



HHS Public Access

Author manuscript

Psychon Bull Rev. Author manuscript; available in PMC 2017 July 28.

Published in final edited form as:

Psychon Bull Rev. 2003 September ; 10(3): 603–609.

A Reversed Word Length Effect in Coordinating the Preparation and Articulation of Words in Speaking

Zenzi M. Griffin

Georgia Institute of Technology & Stanford University

Abstract

Speakers tend to prepare their nouns immediately before saying them rather than 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 speakers delayed starting to say the first name based on the time available to prepare the second name during speech. Before speaking, they spent more time preparing a second name (*carrot*) when the first was monosyllabic (*wig*) rather than multisyllabic (*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 only had to worry about one action at a time. Unfortunately, it's not that easy. Instead, people often must 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 40-50% of the time a person spends speaking is actually silence, with 45% of 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 was the result of selecting all of the words in an utterance before speaking, and only retrieving the words' sounds during speech. In this view, disfluencies occurred when speakers tried to decide the content of their speech or which words to use while speaking (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 suggest that even when fluent, speakers choose some words after they start speaking rather than beforehand (e.g., Huitema, 1993; Kempen & Huijbers, 1983; Lindsley, 1975; Pechmann, 1989; Smith & Wheeldon, 1999). It is only in experiments where speakers are compelled to prepare their speech in advance or recall sentences that evidence of buffering is observed (e.g., Ferreira, 1991; Griffin & Bock, 2000; Wheeldon & Lahiri, 1997).

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 one second to begin naming an object presented in isolation (e.g. Snodgrass & Yuditsky, 1996), suggesting that selecting a name and retrieving its sounds take slightly less time. 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, 1973). The more appropriate names an object has, the longer speakers gaze at it (Griffin, 2001). These gazes¹ in mid-utterance suggest that speakers do not just retrieve word sounds as they speak; they also decide which words to use.

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 one second to prepare every word, they could not prepare and utter one word at a time without hesitating between words. However, all words are not equally difficult to prepare. Due to their high frequency and role in structuring sentences, function words (e.g., *the*, *might*, *on*) are likely to take far less time to prepare than content words (*cow*, *lovely*, *walk*) do. While function 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 words 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 (e.g., Kempen & Hoenkamp, 1987; Smith & Wheeldon, 1999). When the articulation of such a unit provides enough time to prepare the next one, speech would be fluent. When there is not enough time for preparation, speech would be disfluent. However, speakers seem able to modulate their preparation with more precision than either of these strategies suggests. But if they can estimate the time available during speech for preparing a word, what is such an estimate based on?

The experiments reported here test whether speakers can modulate their word preparation based on word length (or another correlate of articulation duration). Speakers were asked to label 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 ms to articulate, speakers do not have the necessary 800 ms to prepare a second name like *carrot* during speech. If speakers begin to say *wig* as soon as

¹It takes approximately 150-200 ms to program and execute an eye movement, so gazes lag behind attention (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).

they have prepared it, *carrot* will not be ready when needed and they will hesitate between nouns. That is, they 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 *carrot* 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 like *wig* should compel them to prepare much of *carrot* before beginning to speak. In contrast, a longer first word like *windmill* would allow speakers to begin speaking earlier, preparing *carrot* less before saying the first name².

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 beginning to speak (e.g., Kempen & Huijbers, 1983). If starting speech is contingent on complete or partial preparation of the second name, it should be equally prepared before the onset of short and long first names.

By using the traditional measures such as response latencies and speech durations, combined with gaze durations as a reflection of word preparation, it is 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 for credit in an introductory or cognitive psychology course. All were native speakers of North American dialects of English with normal or corrected-to-normal vision. Three additional participants could not complete the experiment due to equipment problems, and one 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. Stimuli were displayed on a 21-inch monitor. One computer processed eye image data, sending uncalibrated data to another, which was responsible for timing, presenting stimuli, digitally recording speech, calculating and recording calibrated eye position. Speech was recorded at 12 kHz via a SoundBlaster card, using a headset microphone. Participants placed

²Everything else being equal, short and long object names may have identical object naming latencies (Bachoud-Levi, Dupoux, Cohen, & Mehler, 1998). Meyer, Roelofs, and Levelt (in press) 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.

their foreheads against a rest to prevent movements in depth. Displayed objects subtended a maximum of 8.8° of visual angle horizontally.

