Journal of Psycholinguistic Research, Vol. 35, No. 2, March 2006 (© 2006)
DOI: 10.1007/s10936-005-9011-6

Planning at the Phonological Level during Sentence Production

Tatiana T. Schnur,[1,2,5] Albert Costa,[3] and Alfonso Caramazza[4]

In two picture–word interference experiments we examined whether phrase boundaries affected how far in advance speakers plan the sounds of words during sentence production. Participants produced sentences of varying lengths (short: determiner + noun + verb or long: determiner + adjective + noun + verb) while ignoring phonologically related and unrelated words to the verb of the sentence. Response times to begin producing both types of sentences were faster in the presence of a related versus unrelated distractor. The results suggest that the activation of phonological properties of words outside the first phrase and first and second phonological word affect onset of articulation during sentence production. The results are discussed in the light of previous evidence of phonological planning during multi-word production. Implications for the phonological facilitation effect in the picture–word interference paradigm are also discussed.

KEY WORDS: phonological encoding; planning; picture–word interference paradigm; sentence production.

INTRODUCTION

Although producing speech feels unplanned and spontaneous, our language system considers multiple words before speaking begins. When a speaker wants to produce a sentence, how much phonological information is activated in advance of speaking? Here we investigated whether phrase and word boundaries affected the extent of phonological planning during sentence production. Previous research has shown that multiple words

---

The research reported in this article served as partial fulfillment of the requirements for the PhD degree at Harvard University for Tatiana T. Schnur.

[1] Moss Rehabilitation Research Institute, Korman, Suite 213, 1200 West Tabor Road, Philadelphia, PA 19147, USA.
[2] University of Pennsylvania, Philadelphia, PA, USA.
[3] Universitat de Barcelona, Barcelona, Spain.
[4] Harvard University, Cambridge, MA, USA.
[5] To whom all correspondence should be addressed. e-mail: tschnur@mail.med.upenn.edu

are phonologically planned before a single phrase (e.g., a noun phrase like *the red car*) is produced (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Jescheniak *et al.*, 2003; Miozzo & Caramazza, 1999) but only one word and phrase (e.g., *the arrow*) is phonologically planned in advance of articulation when multiple phrases are produced (e.g., *the arrow and the bag*; Meyer, 1996). The disparate results left open the possibility that the extent of phonological planning is determined by phrase boundaries (major grammatical or phonological phrase boundaries), not by phonological word boundaries (a content word and any unstressed function word) suggesting that entire phrases are phonologically planned before articulation begins. We present results from two experiments that investigated how far in advance of articulation the sounds associated with a sentence are planned when multiple phrases are produced. We used the picture–word interference paradigm where participants produced intransitive sentences of varying lengths (either with a simple or complex subject noun phrase) while ignoring distractor words phonologically related and unrelated to the verb (the second or third phonological word) of the sentence. The results suggest that phonological planning extends across phrase boundaries during sentence production. The sounds associated with several words are retrieved before articulation begins. We discuss the implication of these results for models that predict how far in advance of articulation the language system considers future words.

Because speech errors occur in a systematic fashion, they are a potentially useful database for examining processing constraints on speech production. Evidence for the extent of planning during speech production has come from two types of speech errors, word exchanges and sound exchanges. For example, in the word exchange error *Give a baby to the banana* (example taken from Meyer, 1996), words are exchanged in a sentence and these exchanges can occur across phrases and usually within a clause. These errors usually involve whole words and tend to be restricted to the same syntactic class (e.g., noun for noun, verb for verb). In contrast, sound exchange errors like *Bill snovels show* (vs. *Bill shovels snow*) (taken from Garrett, 1980), involve sound exchanges and occur regardless of grammatical class (see Dell & Reich, 1981 for discussion of errors reflecting both phonological and grammatical processing constraints, i.e., lexical-bias effects). Because speech errors obey different constraints, they are thought to arise at different levels of encoding. In the case of whole word exchanges, because the exchanges are whole words and are not related by sound it is believed that they occur at the grammatical encoding stage when lexical items are selected and ordered. In the case of sound exchanges, the errors are believed to arise at the phonological encoding level where the sounds of words are selected. Phonological errors can thus reveal how much of an utterance is phonologically available at one time.

Patterns in phonological errors suggest that phonological encoding is less extensive than grammatical encoding, in that phonological encoding occurs within phrases (involving adjacent words), as opposed to across clausal boundaries. For example, sound exchanges occur within a phrase and clause, as opposed to across clauses, approximately 87% of the time (where a phrase is defined as a simple noun phrase (NP), or a verb-phrase (VP) and any of its direct object NPs; Garrett, 1980).

Theoretically, how much planning produces fluent speech? Consider production of the sentence *The girl kicks the ball.* Assume that the message has been conceptually formulated, and the speaker now needs to select the words that correspond to the concepts in the message. If each processing level must wait for processing of the entire message before the next level of processing to begin, then all the words corresponding to the message would have to be selected and encoded grammatically before the sounds were retrieved during phonological encoding. Before the utterance could be articulated, all the sounds corresponding to the utterance would have to be selected. This planning strategy might result in long hesitations between bursts of speech. In order to facilitate fluent speech and limit severe dysfluencies it has been suggested that only a minimal amount of information needs to be processed at one level for the next processing level to begin (Levelt, 1989). With incremental planning, words are delivered one at a time from grammatical to phonological encoding. Articulation of the motor program for the first syllable *the* begins when the phonological segments corresponding to *the girl*, (the first phonological word) are syllabified. Onset of articulation depends on how long it takes to construct the first phonological word of the utterance. In most cases, the amount of planning needs to be no more than one lexical or phonological word in advance of articulation. However, there may be some exceptions. Planning may vary as a function of the rate of speech, where slow rates may induce more planning. Alternatively, articulation in some cases may begin with less than a phonological word (Meyer *et al.*, 2003; Schriefers & Teruel, 1999). However, it is argued that in general, speech proceeds one lexical and phonological word at a time, where "execution can follow phonological encoding at a very short distance, a distance smaller than a full phonological phrase. This distance is probably the size of a phonological word... [where] buffering will be minimal or absent" (Levelt, 1989, p. 421). We will refer to this idea as radical incrementality (following Ferreira & Swets, 2002). Thus the question of interest here is the degree to which phonological planning is radically incremental during sentence production, e.g., whether articulation follows phonological encoding at a distance of a phonological word, or whether several elements (words or phrases) are phonologically activated before articulation begins.

