A reversed word length effect in coordinating the preparation and articulation of words in speaking

ZENZI M. GRIFFIN
Georgia Institute of Technology, Atlanta, Georgia
and Stanford University, Stanford, California

Speakers tend to prepare their nouns immediately before saying them, rather than preparing them further in advance. To test the limits of this last-second preparation, speakers were asked to name object pairs without pausing between names. There was not enough time to prepare the second name while articulating the first, so the speakers' delay in starting to say the first name was based on the amount of time available to prepare the second name during speech. Before speaking, they spent more time preparing a second name (e.g., *carrot*) when the first name was monosyllabic (e.g., *wig*) rather than multisyllabic (e.g., *windmill*). When additional words intervened between names, the length of the first name became less important and speech began earlier. Preparation differences were reflected in speech latencies, durations, and eye movements. The results suggest that speakers are sensitive to the length of prepared words and the time needed for preparing subsequent words. They can use this information to increase fluency while minimizing word buffering.

Life would be easier if we had to worry about only one action at a time. But unfortunately, life is not that easy. Instead, people must often ensure that a second action follows a first within a reasonable time period. Success and failure in coordinating preparation and execution is particularly apparent in speech. As much as 40% to 50% of the amount of time a person spends in speaking is actually silent, with 45% of all pauses occurring where there are no grammatical junctures to motivate them (Goldman-Eisler, 1968). These pauses are associated with having a large number of appropriate words from which to choose. Such observations led researchers to suggest that fluent speech is the result of selecting all of the words in an utterance before speaking and retrieving only the words' sounds during speech. In this view, disfluencies occur when speakers try to decide the content of their speech or which words to use while speaking (see, e.g., Butterworth, 1989; Garrett, 1988; Goldman-Eisler, 1968).

This view of the timing of word preparation and execution has been challenged. A number of reaction time studies (e.g., Huitema, 1993; Kempen & Huijbers, 1983; Lindsley, 1975; Pechmann, 1989; Smith & Wheeldon, 1999)

suggest that even when they are fluent, speakers choose some words after they start speaking rather than beforehand. It is only in experiments in which speakers are compelled to prepare their speech in advance or to recall sentences (e.g., Ferreira, 1991; Griffin & Bock, 2000; Wheeldon & Lahiri, 1997) that evidence of buffering is observed.

Studies of speakers' eye movements before and during speech lend further support to the idea that people may speak fluently even when choosing their words shortly before saying them. Speakers take about 1 sec to begin naming an object presented in isolation (see, e.g., Snodgrass & Yuditsky, 1996), which suggests that selecting a name and retrieving its sounds take slightly less than 1 sec. When describing scenes, speakers gaze at referents during the second before naming them (Griffin & Bock, 2000), although it takes about a fraction of this time to identify them (Potter, 1975). The more appropriate names an object has, the longer speakers gaze at it (Griffin, 2001). These gazes in midutterance suggest that speakers do not just retrieve word sounds as they speak, but that they also decide which words to use.[1]

One challenge to any word-by-word account of word preparation in speech comes from the observation that most words in English and other languages take far less than a second to articulate, even in isolation. So if speakers needed 1 sec to prepare every word, they could not prepare and utter each word immediately without hesitating between words. However, all words are not equally difficult to prepare. Because of their high frequency and their role in structuring sentences, function words (e.g., *the*, *might*, *on*) are likely to take far less time to prepare than content words (e.g., *cow*, *lovely*, *walk*) do. Although func-

tion words may require little preparation time themselves, they add to the duration of speech, during which speakers may plan upcoming words. So, despite the relatively long preparation time of content words, speakers may often be able to prepare them one at a time while speaking, with either concurrent or serial and rapid preparation of function words (see, e.g., Kempen & Hoenkamp, 1987; Levelt, 1989).

