ticket from a motor . from ay uh . . . uh . policeman . motorcycle policeman	driver
GS2004	O5	Two mer– . two men are carrying . an individual . . . in a vehicle which I don't know the name of	servants
GS2004	O6	U:h sss– person in the water, i:s u::h about to be devoured by a great white shark	swimmer
GS2004	O7	U::h . ij– . . businessman is . turned around . . . to a waitress who's offering him a drink on a tray	man
GS2004	O8	A l– . clock alarm going off is waking the man	alarm clock
GS2004	O9	The dress is being shown to the f– man at the counter	man
GS2004	O9	The uh . p– p– p– water pitcher is being delivered from one person to another at the table	pitcher
GS2004	Y18	The rezi– . the musician is being given a guitar by a young man	musician
GS2004	Y19	A little girl, is . showing anoth–, uh boy, a piece uh paper	boy
GS2004	Y20	The lawy:er i:s telling the judge . . what should b:e . his id– . his evidence for his client	gun
GS2004	Y21	U:m. a:y girl is hand:ing ay paintbrush t:o uh . gu–, boy on a ladder wh:o's painting the walls	boy
GS2004	Y22	The arm, alarm clock is waking the boy	alarm clock
SG2004	O10	The . cro:wn and th::e . berry pi . . . branch are above the ladder	berries
SG2004	O11	The crib and u:h spoon or roll or no the doorknob are above the . . cane?	doorknob
SG2004	O12	Thee um . scotch tape and the window sil–, window are above the ladder	window
SG2004	O12	Thee, van and the hor– and the donkey are above the ruler	donkey
SG2004	O13	The raid– the speaker and thuh, ruler are above the table	speaker
SG2004	O14	The z– . u:h blimp and th:ee stairs are above thee hat pin	blimp

Note. Erroneous words are underlined. Periods indicate pauses, commas intonation breaks, dashes halted speech, and colons prolonged sounds. G2003a1 = Experiment 1 of Griffin (2003a), G2003a2 = Experiment 2 of Griffin (2003a), G2003b = follow-up experiment described in Griffin (2003b), GS2004 = Griffin & Spieler (2004), SG2004 = Spieler & Griffin (in press). Speakers are individuated with codes starting with "Y" for young adults and "O" for older adults, followed by numbers.

objects were no more fluent than the errors, $t < 1$. Therefore, any differences between error and same-object data cannot be attributed to differences in stimulus properties, such as object complexity, image quality, name agreement, or the frequency of the dominant name, or to contextual differences in fluency or utterance position.

DATA TREATMENT

The original eye movement data from each experiment were used. Intended objects were identified on the basis of corrected names and task instructions. In all experiments, fixation coordinates were automatically categorized with respect to

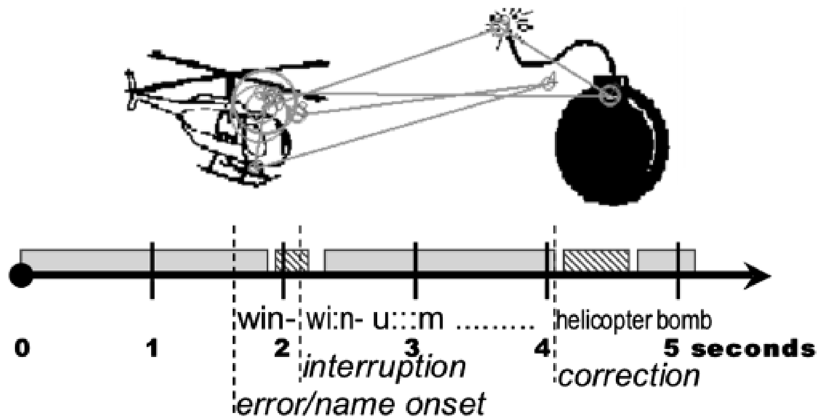

Fig. 1. Example of eye movements and speech from a trial in the corpus of errors. Periods indicate pauses, dashes halted speech, and colons prolonged sounds. The gray circles superimposed on the objects represent fixation locations, with greater diameters indicating longer durations. Gray lines connect successive fixations. The timeline shows the duration of gazes on the helicopter (solid gray bars) and the bomb (striped bars) and when they occurred relative to speech.

experimenter-defined polygons that surrounded objects with a 1° to 2° margin of error. Gazes were defined as beginning at the onset of a fixation within a polygon and ending with the offset of the last fixation within it. The time spent gazing at an object prior to producing a name was calculated by summing the gaze durations on the object prior to the onset of the name. If a gaze spanned the onset of a name, only the time prior to the name's onset was included in this measure. For example, in the timeline of Figure 1, the first gaze on the helicopter continued after the onset of the error, but only the duration to the left of the dashed error line constituted the gaze time before the error's onset. The remainder of the gaze (i.e., the portion to the right of the dashed line) contributed to the time spent gazing after the onset of the error. Error onsets, correct-name onsets, and interruption points were measured with sound-editing software. Two-sided paired t tests comparing errors with the same-speaker (t_1) and same-object (t_2) sets are reported for each measure.