Speech error evidence suggests that phonological information is activated one word in advance of speaking. Although speech error evidence has provided useful insights into how the speech production system is organized, errors may not reflect normal speech processing (see Meyer, 1992). With regard to planning, slips of the tongue may arise precisely because too many words are computed at the same time. They thus erroneously suggest more planning than normally occurs. Although error evidence has provided insight into planning in spontaneous speech for the reasons mentioned above, it is important to complement error evidence with evidence from speech produced without errors, under experimental control.

The question of whether articulation follows phonological encoding within a distance of a phonological word (radical incrementality) has been addressed here and elsewhere using the picture–word interference paradigm. In this paradigm, participants simply name a picture. For example, they may see a picture of a table and produce *table*. Printed on the picture is a distractor word which participants automatically read although they are told to ignore it. A distractor word interferes with picture naming by affecting naming speed. By varying the relationship between the word and the picture, the interference produced by the distractor is systematically varied. For example, participants name the picture *table* and the word written inside is *chair*. These concepts are semantically (categorically) related. Participants are slower to name *table* with *chair* written inside in comparison to an unrelated word. This is referred to as the semantic interference effect (Glaser & Dungelhoff, 1984; Glaser & Glaser, 1989; Lupker, 1979). Naming of the picture in this case is thought to be slower because the closer the word and picture are in meaning, the harder it is to decide which one must be selected for production (Levelt *et al.*, 1999; Roelofs, 1992). While the semantic interference effect slows selection of the correct word, another effect occurs when sounds are retrieved. For example, a participant names the picture *table* with the word *tape* written inside. Here, picture naming is accelerated when the word sounds like the picture name in comparison to an unrelated word. This is referred to as the phonological facilitation effect (Lupker, 1982; Meyer & Schriefers, 1991; Rayner & Springer, 1986). This acceleration of speech could be the result of overlapping of sound representations between the picture name and the written word.

It is important to note that depending on where the phonological facilitation effect occurs during the speech production process, phonological facilitation effects may reflect processes at more than one level of speech production. With the picture–word interference paradigm, evidence suggests that the semantic interference and phonological facilitation effects arise at different levels. Because semantic and phonological effects follow different time courses (depending on when a distractor is displayed relative to the picture), the effects are thought to arise at different levels of the speech process (Levelt *et al.*, 1991; Schriefers *et al.*, 1990; Starreveld & La Heij, 1996). Semantic interference is thought to arise during the selection of the lexical item, at the grammatical encoding stage. Phonological facilitation is thought to occur during selection of the phonological representation of the word to be produced, during phonological encoding. We return to this issue in the General Discussion.

Because the domain of syllabification is the phonological word (Levelt, 1989, 1992; Levelt & Wheeldon, 1994) and this is the smallest prosodic unit, most studies have examined whether one phonological word (radically incremental planning) or more than one phonological word is planned before articulation. Meyer (1996) looked at the extent of planning at the grammatical and phonological encoding stages for Dutch utterances of two objects using the picture–word interference paradigm. For example, participants described pictures of object pairs using phrases like *de pijl en de tas* (the arrow and the bag) and *de pijl staat naast de tas* (the arrow is next to the bag). To look at planning at the grammatical encoding stage, semantically related word distractors to the first (spear for arrow) or the second noun (suitcase for bag) were displayed during picture naming. It was found that response times were slowed when the distractor was semantically related to either noun. This suggested that before participants produced the utterance, they had already selected the first and second noun at the grammatical level. To examine phonological encoding, the distractor word was phonologically related to either the first noun (art for arrow) or the second noun (ball for bag). Meyer found phonological facilitation to the first noun, but no facilitation to the second noun. These results support radical incrementality where only a minimal amount of information is necessary for articulation to begin, i.e., a phonological word.

Further support for radical incrementality has been argued from evidence from a prepared speech paradigm in Dutch when two and three phonological word sentences were produced (Wheeldon & Lahiri, 1997). Participants were instructed to memorize a visually presented noun phrase (e.g., *fresh water*), listen to an auditory question (e.g., *what do you seek*) and after a variable delay, produce a prepared response (e.g., *I seek fresh water*). Response times were assumed to reflect preparation of an entire abstract phonological representation, that is, a phonological representation produced after syntactic structure is determined but before the motor plans corresponding to the syllables to be produced (phonetic encoding) occurs. The influence of the number of phonological words and their complexity was examined, while the number of syllables and content words was held constant. Three phonological word phrases were produced more slowly than two phonological word phrases. These results show that when memorized utterances are produced, phonological planning depends on the number of phonological words in the utterance. In a separate experiment when memorized utterances were produced immediately, planning encompassed the first phonological word, but whether planning extended to two phonological words was not tested.

In contrast to the above results, the activation of phonological properties of words outside the first phonological word was seen to affect the onset of articulation during noun phrase production. Using the picture–word interference paradigm, Costa and Caramazza (2002) found phonological facilitation to both the adjective and noun in both English and Spanish noun phrases (e.g., *the red car* or *el coche rojo*). The words are part of the first and second phonological words. Using the same paradigm, Jescheniak *et al.* (2003) found phonological activation of the object name in noun phrase production in German when it was part of the second and third phonological word (although the effect varied between phonological facilitation and interference). Using a simple picture-naming paradigm, phonological properties outside the first phonological word was seen to affect onset of articulation in noun phrase production in French and Italian (Alario & Caramazza, 2002; Miozzo & Caramazza, 1999).

Thus, evidence for the extent of phonological planning is mixed. The production of single phrases showed evidence that the level of activation of the phonological properties of words belonging to the second and third phonological word affected the onset of articulation (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Miozzo & Caramazza, 1999). When multiple noun phrases were produced, only activation of phonological properties of words in the first phonological word affected articulation, consistent with *radical incrementality* (Meyer, 1996; Wheeldon & Lahiri, 1997).

The difference in the extent of phonological planning between the above sets of experiments when the picture–word interference paradigm was used may be due to the number of phrases produced. Two phrases were produced in Meyer's (1996) study, versus only one phrase in other studies (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Miozzo & Caramazza, 1999). Phonological facilitation may not have occurred for the second object in Meyer (1996) because it occurred across a phrase boundary, in the second noun phrase. It is thus possible that phonological planning extended across multiple phonological words because an entire phrase was planned, as opposed to one or more phonological words.

In the experiments presented here, we addressed whether the difference in previous results was due to the number of phrases produced, e.g., whether the presence of a phrase boundary influences phonological planning. Consistent with the design of the experiments outlined above, we manipulated a phrase boundary defined both syntactically and phonologically.[6] The extent of phonological planning is addressed in terms of phonological representations, e.g., prosodic units under the assumption from modular production models that syntactic structure does not impact processing at the phonological level (Levelt *et al.*, 1999). Thus, we assume here that phonological phrase boundaries are relevant to planning at the phonological level, consistent with previous work. However, whether syntactic structure impacts phonological planning independent of prosodic structure is a separate empirical question.