But how do speakers decide when to begin preparing a content word? One possibility is that they prepare one phonological word at a time (e.g., Wheeldon & Lahiri, 1997) and simply bet on having enough time while articulating the current phonological word to prepare the next one. A phonological word has one stressed syllable and is composed of a content word and adjacent function words, as in *on a car*. Alternatively, speakers might time their preparation according to syntax, preparing one constituent (such as a noun phrase) at a time (see, e.g., Kempen & Hoenkamp, 1987; Smith & Wheeldon, 1999). When the articulation of such a unit provides enough time to prepare the next one, speech will be fluent. When there is not enough time for preparation, speech will be disfluent. However, speakers seem able to modulate their preparation with more precision than either of these strategies might suggest. But if speakers can estimate the time available during speech for preparing a word, what is the basis of such an estimate?

The experiments reported here test whether speakers can modulate their word preparation on the basis of word length (or another correlate of articulation duration). Speakers were asked to name two objects without pausing between their names, as in *wig carrot*. The number of syllables in the name of the first object was manipulated. Because single-syllable words such as *wig* take well under 500 msec to articulate, speakers do not have the necessary 800 msec to prepare a second name such as *carrot* during speech. If speakers begin to say *wig* as soon as they have prepared it, *carrot* will not be ready when needed and speakers will hesitate between nouns. That is, speakers cannot produce the first-object name as soon as it is prepared without sacrificing fluency. However, when the first object has a longer name such as *windmill*, speakers have more time to prepare the second name after speech begins. Can speakers take into consideration both the time afforded by the first word's length and the time typically needed to prepare a second object's name? If they can, a short first word such as *wig* should compel them to prepare much of the second word before beginning to speak. In contrast, a longer first word such as *windmill* would allow speakers to begin speaking earlier and to prepare the second word less before saying the first name.[2]

A simpler alternative would involve preparing both words completely before speaking, thereby avoiding disfluencies and estimates of preparation time. Alternatively, speakers might monitor their preparation more precisely and delay speaking until they have completed some minimum amount of processing for the second name. If speakers can detect when they have selected a name for a second object, they could use it as a criterion for when to begin to speak (see, e.g., Kempen & Huijbers, 1983). If starting speech is contingent on complete or partial preparation of the second name that name should be prepared equally before the onset of both short and long first names.

The traditional measures such as response latencies and speech durations, combined with the measure of gaze duration as a reflection of word preparation, make it possible to determine which criteria speakers use. Although this experimental design does not and cannot address speakers' scope of word preparation or structural planning in conversation, it tests what speakers are capable of considering in their timing of word preparation.

EXPERIMENT 1

Method

Participants. Twenty Stanford undergraduates participated in return for credit in an introductory or cognitive psychology course. All were native speakers of North American dialects of English, and all had normal or corrected-to-normal vision. Three additional participants could not complete the experiment because of equipment problems, and 1 was replaced for not following instructions.

Apparatus. Eye movements were monitored with a remote video-based pupil/corneal reflection system, an ISCAN ETL-400 with a high-speed upgrade sampling at 120 Hz. The stimuli were displayed on a 21-in. monitor. One computer processed eye image data and sent uncalibrated data to another computer, which was responsible for timing, presenting stimuli, digitally recording speech, and calculating and recording calibrated eye position. Speech was recorded at 12 kHz via a SoundBlaster card with the use of a headset microphone. The participants placed their foreheads against a rest to prevent movements in depth. The displayed objects subtended a maximum of 8.8° of visual angle horizontally.

Materials and Design. The pictured objects were line drawings from Snodgrass and Vanderwart (1980), the Philadelphia naming test (Roach, Schwartz, Martin, Grewal, & Brecher, 1996), and Huitema (1996). Every display contained two objects side by side. The length of the left object's dominant name and the difficulty of the right object's name were systematically varied across displays. Thirty-two matched pairs of pictured objects were selected to occupy the left positions of displays and so to be named first. Left-object pairs were composed of an object with a one-syllable name,

Table 1
Mean (and Standard Error of the Mean) Properties of Left-Object Names

Left Name	N	Syllables		Name Agreement		Written Word Frequency		Spoken Word Frequency	
		M	SE	M	SE	M	SE	M	SE
Short	32	1.0	0.0	.947	.009	14.8	2.3	4.6	0.9
Long	32	2.7	0.1	.954	.008	14.3	4.0	6.4	1.4