RESULTS

Mean values for all measures, based on means for speakers and items, are listed in Table 2. Figure 2 plots the proportion of trials on which speakers gazed at the intended object for 8-ms intervals relative to the onset of the erroneous and correct names. Movies illustrating the eye movements and speech in this sample are available on the Web at http://oak.psych.gatech.edu/~zgriffin/.

Before Name Onset

Speakers' gazes on intended objects before uttering erroneous and correct names were very similar. Speakers gazed at the intended object at least once before every error in the corpus. For every error but one, speakers gazed at the intended object within 1 s of the error's onset. There was no difference in mean number of gazes to intended objects before the onset of errors and correct names, ts < 1.3, n.s.

The mean times spent gazing at objects prior to naming them correctly and naming them erroneously were compared in order to evaluate the hypothesis that speakers spent less time gazing at referents when they produced erroneous names than when they produced correct names. Unexpectedly, speakers gazed at intended objects for 30% more time before errors than before correct names, $t_1(32) = 2.75$, $p < .01$. However, the same-object analysis showed a nonsignificant -11-ms difference, $t_2 < 1$; in other words, the mean gaze time on the objects before errors was 99.1% of the mean gaze time before correct names. This discrepancy between the speaker and item analyses suggests that the objects that elicited errors were time-consuming to name in general: All speakers spent a long time gazing at these particular objects, and speakers who produced errors spent more time gazing at them than at other similar objects.

When speakers gaze at an object more than once, the last gaze before referring to it is usually longest and appears to be the one that is most related to word preparation (Griffin & Bock, 2000). Eye-voice spans were calculated as the time from the onset of the last (or only) gaze made to an object prior to production of its name and the onset of the name. (For the example trial in Fig. 1, this period extends from the start of the trial, 0 ms, to the dashed line labeled "error/name onset.") Eye-voice spans did not differ significantly between errors and correct names, lasting about 1 s in both cases, ts < 1.3.

Typically, about 100 to 300 ms before saying an object's name, speakers shift their gaze to the next object to be named (Griffin, 2001; Meyer et al., 1998). It has been argued that this shift coincides with the end of phonological encoding (Meyer et al., 1998; Meyer & Van Der Meulen, 2000). Alternatively, the shift may coincide with the completion of a name check (Van

TABLE 2

Means and 95% Confidence Interval Half-Widths (CI) for Gaze Measures When Objects Were Named Erroneously and Correctly

Measure	Same-speaker data				Same-object data			
	Error	Correct word	Difference	CI	Error	Correct word	Difference	CI
Before word onset								
Number of gazes	1.554	1.337	0.217	0.345	1.600	1.558	0.042	0.249
Gaze time (ms)	1,272	982	290*	215	1,198	1,209	−11	219
Onset of word − onset of last gaze (ms)	1,154	1,045	109	188.5	1,096	1,220	−124	199.5
Offset of last gaze − onset of word (ms)	203	−101	304*	301	211	−73	284†	291
After word onset								
Gaze time (ms)	1,463	638	825*	464	1,367	697	670*	384

†.10 < p < .05. *p < .05.

Der Meulen, 2001). In the present sample, when speakers correctly named the objects, they shifted their gaze away from them about 100 ms before beginning to utter their names. However, when speakers produced errors, they tended to continue gazing at the objects while beginning to utter the erroneous names. Figure 1 illustrates an example of this pattern in that the speaker continued gazing at the helicopter after the onset of "win–." This difference in when gazes ended relative to name onsets was significant in the analysis by speakers and marginal in the analysis by items, $t_1(32) = 2.06$, $p < .05$, and $t_2(39) = 1.97$, $p < .06$.

In sum, speakers' gazes at objects were similar prior to correct and erroneous names. The data offer no support for the prediction that compared with correct names, speech errors are associated with less time gazing at intended objects before naming them. Instead, insofar as there were any differences, speakers spent more time gazing at objects before naming them with erroneous words than before naming them correctly.