Because the extent of phonological planning has primarily been addressed in the production of single words and concatenated phrases, the experiments here further explore phonological planning when sentences are produced. Previous work by Meyer (1996) involved sentence production where the verb was always *is*, a potentially syntactically and phonologically impoverished verb (see Ferreira, 2000). Thus it is also of interest to explore phonological planning during sentence production when non-gerundial verbs are produced.

The picture–word interference paradigm and the phonological facilitation effect were employed to test whether the activation of phonological properties beyond a phonological word (PW) and phrase (P) (e.g., [The girl]$_{PW\ P}$ [jumps]$_{PW\ P}$; Experiment 1) and multiple phonological words (e.g., [[The orange]$_{PW}$ [girl]$_{PW}$]$_{P}$ [jumps]$_{PW\ P}$; Experiment 2) affected onset of articulation.

In Experiment 1, we maintained a property of Costa and Caramazza (2002) where three words were produced (e.g., the red car) while increasing the number of phrases produced. Participants produced intransitive sentences describing an actor and action (e.g., girl jumping) while producing two phrases, a noun phrase (NP) and a verb phrase (VP) (e.g., *the girl jumps*).

[6] Phonological phrases are created from groups of phonological words and are derived from syntactic structure. Specifically, all phonological words that fall within a major grammatical phrase up to the phrase's right boundary are grouped together to form a phonological phrase (the X-max algorithm, Selkirk, 1986; for discussion of definitions of a phonological phrase, see Levelt, 1989; Selkirk, 1986). Phonological phrases are always syntactic phrases; however, the reverse is not always the case (e.g., *She kicks the ball* is considered one phonological phrase but is comprised of several grammatical phrases).

EXPERIMENT 1

In Experiment 1 we tested whether the activation of the phonological properties associated with the verb in a sentence affected articulation onset for sentences like *The girl jumps* where the verb was part of the second phonological word and phrase ([The girl]$_{PW\,P}$ / [jumps]$_{PW\,P}$). Phonologically related distractors to the verb were displayed in comparison to unrelated distractors and a baseline condition of a string of XXXs. The baseline condition was included to verify that the presentation of word distractors affected response times. Participants described pictures depicting four different actors (men, women, boys, and girls) performing intransitive actions. If response times were faster in the phonological condition in comparison to the unrelated condition, this was taken to indicate that the level of activation of the phonological properties of the verb and any preceding content had an impact on the onset of articulation. If articulation followed phonological encoding at a distance of a phonological word (radical incrementality) or only for a single phrase when sentences were produced, then we did not expect response times to differ when naming with a related versus unrelated distractor word. Evidence from Meyer (1996), where no phonological effects were found beyond the first phrase and phonological word, supports this prediction. However, if phonological planning occurred for more than one phonological word and phrase, then we expected participants to name pictures more rapidly in the context of a phonologically related word in comparison to an unrelated word. As a control to verify that participants were processing the word distractors, we expected participants to be slower to name pictures in the presence of unrelated words, in comparison to the baseline condition of a string of XXXs.

Method

Participants

Sixteen Harvard University undergraduate students participated. They were paid or received credit for an introductory psychology course. All were native English speakers. None participated in other experiments.

Materials

Twenty-eight line drawings depicting actions were used as target stimuli (modified from the materials used in Masterson & Druks, 1998) (see Appendix A). All pictures depicted an actor performing an intransitive action. An actor was depicted as either a boy, girl, man, or woman, so that seven of the 28 actions fell into each category. Each picture was presented with four distractor words: (a) phonologically related to the verb (e.g., rust for run); (b) phonologically unrelated to the verb (e.g., shawl for run); (c) a baseline condition (a string of 6 Xs printed inside each picture); and (d) a filler condition. The filler condition was included to reduce the percentage of trials where the distractor and picture were phonologically related to 25% of the total trials. The filler condition included unrelated distractors and was not analyzed. The pictures and the distractors were paired so that each distractor appeared once in the phonologically related and once in the unrelated condition. For example, the picture *the girl jumps* appeared once with the phonologically related distractor *jug*, and once with the unrelated distractor *sneer*. The same distractors were displayed for the picture *the man sneezes*. Except in this case, *sneer* was the phonologically related distractor, and *jug* was the unrelated distractor. The distractors paired with each picture had similar frequencies (Francis & Kucera, 1982). Distractors were chosen so that they did not sound similar to the agent of the sentence except in two cases due to experimenter error. Phonologically related distractors shared the first two segments with the verb of the picture. Because it was difficult to select distractors that satisfied all the above conditions, distractor grammatical class was not controlled for.

The distractors were shown in 28-point boldface capital letters in Geneva font, superimposed on the pictures. Pictures were centered at fixation. Word position varied randomly in the region around fixation to prevent participants from systematically fixating the portion of the picture not containing the distractor. However, for an individual picture, the position of all of its distractors was the same. Distractors and pictures were presented simultaneously.

The experimental stimuli were presented in four different blocks of 32 trials each (28 experimental trials and four warm-up trials) for a total of 128 trials. Each picture was presented once per block. At the beginning of each block, four pictures were included as warm-up trials. The trials were randomized so that (a) the same picture did not occur twice in the same block; (b) the same distractor condition occurred no more than three times in a row; (c) no agent occurred more than twice in a row; and (d) no phoneme in the subject or verb was the same from trial to trial. Four different block orders were designed and presented to participants according to a Latin-square design.

Before the experiment proper, participants had two practice series. In the first series participants were presented with all the pictures with a series of Xs printed inside each picture, to train the subject to use the correct name for each picture. In the second practice series, they were presented with all the pictures with practice distractors printed inside every picture. These practice distractors were not used during the experiment.

Apparatus

The pictures were presented on a Macintosh using the PsychLab program (Bub and Gym, University of Victoria, British Columbia, Canada). Response times (RTs) were measured to the nearest millisecond by means of a voice key (KOSS headset/CMU voicebox) from appearance of the picture until the voice key was triggered.

Procedure

Participants were asked to name the picture using a full sentence (e.g., the girl jumps). Participants were tested individually in a darkened testing room. They were instructed to name pictures "like they normally speak" and as accurately as possible. When participants made mistakes during the practice session, they were asked to name the picture correctly. Each trial proceeded as follows: A fixation point (+) was shown for 700 ms, followed by presentation of the stimulus 300 ms later. Pictures remained on the screen until the microphone was triggered. There was a 2000 ms pause between trials. The experimenter remained in the testing room in order to record incorrect responses and when voice key malfunctions occurred. A session lasted approximately 25 min.