Note—Word frequencies given are occurrences per million.

such as *wig*, and another object with at least two syllables in its name, such as *windmill*. The names in the pairs were matched in (at least) their initial segments and their noun lemma frequencies from the English Celex database (Baayen, Piepenbrock, & Gulikers, 1995). Because word frequency is highly correlated with length (e.g., Zipf, 1949/1965), most of the left-objects had low-frequency names (see Table 1). All objects had name agreement in norms of .78 or greater (Griffin & Huitema, 1999). See the Appendix for a full list of object names.

Object-naming latencies from an unpublished experiment were available for 22 of the 32 pairs of left objects (Bock, Huitema, & Griffin, 1995). A paired *t* test on median item latencies showed that participants waited a nonsignificant 5.8 msec (*SE* = 26.8 msec) longer to begin saying the long names in the pairs. This lack of a length effect suggests that long- and short-named object pairs were well matched. Moreover, for 29 object pairs, reaction times were available from an object decision study (Huitema, 1993). Some researchers argue that deciding that a picture depicts a real object rather than a nonsense one requires object recognition but not naming. The objects with long names were recognized a nonsignificant 7.9 msec (*SE* = 18.6 msec) faster than those with short names. Response accuracy for both long- and short-named objects was also similar; long-named objects were recognized a nonsignificant 0.12% (1.6%) less accurately than were short-named ones.

Each matched pair of short- and long-named left objects was displayed with another two objects to control for any effect of the second object named. For instance, pictures of the wig and the windmill appeared equally often with the pictures of a bear and a carrot across lists. The 32 pairs of right-position objects varied systematically but slightly in word frequency[3] and length. These differences in the right-position pairs were not critical to the hypothesis being tested and had no significant effects. Therefore, the factor of right-object difficulty will not be discussed further. A similar set of 36 displays was included to pilot another experiment. The minimum horizontal distance between objects was approximately 3.2°.

Two stimulus lists were composed so that each short- and long-named object in a left-object pair was displayed with each object in a right-object pair. Every participant named all 64 left- and 64 right-position objects, with 16 items in each cell from the crossing of left-object name length and right-object difficulty. The displays appeared in a fixed pseudo-random order.

Procedure. Two practice trials and one warm-up trial preceded the 100 trials of the experiment. The participants were tested one at a time. Before every display, the participants fixated a validation point located where the left object would appear. Failure to fixate for 800 msec within 2.5° of the validation point led to a recalibration. The participants were instructed to name the two objects from left to right without pausing between names. Eye movement and voice recording coincided with the onset of picture presentation and ended after the experimenter heard the right-object name.

Results

A naive coder listened to all sound files and noted whether speakers produced the intended names, in left-to-right order, without pausing. The coder also noted any nonspeech noises that preceded speech. Disfluencies such as noticeable pauses between names (greater than 200 msec), fillers ("uh"), stutters or false starts ("wuh-wig"), and drawled segments were noted. Speech onsets and offsets were measured with an autocorrelation-based algorithm implemented in MatLab (Banseel, Griffin, & Spieler, 2001). Measurements made with a waveform editor replaced values generated by the algorithm that were likely to have been triggered by noise other than the object names. Effects reported below as significant yielded min F' statistics[4] with probabilities less than .05.

The speakers produced the anticipated names for both objects on 1,146 (89.5%) trials. The following analyses are on the subset of 985 (77.0%) fluent trials. Means and standard errors for dependent measures appear in Table 2.