Materials and design—Pictured objects were line drawings from Snodgrass and Vanderwart (1980), the Philadelphia naming test (Roach et al, 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 position of displays and be named first. Each pair was composed of an object with a one-syllable name such as *wig* and another with at least two syllables such as *windmill*. Names in 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 word pairs were low frequency (see Table 1). All objects had name agreement in norms of .78 or greater (Griffin & Huitema, 1999).

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 non-significant 5.8 (SE = 26.8) ms longer to begin saying the long names in the pairs. This lack of a length effect suggests that long and short object pairs were well-matched. Moreover, for 29 object pairs, reaction times were available from an object decision study (Huitema, 1993). Deciding that a picture depicts a real object rather than a nonsense one is argued to require object recognition but not naming. The objects with long names were recognized a non-significant 7.9 (18.6) ms faster than those with short names. The objects were also similar in response accuracy, with long named objects a non-significant 0.12% (1.6) less accurate than short named ones.

Each matched pair of short- and long-name left objects was displayed with another two objects to control for any effect of the second object named. For instance, the wig and the windmill appeared equally often with a bear and a carrot across lists. The thirty-two pairs of right-position objects varied systematically but slightly in word frequency³ and length. These differences in right-hand pairs were not critical to the hypothesis being tested and had no significant effects. Therefore, the factor of right-object difficulty will not be further discussed. A similar set of 36 displays was included to pilot another experiment. The minimum horizontal distance between objects was about 3.2°.

Two stimulus lists were composed so that each short- and long-name 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. Displays appeared in a fixed pseudo-random order.

Procedure—Two practice trials and one warm-up trial preceded the 100 trials of the experiment. Participants were tested one at a time. Before every display, participants fixated

³Based on noun lemma frequencies in Celex, the mean frequency difference was 26 (6) occurrences per million in spoken language and 53 (11) in written.

a validation point located where the left object would appear. Failure to fixate for 800 ms within 2.5° of the validation point led to a re-calibration. 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 naïve coder listened to all soundfiles and noted whether speakers produced the intended names, in left-to-right order, without pausing. The coder also noted any non-speech noises that preceded speech. Disfluencies such as noticeable pauses between names (greater than 200 ms), 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 be triggered by noise other than object names. Effects reported as significant below yielded *minF* statistics⁴ with probabilities less than .05.

Speakers produced the anticipated names for both objects on 1146 (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—Speakers started naming the left object 78 ms earlier when it had a long name like *windmill* rather than a short one like *wig*, $minF(1,47) = 4.90$. As expected, the duration of speech, from the onset of the left-object name to the offset of right-object name, was 74 longer when left names were long rather than short, $minF(1,46) = 13.12$. Due to counterbalancing second objects, this duration difference should be wholly attributable to the length of first object names. These results are consistent with the idea that speakers prepared second names less before speaking when longer first names provided more time for their preparation during speech. More compelling evidence comes from analyses of eye movement data.

Eye movements—For each display, two large square regions of about 11° by 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 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 speakers spent gazing at right objects before speech onset and during articulation. Before speech, speakers gazed at right objects for 62 ms longer when the left object had a short rather than long name, $minF(1,49) = 8.87$. During articulation, speakers gazed at right objects for 92 ms less time when the left object name was short rather than long, $minF(1,48) = 11.22$. Thus, speakers attended to right objects for more time prior to speech and less

⁴*MinF* combines the results of ANOVAs based on 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 *MSE*, it saves journal space without sacrificing information. See Clark (1973) and Raaijmakers, Schrijnemakers, and Gremmen (1999) for details.

during articulation when the first word they said was short. Assuming that gaze durations reflect word preparation, this suggests speakers varied their pre-speech preparation of a second name based on the length of the preceding name.