Analyses

Three types of responses were classified as errors: (a) production of the wrong name; (b) verbal disfluencies (stuttering, utterance repairs, etc.); and (c) voice key malfunctions. Responses faster than 300 ms and three SDs from a participant's condition mean were also eliminated. Separate ANOVAs were carried out on the errors and response times using either the means per subject or means per item as dependent variables yielding $F1$ and $F2$ statistics, respectively. The variable Type of distractor was analyzed with two levels for all ANOVAs: either phonologically related and unrelated to test the phonological facilitation effect, or unrelated and baseline, to test the effectiveness of the presentation of a word distractor. Type of distractor was considered a within-subject and within-item variable.

Results and Discussion

Table I reports a summary of the data. The naming latencies from one item were removed because it elicited a high percentage of errors (more

Table I. Experiment 1. Mean Response Times (ms), Standard Deviations (SD) and Percentage of Errors (Error %), for Phonologically Related, Unrelated, and Baseline Conditions

Type of distractor	Mean	SD	Error %
Phonologically related	792	83	8.7
Phonologically unrelated	818	91	12.5
Baseline (XXXs)	784	76	7.4
Phonological effect (Related—Unrelated)	−26*		

A significant difference of $p < 0.05$ is indicated by an *.

than 30%). Error rates consisted of 8.7% of the data before outliers were removed and 9.5% of the data after outliers were eliminated. The difference in error rates between phonologically related and unrelated conditions reached near significance [$F1(1, 15) = 9.23$, MSE = 0.2962 $p = 0.008$; $F2(1, 26) = 3.82$, MSE = 0.2962, $p = 0.06$] and error rates were significantly higher for the unrelated versus baseline condition [$F1(1, 15) = 9.13$, MSE = 0.5601, $p = 0.008$; $F2(1, 26) = 11.08$, MSE = 0.5601, $p = 0.002$]. The pattern does not reflect a speed-accuracy trade-off as errors were higher when RTs were slower for both comparisons.

Response times were faster in the phonologically related condition (792 ms) versus the unrelated condition [818 ms; $F1(1, 15) = 6.29$, MSE = 136527, $p = 0.02$; $F2(1, 26) = 5.08$, MSE = 142441, $p = 0.03$]. The baseline condition produced faster naming latencies (784 ms) in comparison to the unrelated condition [818 ms; $F1(1, 15) = 16.32$, MSE = 240646, $p = 0.001$; $F2(1, 26) = 8.62$, MSE = 228128, $p = 0.006$].[7]

We additionally examined whether the phonological facilitation effect changed between the first half and second half of the experiment to test whether participants developed a strategy whereby they intentionally focused their attention on the relationship between distractor and verb, delaying onset of articulation until the verb was phonologically encoded. If this were the case we would expect the facilitation effect to be larger during the second half of the experiment as participants learned the relationship between distractor and target. Response times in the first half

[7] In order to test whether the chosen distractors could reliably produce an effect independent of sentence context, a separate group of participants named the same items, but named the verb in isolation using the 3rd person singular form of the action name (e.g. jumps). When participants named verbs alone, response times were significantly faster in the phonologically related condition in comparison to the unrelated condition [$F1(1, 9) = 10.97$, MSE = 264348, $p = 0.009$; $F2(1, 27) = 13.79$, MSE = 297477, $p < 0.001$].

of the experiment were on average 39 ms longer than in the second half which resulted in a significant effect of experiment half ($F1(1, 15) = 20.51$, MSE = 182620, $p < 0.0001$). Although experiment half interacted significantly with condition ($F1(1, 15) = 13.58$, MSE = 120915, $p < 0.01$) the phonological facilitation effect decreased between the first (−51 ms) and second half of the experiment (−4 ms) suggesting that participants were more susceptible to the distractor effect when they had been less exposed to the experiment, which does not support a strategic interpretation to the phonological facilitation effect seen here. Additionally, related words were presented on only 25% of all trials so that three-quarters of the time there was no relationship between distractor and verb, limiting the usefulness of a strategy of focusing on the distractor–verb relationship. Elsewhere it has been shown that the magnitude of the facilitation effect does not change whether phonologically related distractors comprise 50% or 25% of total trials, suggesting that the facilitation effect is not a strategic one (Meyer & Schriefers, 1991).

The experiment showed that production of a sentence (e.g., the girl jumps) was facilitated with a distractor phonologically related to the verb (jug) compared to an unrelated distractor (sneer). Assuming that the phonological facilitation effect reflects processes at the level of phonological encoding, the phonological properties of the verb, part of the second phonological word and phrase were active before articulation. This result is compatible with results from Costa and Caramazza (2002) and Jescheniak *et al.* (2003), where phonological facilitation was observed for the second phonological word for noun phrases like *the red car*. This result is also compatible with evidence from noun phrase production, where phonological planning extended two phonological words (Alario & Caramazza, 2002; Miozzo & Caramazza, 1999; Roelofs, 1998). Experiment 1 extends these results by showing that phonological planning extends beyond phrase boundaries. The presence of a phrase boundary does not necessarily limit the extent of phonological planning. Put another way, articulation can begin with more than a single phonological word or phrase phonologically encoded.

These results are inconsistent with results of sentence production from Meyer (1996) as more than one phonological word was planned in advance of articulation. The lack of phonological facilitation to the second object in the second phrase in sentences like *de pijl en de tas* (the arrow and *the bag*; Meyer, 1996) may have been due to the number of phonological words in these utterances as opposed to the presence of a phrase boundary. The utterance, *de pijl en de tas* may be produced in two phonological words, $[\text{de pijl en}]_{\text{PW}}$ $[\text{de tas}]_{\text{PW}}$, with the main stress on *pijl* (arrow) and *tas* (bag). However, it is possible that participants provided extra stress on the function word *en* (and), creating a three phonological word utterance, [de pijl]$_{PW}$ [EN]$_{PW}$ [de tas]$_{PW}$. The second object *de tas* (the bag), would have been part of the third phonological word. Did Meyer (1996) fail to find phonological facilitation to the second object because *the bag* was part of the third phonological word? Jescheniak *et al.* (2003) found when complex noun phrases like *the big red dog* were produced, phonologically related distractors to the object of the noun phrase (and third phonological word) produced interference, not facilitation. Thus, we tested whether we could obtain a phonological facilitation effect to the second phrase and third phonological word of a sentence where previous results have found either no significant effects (Meyer, 1996) or interference (Jescheniak *et al.,* 2003).