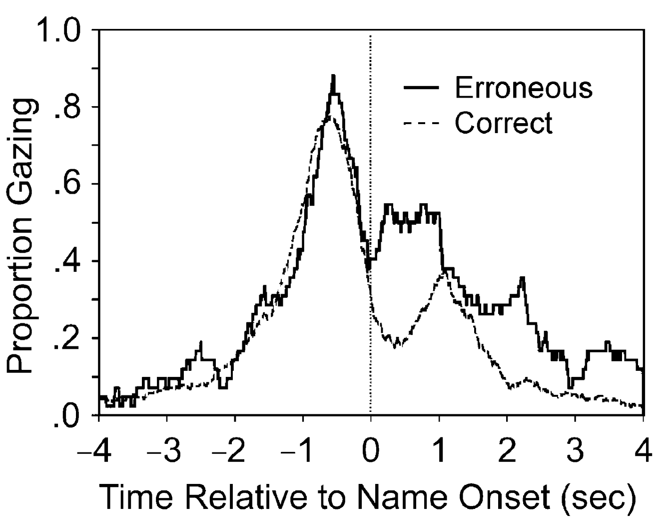

Fig. 2. Proportion of trials on which speakers gazed at intended objects plotted for 8-ms intervals relative to the onset of names that were erroneous (e.g., uttering ''win–'' when the intended object was a helicopter) and correct (uttering ''helicopter''). Correct naming trials come from the same-object set.

After Name Onset

Speakers often interrupted themselves in the middle of articulating erroneous names and produced corrections soon after. In such cases, the interruptions occurred a mean 519 (range: 93–2,608) ms after the onset of erroneous words. Corrections were uttered 544 (40–2,439) ms later. About half of the interruptions occurred while speakers gazed at intended objects. Levelt (1983) posited that speakers interrupt their speech as soon as an error is detected. Subsequent researchers have argued that this cannot consistently be the case, because new speech often begins within milliseconds of interruptions (Blackmer & Mitton, 1991). These researchers claim that speakers must begin planning corrections before the interruption (or in parallel with it; Hartsuiker & Kolk, 2001). Gazing at an object while uttering an error might help a speaker detect the error. If so, one would expect shorter latencies from error onset to speech interruption when speakers continued to gaze at an object after the onset of an error than when they looked away before the error. Instead, there was a numerical trend in the opposite direction. Specifically, interruptions occurred a mean 466 ms after error onset for the 24 items that speakers looked away from before articulating errors and a mean 546 ms after error onset for the 14 items that speakers continued gazing at (2 items lacked clear interruption points). The difference did not approach significance in an unpaired t test, $t_2 < 1$. The longer that a gaze overlapped with articulation of an error, the longer the latency between the error and the interruption, $r(n = 38) = .31$, $p < .06$. Thus, gazing at a referent at error onset was not associated with a speedy interruption.

Speakers gazed at objects for twice as long after the onsets of errors than after the onsets of correct names, $t_1(32) = 3.62$ and $t_2(39) = 3.53$, $ps < .002$. The additional time spent gazing at objects after the onset of errors seems associated with preparing corrections. Figure 3 plots the proportion of trials on

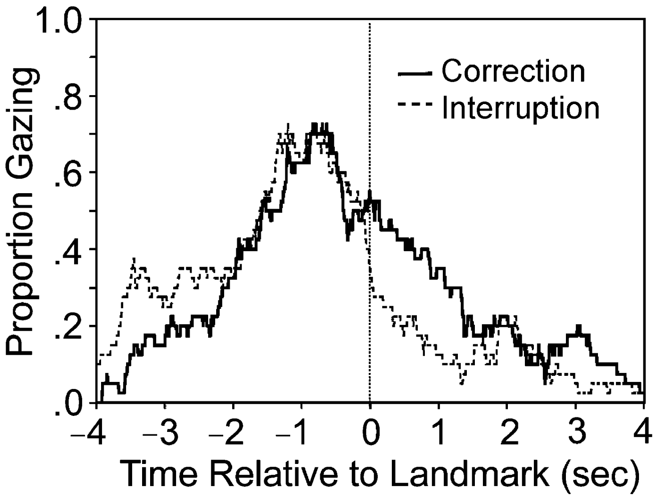

Fig. 3. Proportion of trials on which speakers gazed at intended objects plotted for 8-ms intervals relative to interruption points (e.g., the end of ''win–'') and the onset of corrections (e.g., ''helicopter'').

which speakers gazed at intended objects as a function of time, relative to interruption points and the onsets of corrected words. Across items, the time spent gazing at referents after error onset was positively correlated with the latency from interruption to correction, $r(n = 36) = .50, p < .005$. The time spent gazing at objects after the onsets of corrections, 555 ms, did not differ significantly from the time that the same speakers spent gazing at objects after initially correct words, $t_1(30) < 1$, but was marginally less than the 706 ms that other speakers spent gazing at these objects after correct names, $t_2(37) = 1.81, p < .08$.