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, in press).⁵

Discussion

When the first word spoken was short (*wig*) rather than long (*windmill*), speakers prepared second object names more 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 less time spent gazing at the right object during speech. The next experiment replicates and extends this result.

Experiment 2

To ensure that the results of Experiment 1 were due to word length and preparation rather than 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 names. 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 the age of 3 years. Data from three additional participants could not be used due to equipment problems.

Materials, apparatus, procedure, and design—The same stimuli and equipment were used as before. Four stimulus lists were created to counterbalance the pairing of long and short name left objects with right objects and whether pairs appeared in the first or second block of naming. The order of the blocks with immediate naming and “next to” inserted was counterbalanced. In other respects, the procedure was the same as in Experiment 1.

Results

Speakers produced target responses within 4 s on 89.0% of all trials. The analyses reported below are based the 1561 trials (76.2%) fluent trials. Initial 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

⁵It 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).

analyses for the immediate and “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, 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 first word length when the words “next to” intervened between the object names, $F_s < 1$. Utterances containing longer names and more words took significantly more time to say.

Eye movements—Time spent gazing at second named objects before and during speech is plotted in Figure 2. As before, when naming the objects in immediate succession, speakers gazed at right objects longer prior to speaking when the left object had a shorter name, $\min F(1,53) = 4.31$. When “next to” intervened, the difference was reduced and not significantly different from zero. Speakers gazed longer at right objects after starting to speak when 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, for immediate naming, $\min F(1,44) = 3.69, p < .07$, and for “next to” trials, $\min F(1,42) = 3.94, p < .06$. Again, the time spent gazing at left objects did not vary significantly as 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, speakers spent 66 ms less time gazing at second objects prior to speaking, $\min F(1,57) = 7.12$, and 166 ms 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 names showed that the reversed length effect could be reduced by allowing speakers more time to prepare second names during speech. Combined with the equal gaze durations on left object, this makes it unlikely that left objects differed across important dimensions other than name length. Lengthening the utterances with “next to” allowed speakers to postpone more of second name preparation until after speech began.

Comparing the immediate naming data of Experiment 2 to 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 ms longer than Stanford students did gazing at the first object, $\min F(1,64) = 8.88$, and 231 ms less time gazing at the second object, $\min F(1,66) = 22.44$. Tech students could afford to do less pre-speech preparation of second names because they spent 121 ms longer articulating the same object names, $\min F(1,54) = 6.96$. During speech, they gazed at second objects for 176 ms longer, $\min F(1,58) = 11.35$. These data suggest that regional or individual differences in speech rate may also modulate the timing of preparation.

General Discussion

Speakers varied when they began to speak based on 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 their first word was short (e.g., *wig carrot*) rather than long (*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, speakers then spent less time gazing at it after speech began. Furthermore, adding words between object names allowed 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 buffering⁶ of the first word by only preparing 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, speakers who produced sentences to describe pictured events showed a similar sensitivity to the time available to prepare nouns during the articulation of preceding words (e.g., Griffin & Bock, 2000). This is apparent using a slightly different way of expressing the consistent eye-voice span originally reported. Based on 129 trials with fluent second nouns, 38% of the variance in when speakers shifted their gaze to the referent of their second noun was captured by equation below.

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

The onset of the first noun is the zero point in time. When less than 942 ms (999/1.06) intervened between nouns, gazes to the second referent preceded the onset of the first noun. With more intervening time, the gazes followed its onset. For every additional millisecond (1.06) of speech between the first noun and the second, speakers shifted their gazes to the referent of the second noun a millisecond later. Because disfluent trials were excluded from this analysis, this timing is not due to speakers delaying the onset of the second noun due to problems in preparation. Also, unlike the experiments reported here, the speakers in this earlier study 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 the relative animacy of the entities involved. Thus, these timing effects are not limited to naming unrelated objects in a pre-specified order.