EXPERIMENT 2

In order to test whether phonological planning extends to three phonological words and second phrase in an utterance, in this experiment participants described the same pictures as in Experiment 1, but they also named the color the person was drawn in, e.g., *The orange girl jumps*. The verb in the sentence is part of the third phonological word and second phrase (e.g., [[the orange]$_{PW}$ [girl]$_{PW}$]$_{P}$ [jumps]$_{PW\ P}$). If phonological planning extends to three phonological words during sentence production, then we expected to see the same pattern of results seen in Experiment 1—phonological facilitation to the verb. However, if phonological planning does not extend beyond the second phonological word we did not expect facilitation effects to the verb.

Method

Participants

Twenty-one Harvard University undergraduate students were paid, or received course credit for their participation. All were native English speakers. None participated in other experiments.

Materials

The same materials and distractors from Experiment 1 were used in this experiment. However, the actors were depicted in one of four colors: pink, red, orange, and yellow (see Appendix A). Colors were assigned to pictures so that no words in the sentence began with the same sound or rhymed. The same color was used in a single picture for all conditions except for the filler condition. In this case, an alternative color (one of the other three) was used. For example, for the picture of *the girl jumps*, the girl appeared in orange for the phonologically related, unrelated, and baseline conditions. However, when the picture was displayed in the filler condition, the picture appeared in a different color, e.g., *the yellow girl jumps*. This was to mitigate any participant strategy where a picture was memorized with regard to its color. Although unlikely considering there were 32 different pictures, by having a picture appear in a different color for 25% of the total trials, we thought this would ensure spontaneous sentence description. All colors were named an equal number of times.

The experimental stimuli were presented in four different blocks, for a total of 128 trials (112 experimental trials and 16 filler trials), as in Experiment 1. The trials were randomized as in Experiment 1 with the further restriction that no color occurred more than twice in a row.

Procedure

Participants were asked to name the picture using a full sentence (e.g., the orange girl jumps). All other aspects of the experiment were the same as Experiment 1.

Results and Discussion

Table II reports a summary of the data. The naming latencies from one item and subject were removed because of a high percentage of errors (more than 30%). Because of experimenter error, the sixth replication of the Latin-square design was not completed.

Table II. Experiment 2. Mean Response Times (ms), Standard Deviations (SD) and Percentage of Errors (Error %), for Phonologically Related, Unrelated, and Baseline Conditions

Type of distractor	Mean	SD	Error %
Phonologically related	861	139	12.9
Phonologically unrelated	898	169	18.7
Baseline (XXXs)	857	159	13.5
Phonological effect (Related–Unrelated)	−37		

A significant difference of $p < 0.05$ is indicated by an *.

Error rates consisted of 14.2% of the data before outliers were removed and 15% of the data after outliers were eliminated.[8] Error rates were significantly different between phonologically related and unrelated conditions [$F1(1, 19) = 10.50$, MSE $= 0.8898$, $p = 0.004$; $F2(1, 26) = 7.75$, MSE $= 0.8898$, $p = 0.009$]. The difference between errors rates for unrelated and baseline conditions was marginally significant [$F1(1, 19) = 4.07$, MSE $= 0.7259$, $p = 0.06$; $F2(1, 26) = 4.87$, MSE $= 0.7259$, $p = 0.04$)]. As in Experiment 1, there was no speed-accuracy tradeoff.

Response times were faster in the phonologically related condition (861 ms) than the unrelated condition [898 ms; $F1(1, 19) = 10.56$, MSE $= 469104$, $p = 0.004$; $F2(1, 26) = 5.86$, MSE $= 31306$, $p = 0.02$]. The baseline condition produced faster naming latencies (857 ms) in comparison to the unrelated condition [898 ms; $F1(1, 19) = 15.89$, MSE $= 469104$, $p < 0.001$; $F2(1, 26) = 4.40$, MSE $= 326070$, $p = 0.04$]. We examined whether the phonological facilitation effect increased in magnitude from the first half to the second half of the experiment as in Experiment 1. Response times were 49 ms slower in the first half than second half of the experiment which resulted in a significant effect of experiment half ($F1(1, 15) = 17.14$, MSE $= 524735$, $p < 0.001$). However, experiment half did not interact with condition ($F1 < 1$).

As in Experiment 1, phonologically related distractors to the verb facilitated sentence production compared to unrelated distractors. Here however, the phonological properties corresponding to the verb were part of the third phonological word, and second phrase. Experiment 2 showed that in the case of sentence production, phonological planning extends three phonological words, and two phrases. The results from Experiment 2 again suggest that phrase boundaries do not play a role in the extent of phonological planning. The lack of effect to the second object in Meyer (1996) cannot have been due to the object being outside the extent of phonological planning if that extent is defined by a phrase *or* phonological word boundary. The direction of the phonological effect (facilitation) in Experiment 2 was in contrast to the inhibition seen by Jescheniak *et al.* (2003). In the General Discussion we will discuss possible reasons for the varying effects in phonological planning.

[8] Although error rates are higher than in Experiment 1, they are within the range seen in multi-word production studies (an average 14% error reported by Meyer (1996) and an average of 9.6% reported by Jescheniak *et al.* (2003)).

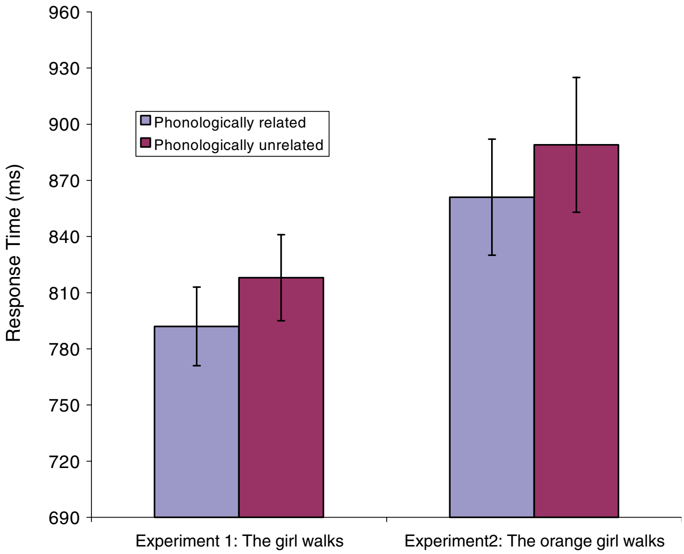

Fig. 1. Experiments 1 and 2: Response times and standard error bars for sentences where distractors phonologically related and unrelated to the verb were displayed, when the verb was part of the second phrase.