Speech. The speakers started naming the left object 78 msec earlier when it had a long name such as *windmill* rather than a short one such as *wig* [min $F'(1,47) = 4.90$]. As expected, the duration of speech, from the onset of the left-object name to the offset of the right-object name, was 74 msec longer when left names were long rather than short [min $F'(1,46) = 13.12$]. Because we counterbalanced the second objects, this duration difference should be wholly attributable to the length of first-object names. These results are consistent with the idea that the speakers prepared second names less before speaking when longer first names provided more time for their prepara-

Table 2
Mean (and Standard Error of the Mean) as a Function of Left-Object Name Length (in Milliseconds)

| | | | | | Right Object | | | | Left Object | | | |
| | Speech Onset | | Articulation Duration | | Gaze Time Prespeech | | Gaze Time During Articulation | | Gaze Time Prespeech | | Gaze Time During Articulation | |
Left Name	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
					Experiment 1							
Short	1,198	44	894	20	482	29	567	27	648	21	217	19
Long	1,120	38	967	22	420	27	659	29	630	18	196	18
Difference	78**		−73†		62†		−92†		18		21	
					Experiment 2: Immediate Naming							
Short	1,102	45	1,039	32	252	28	759	35	782	31	193	24
Long	1,032	42	1,059	32	205	25	813	34	761	31	172	23
Difference	70*		−20		47**		−54*		21		21	
					Experiment 2: "Next To" Intervening							
Short	1,015	44	1,260	42	174	26	920	46	780	32	256	32
Long	1,019	49	1,326	41	153	28	984	46	807	31	242	41
Difference	−4		−66**		21		−64*		−27		14	

Note—Significance values based on min F' combining analyses of variance on the basis of subject and item means.   *p < .10.   **p < .05.   †p < .005.
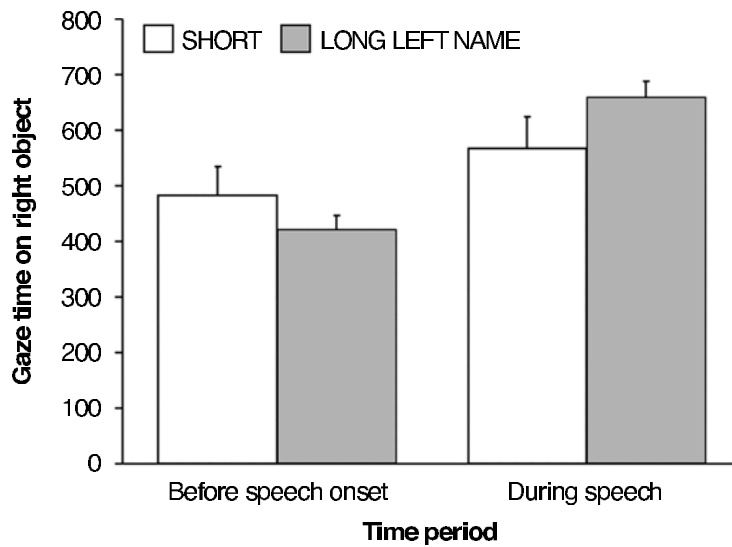
Figure 1. Mean amount of time (in milliseconds) speakers gazed at right (second-named) objects before and during speech, by left-object name length in Experiment 1. Bars indicate standard errors.

tion during speech. More compelling evidence comes from the analyses of eye movement data.

Eye movements. For each display, two large square regions of about 11° × 6° of visual angle, corresponding to object locations, were defined. A gaze was considered to begin with the onset of the first fixation within a region and to end with the start of a saccade leaving the region.

The eye movement data support the speech data. Figure 1 shows the amount of time that the speakers spent gazing at right objects before speech onset and during articulation. Before speech, the speakers gazed at right objects 62 msec longer when the left object had a short name rather than a long name [min $F'(1,49) = 8.87$]. During articulation, the speakers' gaze durations for right objects were 92 msec shorter when the left-object name was short rather than long [min $F'(1,48) = 11.22$]. Thus, the speakers attended to right objects longer prior to speech and less during articulation when the first word that they spoke was short. If we assume that gaze durations are a reflection of word preparation, this result suggests that speakers varied their prespeech preparation of a second name on the basis of the length of the preceding name.