Findings in the motor control literature resemble the reversed length effect observed here. People took more time to start an arm movement when they were instructed to pause for 50 rather than 200 ms before performing another movement (Ketelaars, Garry, & Franks, 1997). So, the less time available for programming a second movement after a first began, the later

⁶Surprisingly, 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.

the first one started. The fact that this effect was due to pause duration rather than additional movements suggests that the causal correlate of word length in the present studies may be based on timing rather than the number of syllables or movements.

Returning to the language literature, the present findings argue against the notion that within a particular situation speakers consistently prepare a phonological word (e.g., Wheeldon & Lahiri, 1997) or major constituent (e.g., Smith & Wheeldon, 1999) before beginning to speak. Speakers may prepare multiple content words before speaking 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 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.

Acknowledgments

Thanks to Amit Mookerjee and Prakeet Banseel for their excellent programming, to Nadine de Lassus, Alexia Galati, Kristin Garton, A. J. Margolis, Azucena Rangel, and Russell Smith for their help in data collection and coding, and to Azucena Rangel and three helpful reviewers for comments on the manuscript. This research was supported by a grant from the National Institutes of Health, R03 MH61318-01.

Appendix

Table of experimental picture names.

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

Set	Left object		Right object	
	Short	Long	Easy	Hard
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

References

- Baayen, RH., Piepenbrock, R., Gulikers, L. The CELEX Lexical Database (Version 2) [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania; 1995. Distributor
- Bachoud-Levi AC, Dupoux E, Cohen L, Mehler J. Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language*. 1998; 39:331–346.
- Banseel, P., Griffin, ZM., Spieler, DH. MatLab routine for measuring speech onset and offset (Version 1.0). 2001. <http://psychology.gatech.edu/spieler/software.html>
- Bock K, Huitema JS, Griffin ZM. Repetition priming and the processing components of object naming. 1995 Unpublished data.
- Butterworth, B. Lexical access in speech production. In: Marslen-Wilson, WD., editor. *Lexical representation and process*. Cambridge, MA: MIT Press; 1989. p. 108-135.
- Clark HH. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*. 1973; 12:335–359.
- Ferreira F. Effects of length and syntactic complexity on initiation times to prepared utterances. *Journal of Memory and Language*. 1991; 30:210–233.
- Fischer, B. Attention in saccades. In: Wright, RD., editor. *Visual attention*. New York, NY: Oxford University Press; 1998. p. 289-305.
- Garrett, MF. Processes in language production. In: Newmeyer, FJ., editor. *Linguistics: The Cambridge survey, III: Language: Psychological and biological aspects*. Vol. 3. Cambridge: Cambridge University Press; 1988. p. 69-96.
- Goldman-Eisler, F. *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press; 1968.
- Griffin ZM. Gaze durations during speech reflect word selection and phonological encoding. *Cognition*. 2001; 82:B1–B14. [PubMed: 11672707]
- Griffin ZM, Bock K. What the eyes say about speaking. *Psychological Science*. 2000; 11:274–279. [PubMed: 11273384]
- Griffin, ZM., Huitema, JS. Beckman Spoken Picture Naming Norms. 1999. <http://langprod.cogsci.uiuc.edu/~norms/>
- Huitema, JS. Unpublished Dissertation. University of Massachusetts; 1993. Planning referential expressions in speech production.