GENERAL DISCUSSION

The principal finding of the work presented here is that during sentence production phonological planning crosses phonological word and phrase boundaries extending several items in advance of speaking. In two separate experiments participants were asked to name two-phrase sentences (NP + VP) of two phonological words (e.g., [The girl]$_{PW}$ [jumps]$_{PW}$) or three phonological words (e.g., [The orange]$_{PW}$ [girl]$_{PW}$ [jumps]$_{PW}$). Participants were faster to produce both types of sentences in the presence of a phonologically related distractor to the verb in comparison to an unrelated distractor. The results of Experiments 1 and 2 (see Fig. 1) extend previous evidence by demonstrating that modification of the level of activation of phonological properties of the second phrase and second and third phonological word in utterances affected the onset of articulation when sentences were produced. This evidence is not consistent with *radical incrementality* where articulation follows phonological encoding at a distance of one phonological word. These results suggest that several items are buffered instead of the buffer being "minimal or absent" (Levelt, 1989).

Previous investigations of phonological planning have yielded mixed results. When *multiple phrases* were produced, phonological planning extended *one* phonological word (Meyer, 1996). Similarly, evidence from a prepared speech paradigm where two and three phonological word utterances were produced has also been used to support the phonological word as the preferred extent of planning during sentence production (Wheeldon & Lahiri, 1997). However, when one phrase consisting of two or three phonological words was produced, phonological planning extended to all phonological words (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Miozzo & Caramazza, 1999). We tested whether differences in results were due to the number of phonological words or phrases. Differences between results cannot be due to the presence of phrase boundaries or the number of phonological words produced. The results from Experiments 1 and 2, where the extent of phonological planning included the second phrase and second and third phonological words, rule out this possibility.

Why the discrepancy between results showing phonological planning of one phonological word versus more than one word before articulation begins? In the following we discuss two possible explanations. One concerns the fact that for the experiments here *verbs* were phonologically facilitated, and verbs may play a special role during grammatical and therefore phonological encoding. The second possibility concerns a qualitative difference when naming pictures that depict integrated versus discrete entities.

In the Experiments presented in this paper distractors were displayed related to the *verb* in comparison to distractors being related to an adjective or noun in previous work. Ferreira (2000) in her model of syntactic production suggested that the main verb of a sentence is crucial for creation of a sentence's syntactic structure. In her model, a verb must be part of the syntactic representation of the sentence before the subject NP is grammatically encoded. Thus, a main verb is lexically selected before other elements in the sentence. Elsewhere it has been suggested that a verb is selected either at the same time as a subject NP or soon afterwards (Bock, 1987; Griffin, 2000). If we assume that phonological properties of selected lexical items are automatically activated, under this framework even if the verb is not the first item to be produced, its phonological properties are activated before elements in the subject NP. By presenting a distractor phonologically similar to the verb, it is possible this may have sped up retrieval of its phonological properties, freeing up resources for the retrieval of phonological properties corresponding to the subject NP.

Although we are not aware of any specific evidence that speaks to whether verbs are selected before other elements in a sentence, other results have shown that the grammatical "importance" of the primed item in the picture–word interference task does not independently account for phonological facilitation effects. Costa and Caramazza (2002) asked participants to produce noun phrases in both English (the red car) and Spanish (el coche roja). Participants were faster to produce the noun phrase in English when a distractor was phonologically related to the word *car*. The phonological facilitation effect may have been seen for car in English because *car* was the lexical head of the noun phrase. To address this issue they also had participants name the same pictures in Spanish, where the word order is reversed. In Spanish, the adjective *roja* occurs after the noun in *el coche roja*. If *car* were phonologically active only because it is the lexical head of the noun phrase then one would *not* expect to see phonological facilitation to an item after the lexical head. Results showed in Spanish that both *car* and *red* were phonologically active. These results suggest that priming a lexical head of phrase phonologically does not alone account for the extent of phonological planning.

A verb could be conceptually or grammatically important independent of its status as a lexical head of phrase and thus it is possible this is why we found phonological facilitation for the verb when it was part of the second and third phonological word. However, some evidence suggests that planning is not determined by the grammatical status of the primed element during sentence production. Schriefers *et al.* (1998) examined the semantic interference effect for verbs during sentence production using the picture–word interference task in German. Semantic interference effects were only found when a verb occupied the first position in an utterance, and only when it was transitive. Schriefers *et al.* interpreted these results to suggest that articulation of a sentence was not dependent on the lexical selection of the verb of the sentence. Unfortunately the lack of a semantic interference effect is hard to interpret because semantic interference effects to verbs when produced in isolation have been inconsistent (Schnur *et al.*, 2002). In sum, although the grammatical importance (i.e., lexical head of phrase) does not appear to account for the extent of phonological planning seen in the experiments presented here, it is not clear from current evidence whether the conceptual or grammatical importance of verbs was a factor.

A second difference between these experiments was the number of distinct entities named e.g., two in Meyer (1996) (the *arrow* and the *bag*) versus one elsewhere (the *red car*; the *girl kicks*). Although this distinction is not fully formulated, intuitively the concepts depicted in pictures of a girl jumping or a red car are integrated in a way that *arrow* and *bag* are not. The action is not independent of the actor, or the color independent of the object. The difference in extent of phonological planning might be due to a difference in the number of distinct entities named. The extent of phonological planning depends on, among other factors, its input from conceptual and grammatical encoding. In picture naming, articulation may not follow phonological encoding at a distance of multiple phonological words. Instead the amount phonologically planned may correspond to the number of phonological words that describe a single conceptual unit. Articulation may begin when the first conceptual unit is fully phonologically planned. This may be another possibility for why the experiments presented here pattern with those of Costa and Caramazza (2002) and Jescheniak *et al.* (2003) in comparison to Meyer (1996). Evidence from eye gaze durations where two objects are named sequentially suggests that speakers processed the left-most object until its grammatical and/or phonological properties were retrieved before accessing information about the right-most object (Griffin, 2001; Meyer *et al.*, 1998; Meyer & van der Meulen, 2000). In this case, single discrete entities were processed sequentially. A future experiment to further understand whether phonological planning proceeds one distinct entity at a time, could examine eye-gaze durations while complex NPs or sentences are being produced.