The analyses of gazes to left objects showed no significant effects of length in either time period (see Table 2). This lack of a length effect for objects that were matched on name agreement and frequency is consistent with the observation that length effects in production do not appear in mixed length lists (Meyer, Roelofs, & Levelt, 2003).[5]

Discussion

When the first word spoken was short (e.g., *wig*) rather than long (e.g., *windmill*), the speakers prepared second object names for more time before beginning to speak. This extra preparation was reflected in additional time spent gazing at the second object prior to speech and in delayed speech onsets. The benefit of this preparation was apparent in the eye movement data, with less time spent gazing at the right object during speech. Experiment 2 replicated and extended this result.

EXPERIMENT 2

To ensure that the results of Experiment 1 were due to word length and preparation rather than to other properties of the long- and short-named objects, a new group of speakers named the objects in immediate succession and with the words "next to" inserted between the names. Inserting the additional words between the names was predicted to reduce or eliminate the reversed length effect in speech onsets by allowing speakers more time to prepare second-object names while speaking, even when the first object had a short name, as in *wig next to carrot*.

Method

Participants. Thirty-two students from introductory psychology courses at Georgia Institute of Technology participated for extra credit. All but one learned English before 3 years of age. Data from 3 additional participants could not be used because of equipment problems.

Materials, Apparatus, Procedure, and Design. The same stimuli and equipment were used as those in the previous experiment. Four stimulus lists were created to counterbalance the pairing of long- and short-named left objects with right objects and whether pairs appeared in the first or the second block of naming. The order of the blocks with immediate naming and with "next to" inserted was counterbalanced. In other respects, the procedure was the same as that in Experiment 1.

Results

The speakers produced target responses within 4 sec on 89.0% of all trials. The analyses reported below are based

on the 1,561 (76.2%) fluent trials. Initial analyses of variance (ANOVAs) tested the factors of name length, utterance form, and form order. The predicted interactions between name length and utterance form did not consistently reach significance. Reported below are analyses for the immediate and the "next to" utterances, tested separately. There were no significant main effects of form order or interactions between order and length. Means and standard errors appear in Table 2.

Speech. As before, when naming the two objects in succession, the speakers started speaking marginally later when the first object had a short rather than a long name [$\min F'(1,42) = 3.00, p < .10$]. Speech onset was less affected by the length of the first word when the words "next to" intervened between the object names ($Fs < 1$). Utterances containing longer names and more words took significantly more time to say.

Eye movements. The time spent gazing at the second-named objects before and during speech is plotted in Figure 2. As before, when naming the objects in immediate succession, the speakers gazed at right objects longer prior to speaking when the left object had a short name [$\min F'(1,53) = 4.31$]. When "next to" intervened, the difference was reduced and was not significantly different from zero. The speakers gazed longer at right objects after starting to speak when the left objects had long names. This difference between long and short left names after speech began was similar for both utterance forms and approached significance for both utterance forms for immediate naming [$\min F'(1,44) = 3.69, p < .07$] and for "next to" [$\min F'(1,42) = 3.94, p < .06$] trials. Again, the time spent gazing at left objects did not vary significantly as a function of their name lengths.

The initial ANOVAs showed significant main effects of utterance form in time spent gazing at second objects. When "next to" intervened, the speakers' gaze durations for second objects were 66 msec shorter prior to speaking [$\min F'(1,57) = 7.12$] and 166 msec longer during speech [$\min F'(1,36) = 6.34$]. The additional words provided more preparation time during speech.

Discussion

Experiment 2 replicated the results of Experiment 1. In addition, naming the objects with "next to" inserted between the names showed that the reversed length effect could be reduced by allowing the speakers more time to prepare second names during speech. This result, combined with the equal gaze durations on left objects, makes it unlikely that left objects differed across important dimensions other than name length. Lengthening the utterances with "next to" allowed the speakers to postpone more of second-name preparation until after speech began.