- Huitema JS. The Huitema picture collection. 1996 Unpublished.
- Kempen G, Hoenkamp E. An incremental procedural grammar for sentence formulation. *Cognitive Science*. 1987; 11:201–258.
- Kempen G, Huijbers P. The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*. 1983; 14:185–209.
- Ketelaars MAC, Garry MI, Frank IM. On-line programming of simple movement sequences. *Human Movement Science*. 1997; 16:461–483.
- Levelt, WJM. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press; 1989.
- Lindsay JR. Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*. 1975; 7:1–19.
- Martin, RC., Vu, H., Miller, M., Freedman, M. Working memory in language production. Paper presented at Architectures and Mechanisms of Language Processing; Leiden, The Netherlands. 2000 Sep.
- Meyer AS, van der Meulen FF. Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin and Review*. 2000; 7:314–319. [PubMed: 10909139]
- Meyer AS, Roelofs A, Levelt WJM. Word length effects in picture naming: The role of a response criterion. *Journal of Memory and Language*. in press.
- Meyer AS, Sleiderink A, Levelt WJM. Viewing and naming objects: Eye movements during noun phrase production. *Cognition*. 1998; 66:B25–B33. [PubMed: 9677766]
- Pechmann T. Incremental speech production and referential overspecification. *Linguistics*. 1989; 27:89–110.
- Potter M. Meaning in visual search. *Science*. 1973; 187:965–966.
- Raaijmakers JGW, Schrijnemakers JMC, Gremmen F. How to deal with ‘The Language-As-Fixed-Effect Fallacy’: Common misconceptions and alternative solutions. *Journal of Memory and Language*. 1999; 41:416–426.
- Roach A, Schwartz MF, Martin N, Grewal RS, Brecher A. The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology*. 1996; 24:121–133.
- Smith M, Wheeldon L. High level processing scope in spoken sentence production. *Cognition*. 1999; 73:205–246. [PubMed: 10585515]
- Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*. 1980; 6:174–215. [PubMed: 7373248]
- Snodgrass JG, Yuditsky T. Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, and Computers*. 1996; 28:516–536.
- Wheeldon L, Lahiri A. Prosodic units in speech production. *Journal of Memory and Language*. 1997; 37:356–381.
- Zelinsky GJ, Murphy GL. Synchronizing visual and language processing: An effect of object name length on eye movements. *Psychological Science*. 2000; 11:125–131. [PubMed: 11273419]
- Zipf, GK. *Human behavior and the principle of least effort*. New York: Hafner; 1949/1965.

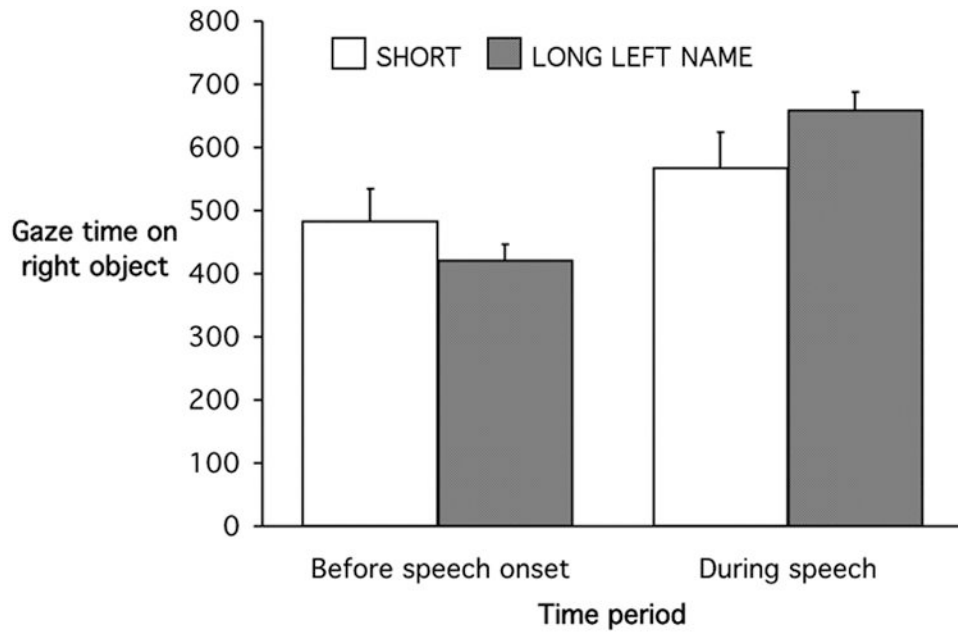


Figure 1. Mean time speakers gazed at right (second-named) objects before and during speech, by left object name length in Experiment 1. Bars indicate standard errors.