Other results suggest different factors to explain the disparate results (Jescheniak *et al.*, 2003). As mentioned previously, Jescheniak *et al.* (2003) conducted a picture–word interference experiment where participants described pictures in German with noun phrases of varying complexity. For example, participants either produced *Hund* (dog), *der hund* (the dog), or *der groBe rote hund* (the big red dog), referred to as, respectively, bare noun, simple noun phrase, and complex noun phrase utterance formats. Auditory distractors phonologically related to the noun of the utterances were presented at several different SOAs including 0 ms. Their results showed that more than one phonological word was phonologically encoded before articulation began, in accordance with our results and others (Alario & Caramazza, 2002; Costa & Caramazza, 2002; Miozzo & Caramazza, 1999). However, as the target item occurred later in the sentence, the magnitude of the phonological facilitation effect changed. For the SOA of 0 ms, Jescheniak *et al.* found that the phonological facilitation effect was largest for bare noun naming, reduced in simple noun phrase naming, and was inhibitory during complex noun phrase production. In a post hoc analysis, Jescheniak *et al.* divided the items into two groups, items that produced smaller facilitation effects in bare noun naming versus items that produced larger facilitation effects. They found that items that produced a small facilitation effect when nouns were produced in isolation (27 ms) produced statistically significant inhibition for complex NP utterances (41 ms). In contrast, items that produced a "large" phonological facilitation in isolation (71 ms) showed a non-significant inhibition effect (16 ms) when the items were produced in complex NP utterance. Jescheniak *et al.* suggested that the "base priming effect" and the relative position of the word in the utterance predicted the magnitude of phonological facilitation. It was proposed that when multi-word utterances are produced, each word receives graded phonological activation whose strength corresponds to position in the utterance, e.g., more activation for earlier positions and less activation for later positions. Disparate results from Meyer (1996), Costa and Caramazza (2002) and others were explained in terms of both the base priming effect, and relative position of the primed item in the utterance.

Although our results are in accordance with Jescheniak *et al.* (2003) where phonological planning extended to the third phonological word, they do not show the predicted pattern of inhibition for the targeted item as it moved to later positions in the utterance. In a pre-test experiment to ensure that the phonologically related distractors produced phonological facilitation in comparison to unrelated distractors when the verb was named in isolation, we found a facilitation effect of 42 ms. However, instead of increasing degrees of inhibition (e.g., less facilitation) as the verb was produced in later positions in the utterance, we found phonological facilitation effects of 26 and 37 ms, for the simple NP + V and complex NP + V utterances. A post hoc analysis showed that the size of the phonological facilitation effect did not interact with the position of the verb in the utterance (alone, second phonological word, third phonological word; $F1$ and $F2 < 1$).

We also did a post hoc analysis to see whether a "base priming effect" for each item could predict the amount of phonological facilitation when the item occurred in complex NP + V utterances (e.g., *the orange girl jumps*). We divided the items evenly based on whether the produced larger or smaller amounts of facilitation in isolation (−98 and + 12 ms on average of facilitation). "Larger" facilitation items produced −46 ms of facilitation and "smaller" facilitation items produced −24 ms of facilitation when they were produced as the third phonological word, and latest position in the utterance. The base priming effect of items in isolation (large versus small) did not interact with the phonological facilitation effect for *the orange girl jumps* ($F1$ and $F2 < 1$). It is not clear whether both "position in utterance" and "base priming effect" predict phonological facilitation effects. However, our experiments were not specifically designed to address these questions and a further exploration of Jescheniak *et al.* model should include investigation of the magnitude of phonological facilitation during sentence production, as opposed to single NP production.

Whether the phonological facilitation effect reflects lexical selection or phonological encoding depends on assumptions made about processing during word production. Let us assume that the phonologically related distractor activates the phonological representation of the word to be produced. If *interactivity* is assumed, where phonological information sends activation back to the grammatical level (i.e., activation flows in two directions: Dell, 1986; Dell *et al.*, 1997; Stemberger, 1985), then the phonological facilitation effect may reflect the influence of the distractor at the grammatical level where the target item is lexically selected. For example, when a picture is shown with a phonologically related distractor, speech could be accelerated because the appropriate target picture name received feedback from its corresponding phonological segments, which were activated by the distractor. This feedback would facilitate both *lexical selection* and the *phonological representation* corresponding to the target word. In contrast, when no interactivity is assumed between phonological and grammatical encoding levels, then the facilitation effect can be assumed to be solely a phonological one. For example, if processing during word production is strictly *discontinuous and serial* (activation only flows in one direction from grammatical to phonological levels; Garrett, 1980; Levelt, 1989, 1992; Levelt *et al.*, 1991; Roelofs, 1992, 1997; Schriefers *et al.*, 1990) or *cascaded* (processing is continuous in one direction where phonological representations corresponding to more than one word are activated during word production; Caramazza, 1997; Morsella & Miozzo, 2002; Peterson & Savoy, 1998) a phonologically related distractor can only influence processing at the phonological encoding level. This is an area of much debate (see Dell & O'Seaghdha, 1992; Levelt *et al.*, 1991, 1999) and there is different evidence to support each position (see Rapp & Goldrick, 2000; Ruml *et al.*, 2000; Starreveld, 2000). Exactly how phonological facilitation takes place during phonological encoding is also still debated (see Starreveld, 2000 for a discussion on this point). However, whichever mechanism one posits to account for phonological facilitation, all involve activation of the phonological properties associated with the target word. Although the phonological facilitation effect may reflect some influence at the level of lexical selection, we assume the bulk of the effect lies at the level of phonological encoding in accordance with the general view of the picture–word interference paradigm.

In sum, if the phonological facilitation effect primarily reflects activation at the phonological encoding level, then the experimental evidence to date does not support the role of the phonological word as a preferred phonological planning unit. Phonological planning is not restricted to single phrases and appears to vary, extending from one to several phonological words and phrases. It has been argued that there is no fixed unit of phonological (Schriefers & Teruel, 1999) or conceptual/grammatical planning (Ferreira & Swets, 2002). Ferreira and Swets (2002) proposed that language speakers are "strategically incremental", in that they weigh how much speech to plan in advance with how quickly they can begin speaking. Strategic incrementality may partially result from a strategy where participants name pictures using different response deadlines (Lupker *et al.*, 1997; Meyer *et al.*, 2003). Schriefers and Teruel (1999) suggested that the amount phonologically planned depends on the utterance produced, and the experimental and speaking demands. We suggest that the extent of phonological planning when measured during picture naming might depend on the number of integrated conceptual units named where articulation begins after the phonological properties corresponding to an integrated conceptual unit are activated. It is also possible that priming verbs phonologically during sentence production may extend the extent of phonological planning. The results presented here show that phonological planning is not restricted by the presence of a phrase boundary and extends multiple words and phrases. These results in conjunction with past evidence do not support a fixed unit of phonological encoding.