Comparing the immediate naming data of Experiment 2 with that of Experiment 1 revealed several differences between the timing of Stanford and Georgia Tech students in producing identical object names. The populations did not differ significantly in speech onsets. Prior to speaking, however, Tech students spent 132 msec longer than Stanford students did gazing at the first object [$\min F'(1,64) = 8.88$] and 231 msec less time gazing at the second object [$\min F'(1,66) = 22.44$]. Tech students could afford to do less prespeech preparation of second names because they spent 121 msec longer articulating the same object names [$\min F'(1,54) = 6.96$]. During speech, they gazed at second objects for 176 msec longer [$\min F'(1,58) = 11.35$] than did the Stanford students. These data suggest that re-
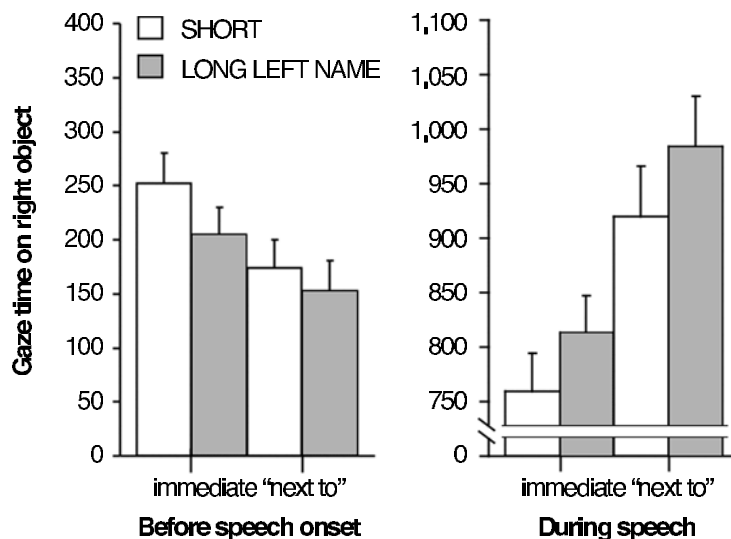

Figure 2. Mean amount of time speakers spent gazing at right (second-named) objects before and during speech as a function of left-object name length and utterance form in Experiment 2. Bars show standard errors.

gional or individual differences in speech rate may also modulate the timing of preparation.

GENERAL DISCUSSION

The speakers varied speech onset on the basis of the amount of time available for preparing words after speech began. Specifically, they gazed longer at the referent of a second noun (e.g., a picture of a carrot) prior to speech when the first word in the pair was short (e.g., *wig carrot*) rather than long (e.g., *windmill carrot*). As a result, they spoke later when the first word was short. Having attended to the second referent longer before saying a short noun, the speakers then spent less time gazing at it after speech began. Furthermore, adding words between the object names allowed the speakers to begin speaking even earlier than they did when naming objects in immediate succession and eliminated the reversed length effect. These results indicate that speakers were able to minimize their buffering[6] of the first word by preparing only as much of the second word before speaking as they needed to. This suggests that speakers can use word length (or something highly correlated with it) to coordinate the timing of word preparation and articulation more generally.

In other studies (e.g., Griffin & Bock, 2000), speakers who have produced sentences to describe pictured events have shown a similar sensitivity to the time available to prepare nouns during the articulation of preceding words. This is apparent using a slightly different way of expressing the consistent eye–voice span originally reported. On the basis of 129 trials with fluent second nouns, 38% of the variance in when speakers shifted their gaze to the referent of the second noun was captured by the following equation:

$$\text{Gaze onset}_{\text{Referent2}} = 1.06\,(\text{Onset}_{\text{Noun2}}) - 999.$$

The onset of the first noun is the zero point in time. When less than 942 msec (999/1.06) intervened between nouns, gazes to the second referent preceded onset of the first noun. With more intervening time, the gazes followed its onset. For every additional millisecond (1.06) between the first noun and the second, speakers shifted their gazes to the referent of the second noun 1 msec later. Because disfluent trials were excluded from this analysis, this difference in timing is not due to speakers' delaying the onset of the second noun because of problems in preparation. Also, unlike those in the experiments reported here, the speakers in Griffin and Bock (2000) showed signs of recognizing the second referent prior to speech. Their choice of subject (first) nouns was primarily based on who-did-what-to-whom and on the relative animacy of the entities involved. Thus, these timing effects are not limited to naming unrelated objects in a prespecified order.