DISCUSSION

There were several reasons to predict that erroneous names, compared with correct names, would be associated with less time gazing at referents before naming them. For example, the continued availability of perceptual information during gazes reinforces the selection and retrieval of correct object names in many models of naming (e.g., Humphreys & Forde, 2001; Humphreys, Riddoch, & Price, 1997; Roelofs, 1992; Van Der Meulen, 2001). The time spent gazing at an object before naming it appears highly related to preparing its name (see Griffin, 2004; Meyer & Lethaus, 2004). So, if the speech errors in this corpus were due to hurried or insufficient word preparation, as in Dell's (1986) and MacKay's (1982) models, gaze times before errors should have been shorter than gaze times before correct names. To the extent that gaze times also include checking the appropriateness of a prepared name against the image of the object (Van Der Meulen, 2001), gaze times should also have been shorter before errors than before correct names. Counter to such predictions, speakers spent as much time gazing at objects before erroneous names as before correct ones. The results suggest that the speech errors people make while describing visual information are not systematically due to shorter gazes on referents or, by extension, to shorter word-preparation times.

The results have three immediate implications. First, they support the common assumption (e.g., Dell, 1995; Ferreira & Griffin, 2003) that aside from environmental intrusions (Harley, 1990), word-substitution errors reflect problems stemming from linguistic processes rather than extralinguistic ones like visual attention. Similarly, studying a different modality, McNeill (1985) provided examples in which a speaker's gesture conveyed correct information while the accompanying speech was in error.[5] These cases constitute objective data to support the subjective sense that speakers know what they wish to say at a conceptual or message level even when the words that they say are wrong. Second, the results are inconsistent with the idea that, as a rule, speech errors involve omitting or rushing through word-production processes.[6] Third, the results suggest that gazing at referents while preparing their names does nothing to ensure that the prepared names are correct. Similarly, speakers in another study were at least as fast, fluent, and accurate in telling time from clocks displayed for 100 ms as from clocks that remained in view throughout their utterances (Bock, Irwin, Davidson, & Levelt, 2003). Along with a number of other experiments and observations, such results suggest that when speakers refer to objects that are visually present, they tend to gaze at them although such gazes may be epiphenomenal (see Griffin, 2004, for detailed discussion).

Epiphenomenal or not, speakers' gazes are informative to listeners (e.g., Baldwin, 1993). Speakers' gazes have even been incorporated into a speech recognition system to aid in disambiguating references in human-computer interactions (Campana et al., 2001). The present results indicate that when speech errors occur, knowing which object a speaker gazed at may be more informative about the speaker's intentions than knowing the words the speaker said. Further, knowing both the meaning of a speaker's words and the object of the speaker's gaze may permit automatic detection of speech errors that the speaker does not notice. The errors that speakers are most likely to leave uncorrected are word substitutions (e.g., "hammer" for "axe"), which occur relatively rarely, at a rate of about 2.5 per 10,000 words (Deese, 1984). Even so, a speaker could easily produce one or two uncorrected substitutions every day,[7] which could cause serious misunderstandings for a system that considered speech input only.

Of course, one should be cautious in drawing generalizations from these results because the data include only errors that speakers made while describing line-drawn objects in

[5]Note that gesture-speech mismatches also occur during cognitive change or learning (Goldin-Meadow, Alibali, & Church, 1993) and are not always speech errors.

[6]In fact, increases in speech rate may be accomplished without reducing the time taken to prepare words (see Griffin, 2003a, 2003b).