APPENDIX A

Appendix A. Stimuli for Experiments 1 and 2. Frequency (FREQ), letter length (LTTR), and number of syllables (SYLL) of the written distractors paired with the target pictures

PICTURES			Phonologically related				Unrelated			
COLOR	AGENT	VERB		FREQ	LTTR	SYLL		FREQ	LTTR	SYLL
yellow	man	BEG	bend	50	4	1	plate	44	5	1
yellow	woman	CRY	crisis	102	6	2	fill	184	4	1
orange	woman	DANCE	dam	5	3	1	swig	2	4	1
orange	man	DIG	dish	36	4	1	shade	39	5	1
Pink	man	DIVE	dice	1	4	1	sleet	1	5	1
yellow	boy	DREAMS	dress	67	5	1	sick	51	4	1
pink	boy	FISH	fill	184	4	1	crisis	102	6	1
orange	girl	JUMP	jug	6	3	1	sneer	3	5	1
red	boy	LAUGH	ladder	19	6	2	yarn	20	4	1
yellow	boy	PLAY	plate	44	5	1	bend	50	4	1
red	woman	POINT	poise	12	5	1	sizzle	5	6	1
yellow	woman	PRAY	praise	21	6	1	scare	26	5	1
orange	woman	RUN	rust	7	4	1	shawl	5	5	1
yellow	man	SHAVE	shade	39	5	1	dish	36	4	1
orange	man	SHOOT	shawl	5	5	1	rust	7	4	1
orange	boy	SING	sick	51	4	1	dress	67	5	1

Appendix A. Continued

PICTURES			Phonologically related				Unrelated			
COLOR	AGENT	VERB		FREQ	LTTR	SYLL		FREQ	LTTR	SYLL
pink	girl	SIT	sizzle	5	6	1	poise	12	5	1
red	girl	SKATE	scare	26	5	1	praise	21	6	1
pink	woman	SKI	scheme	39	6	1	waste	31	5	2
pink	man	SLEEP	sleet	1	5	1	dice	1	4	2
pink	boy	SLIDE	slice	12	5	1	walnut	16	6	1
red	man	SNEEZE	sneer	3	5	1	jug	6	3	1
red	woman	SWIM	swivel	2	6	2	wig	1	3	1
yellow	girl	SWING	swig	2	4	1	dam	5	3	1
orange	girl	WALK	walnut	16	6	1	slice	12	5	1
pink	girl	WAVE	waste	31	5	1	scheme	39	6	1
red	boy	WINK	wig	1	3	1	swivel	2	6	1
red	girl	YAWN	yarn	20	4	1	ladder	19	6	1

REFERENCES

Alario, F. X., & Caramazza, A. (2002). The production of determiners: Evidence from French. *Cognition, 82*(3), 179–223.

Bock, J. K. (1987). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the Psychology of Language* (Vol. 3, pp. 337–390). London: Erlbaum.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology, 14,* 177–208.

Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language, 46*(1), 178–198.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*(1), 123–147.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition, 42,* 287–314.

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior, 20,* 611–629.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheeldon, (Ed.), *Aspects of Language Production. Studies in Cognition Series.* (pp. 291–330). Philadelphia, PA, US: Psychology Press/Taylor & Francis.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language, 46*(1), 57–84.

Francis, W. N., & Kucera, H. (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language Production* (Vol. 1, pp. 177–220). London: Academic Press.

Glaser, W. R., & Dungelhoff, F. J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118*(1), 13–42.

Griffin, Z. M. (2000). *How Early are Verbs Selected in Extemporaneous English Sentence production?* Paper presented at the Architectures and Mechanisms of Language Processes conference, Leiden, The Netherlands.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1–B14.

Jescheniak, J. D., Schriefers, H., & Hantsch, A. (2003). Utterance format affects phonological priming in the picture–word task: Implications for models of phonological encoding in speech production. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 441–454.

Levelt, W. J., Schriefers, H., Vorberg, D., Meyer, A. S., & *et al.* (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review, 98*(1), 122–142.

Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M. (1992). Accessing words in speech production: Stages, processes and representations. *Cognition, 42,* 1–22.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary? *Cognition, 50,* 239–269.

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

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology, 36,* 349–367.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 570–590.

Masterson, J., & Druks, J. (1998). Description of a set of 164 nouns and 102 verbs matched for printed word and frequency, familiarity and age-of-acquisition. *Journal of Neurolinguistics, 11*(4), 331–354.

Meyer, A. S. (1992). Investigation of phonological encoding through speech error analyses — achievements, limitations, and alternatives. *Cognition, 42*(1–3), 181–211.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture – word interference experiments. *Journal of Memory and Language, 35,* 477–496.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language, 48,* 131–147.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture–word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 17*(6), 1146–1160.

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

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

Miozzo, M., & Caramazza, A. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning Memory and Cognition, 25*(4), 907–922.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(3), 555–563.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 24,* 539–557.

Rapp, B. C., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*(3), 460–499.

Rayner, K., & Springer, C. J. (1986). Graphemic and semantic similarity effects in the picture–word interference task. *British Journal of Experimental Psychology, 77,* 207–222.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1–3), 107–142.

Roelofs, A. (1997). Syllabification in speech production: Evaluation of WEAVER. *Language and Cognitive Processes, 12*(5–6), 657–693.

Roelofs, A. (1998). Rightward incrementality in encoding simple phrasal forms in speech production: Verb–particle combinations. *Journal of Experimental Psychology: Learning Memory and Cognition, 24*(4), 904–921.

Ruml, W., Caramazza, A., Shelton, J. R., & Chialant, D. (2000). Testing assumptions in computational theories of aphasia. *Journal of Memory and Language, 43,* 217–248.

Schnur, T., Costa, A., & Caramazza, A. (2002). Verb production and the semantic interference effect. *Journal of Cognitive Science, 3*(1), 1–26.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory and Language, 29*(1), 86–102.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology, 11*(1), 17–50.

Schriefers, H., Teruel, E., & Meinshausen, R. M. (1998). Producing simple sentences: Results from picture–word interference experiments. *Journal of Memory and Language, 39*(4), 609–632.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook, 3,* 371–405.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language, 42*(4), 497–525.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 896–918.

Stemberger, J. P. (1985). An interactive activation model of language production. In A. Ellis (Ed.), *Progress in the Psychology of Language* (pp. 143–186). London: Erlbaum.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37*(3), 356–381.