Findings in the motor control literature resemble the reversed length effect observed in the present study. People take more time to start an arm movement when they are instructed to pause for 50 msec rather than for 200 msec before performing another movement (Ketelaars, Garry, & Frank, 1997). So the less the time available for pro-

gramming a second movement after the first begins, the later the first one starts. The fact that this effect is due to pause duration rather than to additional movements suggests that the causal correlate of word length in the present studies may be based on timing rather than on the number of syllables or speech movements.

The present findings argue against the notion that within a particular situation speakers consistently prepare a phonological word (see, e.g., Wheeldon & Lahiri, 1997) or major constituent (see, e.g., Smith & Wheeldon, 1999) before beginning to speak. Speakers may prepare multiple content words before speaking in order to produce a fluent utterance despite short utterance-initial words rather than to obey a requirement of syntactic processing. If so, the interpretations of important works such as that of Kempen and Huijbers (1983) may change. Of course, the range of circumstances under which speakers spontaneously modulate their timing according to available preparation time remains to be seen. Nonetheless, when preparing actions takes longer than executing them, timing may be everything.

REFERENCES

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (Version 2) [CD-ROM]. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bachoud-Lévi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory & Language*, **39**, 331-346.

Banseel, P., Griffin, Z. M., & Spieler, D. H. (2001). *Epd: MatLab wave file parsing software.* Available from http://oak.psych.gatech.edu/~spieler/software.html

Bock, K., Huitema, J. S., & Griffin, Z. M. (1995). *Repetition priming and the processing components of object naming.* Unpublished manuscript.

Butterworth, B. (1989). Lexical access in speech production. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108-135). Cambridge, MA: MIT Press.

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

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times to prepared utterances. *Journal of Memory & Language*, **30**, 210-233.

Fischer, B. (1998). Attention in saccades. In R. D. Wright (Ed.), *Visual attention* (pp. 289-305). New York: Oxford University Press.

Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey, Vol. 3. Language: Psychological and biological aspects* (pp. 69-96). Cambridge: Cambridge University Press.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* London: Academic Press.

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

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

Griffin, Z. M., & Huitema, J. S. (1999). *Beckman spoken picture naming norms.* Available from http://langprod.cogsci.uiuc.edu/~norms/

Huitema, J. S. (1993). *Planning referential expressions in speech production.* Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Huitema, J. S. (1996). [The Huitema picture collection]. Unpublished raw data.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, **11**, 201-258.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sen-

tence production and naming: Indirect election of words. *Cognition*, **14**, 185-209.

KETELAARS, M. A. C., GARRY, M. I., & FRANK, I. M. (1997). On-line programming of simple movement sequences. *Human Movement Science*, **16**, 461-483.

LEVELT, W. J. M. (1989). *Speaking: From intention to articulation.* Cambridge, MA: MIT Press.

LINDSLEY, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, **7**, 1-19.

MARTIN, R. C., VU, H., MILLER, M., & FREEDMAN, M. (2000, September). *Working memory in language production.* Paper presented at the conference on Architectures and Mechanisms of Language Processing, Leiden.

MEYER, A. S., ROELOFS, A., & LEVELT, W. J. M. (2003). Word length effects in picture naming: The role of a response criterion. *Journal of Memory & Language*, **48**, 131-147.

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.

PECHMANN, T. (1989). Incremental speech production and referential overspecification. *Linguistics*, **27**, 89-110.

POTTER, M. (1975). Meaning in visual search. *Science*, **187**, 965-966.

RAAIJMAKERS, J. G. W., SCHRIJNEMAKERS, J. M. C., & GREMMEN, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory & Language*, **41**, 416-426.

ROACH, A., SCHWARTZ, M. F., MARTIN, N., GREWAL, R. S., & BRECHER, A. (1996). The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*, **24**, 121-133.

SMITH, M., & WHEELDON, L. (1999). High level processing scope in spoken sentence production. *Cognition*, **73**, 205-246.

SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.

SNODGRASS, J. G., & YUDITSKY, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, **28**, 516-536.

WHEELDON, L., & LAHIRI, A. (1997). Prosodic units in speech production. *Journal of Memory & Language*, **37**, 356-381.

ZELINSKY, G. J., & MURPHY, G. L. (2000). Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science*, **11**, 125-131.

ZIPF, G. K. (1965). *Human behavior and the principle of least effort.* New York: Hafner. (Original work published 1949)

NOTES

1. It takes approximately 150–200 msec to program and execute an eye movement, so gazes lag behind attention (see, e.g., Fischer, 1998). Gaze durations in speaking are also affected by variables associated with retrieving the sounds of words (Meyer, Sleiderink, & Levelt, 1998; Meyer & van der Meulen, 2000).

2. Everything else being equal, short and long object names may have identical object-naming latencies (Bachoud-Lévi, Dupoux, Cohen, & Mehler, 1998). Meyer, Roelofs, and Levelt (2003) demonstrated that length effects in object-naming latencies vary with list composition. Because mixed length lists are used here, the only anticipated difference between long and short names was the time needed to articulate them and hence the time they allow during speech for preparing the next word. Meyer et al. also replicated the results of Experiment 1.

3. On the basis of noun lemma frequencies in Celex (Baayen et al., 1995), the mean frequency difference was 26 ($SE = 6$) occurrences per million in spoken language and 53 ($SE = 11$) in written language.

4. MinF' combines the results of ANOVAs on the basis of variability across subjects and items. It is interpreted like other F statistics, but p values of .05 in subject and item analyses result in a minF' corresponding to $p > .05$. Because its degrees of freedom include MS_e, it saves journal space without sacrificing information. See Clark (1973) and Raaijmakers, Schrijnemakers, and Gremmen (1999) for details.

5. It is only when participants silently retrieve and rehearse object names for memory tests that gaze durations on the objects themselves seem to reflect word length (Zelinsky & Murphy, 2000).

6. Surprisingly, people with severely impaired working memory spans may generate multiword utterances. For example, Martin, Vu, Miller, and Freedman (2000) found that an individual with a memory span of one item could successfully produce picture descriptions of the form *the car is blue* but had trouble with *the blue car*. The present results suggest that the adjective–noun sequence probably required more buffering of content words than the noun–copula–adjective sequence did. This tiny difference in buffering may have overwhelmed the individual's limited capacity. However, adjective–noun sequences with longer adjectives, such as *the beautiful car*, should be easier for this individual to produce.

APPENDIX
Table of Picture Names in the Present Study

Set	Left Object		Right Object	
---	Short	Long	Easy	Hard
1	chef	chandelier	tank	cactus
2	comb	compass	bat	tractor
3	drum	dragon	baby	zebra
4	fence	finger	boot	toaster
5	frog	flashlight	clock	pencil
6	harp	hamburger	window	monkey
7	kite	kangaroo	cigarette	toothbrush
8	lock	ladder	computer	cherry
9	owl	octopus	knife	candle
10	pear	pineapple	queen	camel
11	pen	piano	truck	bullet
12	rake	refrigerator	witch	barrel
13	skunk	screwdriver	knot	balloon
14	snail	snowman	nun	angel
15	spoon	spider	well	peanut
16	wig	windmill	bear	carrot
17	tie	toilet	saw	hammock
18	vest	volcano	bar	whistle
19	sled	slingshot	chain	turtle
20	broom	butterfly	can	shovel
21	ham	helicopter	bomb	igloo
22	bench	button	tire	canoe
23	cake	camera	leaf	anchor
24	cow	calculator	lip	giraffe
25	clown	calendar	lamp	scissors
26	ghost	guitar	bird	mailbox
27	pie	penguin	bridge	handcuffs
28	safe	sandwich	fish	hammer
29	scarf	skeleton	pipe	pumpkin
30	maze	motorcycle	apple	ruler
31	tent	typewriter	bone	castle
32	axe	ambulance	bell	blender