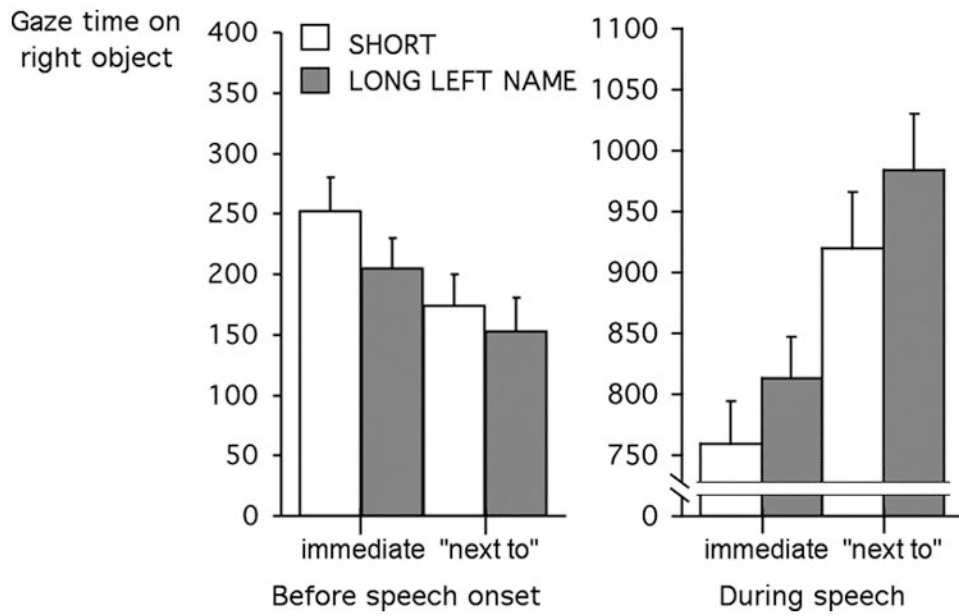


Figure 2. Mean time speakers spent gazing at right (second named) objects before and during speech as a function of left object name length and utterance form in Experiment 2. Bars show standard errors.

Table 1

Mean (and standard error of mean) properties of left-object names.

Left name	Syllables	Name agreement	Written/mln	Spoken/mln
Short	1.0 (0.0)	.947 (.009)	14.8 (2.3)	4.6 (0.9)
Long	2.7 (0.1)	.954 (.008)	14.3 (4.0)	6.4 (1.4)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Mean (and standard error of mean) as a function of left object name length.

Left name	Speech onset	Articulation duration	Right object			Left object		
			Gaze time pre-speech	Gaze time during	Gaze time pre-speech	Gaze time pre-speech	Gaze time during	
Experiment 1								
Short	1198 (44)	894 (20)	482 (29)	567 (27)	648 (21)	217 (19)		
Long	1120 (38)	967 (22)	420 (27)	659 (29)	630 (18)	196 (18)		
Diff	78 **	-73 ***	62 ***	-102 ***	18	21		
Experiment 2 immediate naming								
Short	1102 (45)	1039 (32)	252 (28)	759 (35)	782 (31)	193 (24)		
Long	1032 (42)	1059 (32)	205 (25)	813 (34)	761 (31)	172 (23)		
Diff	70 *	-20	47 **	-54 *	21	21		
Experiment 2 "next to" intervening								
Short	1015 (44)	1260 (42)	174 (26)	920 (46)	780 (32)	256 (32)		
Long	1019 (49)	1326 (41)	153 (28)	984 (46)	807 (31)	242 (41)		
Diff	-4	-66 **	21	-64 *	-27	14		

Note. Significance values based on *minF* combining ANOVAs based on subject and item means.

* indicates $p < .10$,

** indicates $p < .05$, and

*** indicates $p < .005$.