[7]For instance, Wagner (1985) reported that 9-year-old children produced about 20,000 words a day. Estimates for younger and older children were similar.

eye-tracking experiments. In addition, analyses included only errors that speakers detected, and these may differ in important respects from the errors that go undetected (see Deese, 1984). Also, because many of the erroneous words were interrupted, it was impossible to determine what types of errors many of them were. Moreover, many of the implications rest on the assumption that gaze times on referents before their names are produced are good reflections of the time speakers spend preparing the names. How valid this assumption is remains to be seen. With the increasing numbers of researchers using eye tracking to study language production, the opportunity to collect spontaneous errors will permit more specific error types to be examined (e.g., Van Der Meulen, 2003). Despite these limitations, however, the present results provide strong evidence against the hypothesis that speech errors are systematically associated with decreased gaze on referents prior to saying their names.

Acknowledgments—The author thanks Vic Ferreira, Kristin Garton, Chris Herzog, Bob Slevc, and Daniel Spieler for comments and advice, and Kay Bock for suggesting studying language production with eye tracking. This work was supported by a grant from the National Institute of Mental Health, R03 MH61318-01.

REFERENCES

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

Blackmer, E.R., & Mitton, J.L. (1991). Theories of monitoring and the timing of repairs in spontaneous speech. *Cognition*, *39*, 173–194.

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.

Brédart, S. (1991). Word interruption in self-repairing. *Journal of Psycholinguistic Research*, *20*, 123–138.

Campana, E., Baldridge, J., Dowding, J., Hockey, B.A., Remington, R.W., & Stone, L.S. (2001, November). *Using eye movements to determine reference in a spoken dialogue system*. Poster presented at Workshop on Perceptive User Interfaces, Orlando, FL.

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.

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

Deese, J. (1984). *Thought into speech: The psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*, 283–321.

Dell, G.S. (1995). Speaking and misspeaking. In L.R. Gleitman & M. Liberman (Eds.), *An invitation to cognitive science: Vol. 1. Language* (2nd ed., pp. 183–208). Cambridge, MA: MIT Press.

Exline, R.V., & Fehr, B.J. (1978). Applications of semiosis to the study of visual interaction. In A.W. Seigman & S. Feldstein (Eds.), *Nonverbal behavior and communication* (pp. 117–157). Hillsdale, NJ: Erlbaum.

Ferreira, V.S., & Griffin, Z.M. (2003). Phonological influences on lexical (mis)selection. *Psychological Science*, *14*, 86–90.

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

Goldin-Meadow, S., Alibali, M.W., & Church, R.B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review*, *100*, 279–297.

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

Griffin, Z.M. (2003a). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, *10*, 603–609.

Griffin, Z.M. (2003b, March). *Speaking sooner or later: Availability, word length, and fluency*. Paper presented at the 16th Annual CUNY Conference on Human Sentence Processing, Cambridge, MA.

Griffin, Z.M. (2004). Why look? Reasons for eye movements related to language production. In J.M. 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., & Spieler, D.H. (2004, April). *Sentence complexity in event descriptions of younger and older speakers*. Poster presented at the 10th Cognitive Aging Conference, Atlanta, GA.

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

Hartsuiker, R.J., & Kolk, H.H.J. (2001). Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology*, *42*, 113–157.

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., Price, C.J., & Riddoch, M.J. (2000). On the naming of objects: Evidence from cognitive neuroscience. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 143–163). Philadelphia: Psychology Press.

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 of the Royal Society of London: Biological Sciences*, *352*, 1275–1282.

Levelt, W.J.M. (1983). Monitoring and self-repair in speech. *Cognition*, *14*, 41–104.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*, 483–506.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*, 350–371.

Meyer, A.S., & Lethaus, F. (2004). The use of eye tracking in studies of sentence generation. In J.M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 191–211). New York: Psychology Press.

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 & Review*, *7*, 314–319.

Motley, M.T., Camden, C.T., & Baars, B.J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *21*, 578–594.

Oomen, C.C.E., & Postma, A. (2001a). Effects of divided attention on the production of filled pauses and repetitions. *Journal of Speech, Language, & Hearing Research*, *44*, 997–1004.

Oomen, C.C.E., & Postma, A. (2001b). Effects of time pressure on mechanisms of speech production and self-monitoring. *Journal of Psycholinguistic Research*, *30*, 163–184.

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

Spieler, D.H., & Griffin, Z.M. (in press). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes*.

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. (2003, August). *Stability of speech-to-gaze alignment in multiple object naming: Beautiful or boring?* Paper presented at the 12th Annual European Conference on Eye Movements, Dundee, Scotland.

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.

Wagner, K.R. (1985). How much do children say in a day? *Journal of Child Language*, *12*, 475–487.

(Received 7/24/03; Revision accepted 3/4/04)