Planning in sentence production: Evidence for the phrase as a default planning scope

Randi C. Martin [a,*], Jason E. Crowther [a], Meredith Knight [a], Franklin P. Tamborello II [a], Chin-Lung Yang [b]

[a] Department of Psychology, Rice University, Houston, TX, USA
[b] Department of Chinese, Translation and Linguistics, City University of Hong Kong, Kowloon, Hong Kong

ARTICLE INFO

ABSTRACT

Controversy remains as to the scope of advanced planning in language production. Smith and Wheeldon (1999) found significantly longer onset latencies when subjects described moving-picture displays by producing sentences beginning with a complex noun phrase than for matched sentences beginning with a simple noun phrase. While these findings are consistent with a phrasal scope of planning, they might also be explained on the basis of: (1) greater retrieval fluency for the second content word in the simple initial noun phrase sentences and (2) visual grouping factors. In Experiments 1 and 2, retrieval fluency for the second content word was equated for the complex and simple initial noun phrase conditions. Experiments 3 and 4 addressed the visual grouping hypothesis by using stationary displays and by comparing onset latencies for the same display for sentence and list productions. Longer onset latencies for the sentences beginning with a complex noun phrase were obtained in all experiments, supporting the phrasal scope of planning hypothesis. The results indicate that in speech, as in other motor production domains, planning occurs beyond the minimal production unit.

© 2010 Elsevier B.V. All rights reserved.

1. Evidence for a phrasal scope of planning in speech production

In cognitive tasks involving motor output, be it language production, problem solving, or skilled motor performance (such as playing a musical instrument), people must execute a sequence of actions toward some goal. Crucial issues in all these cognitive domains concern the levels of representation at which advance planning takes place and the extent or scope of such planning (e.g., Catrambone, 1998; Rosenbaum, 2010; Smith & Wheeldon, 1999). There are major benefits to planning ahead – such as avoiding

mistakes. One can insure that in problem solving, for instance, one will not end up in a game position from which there is no legal move without backtracking. In piano playing one can avoid choosing fingering for the first notes of a run that would impede rapid execution of the entire arpeggio. In language production, one can avoid becoming tongue-tied because of the difficulty in finding an appropriate word or phrase to complete a thought given what has already been produced. Advance planning has its downsides as well. Advance planning at multiple levels is computationally expensive and could lead to cognitive overload. Also, planning of several production units simultaneously could lead to interference between the current unit being produced and other units that are planned and awaiting production. Thus, there are tradeoffs between planning ahead and planning at the smallest increment possible.

* Corresponding author. Address: Department of Psychology, MS-25, Rice University, Houston, TX 77005, United States. Tel.: +1 713 348 3417; fax: +1 713 348 5221.
  E-mail address: rmartin@rice.edu (R.C. Martin).

In problem solving and motor performance, there is considerable evidence for advance planning (e.g., Catrambone, 1998; Rosenbaum, 2010) and evidence for the use of hierarchical structure in advance planning (Rosenbaum, Kenny, & Derr, 1983). Thus, in these domains it appears that the benefits of advanced planning outweigh the costs. However, in these domains, the sequences to be produced may be highly practiced or there may be a limited repertoire of multi-unit structures and a small set of units to be fit into these structures during execution. In contrast, in the case of language production, there are many syntactic structures and many lexical elements that might be selected. Thus, it is possible that the downsides of advanced planning outweigh the advantages in this domain. In fact, considerable attention has been given to the issue of the extent of advance planning in speech production. In language production, it is clear that speakers are incremental to some degree, at least at the level of grammatical planning and word selection. That is, although speakers may plan at the conceptual or message level what becomes a sentence or clause, lexical items are retrieved and syntactic structure is developed online, influenced by the emphasis given to various elements in the message and the momentary availability of the words to express these elements (e.g., Bock, 1982; Chang, Dell, & Bock, 2006). Evidence suggesting a clausal scope of planning, such as longer pauses before more syntactically complex utterances (Ford, 1982), can be attributed to greater complexity of the message level representation for complex utterances (see Griffin & Ferreira, 2006, for discussion).

Although there may be general agreement about incrementality, the (sub-clausal) unit of planning in language production has been a topic of considerable debate in recent years (Costa & Caramazza, 2002; Ferreira & Swets, 2002; Griffin & Bock, 2000). Some researchers have argued that in lexical planning, speakers proceed in a word-by-word fashion (Griffin, 2003; Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998) whereas others argue that the phrase is the basic processing unit at the lexical level (Martin & Freedman, 2001; Martin, Miller, & Vu, 2004; Smith & Wheeldon, 1999).

Those taking a word-by-word approach emphasize the difficulty of word selection and the consequent need to focus attention on retrieving and beginning production of only one word at a time in order to maintain the message in mind while focusing on the next word to be produced, thereby avoiding interference in the selection of words in their proper sequence (Griffin, 2003, 2004).[1] The strongest empirical support for word-by-word planning has come from studies in which participants' eye movements were recorded while they described pictured scenes[2] (Griffin & Bock, 2000; Meyer et al., 1998). Studies using this methodology have shown a tight linkage between the timing of participants' fixation of an object in the scene and the onset of the word corresponding to the fixated object. If participants were planning several words simultaneously, one might expect that gaze durations and the time delay between object fixation and word onset would be shorter for words occurring later in the utterance. Instead, Griffin and Bock (2000) showed that gaze durations on each object and the delay between fixation and word onset were similar, irrespective of the word's position in the utterance.

Moreover, Griffin and Bock (2000) found that gaze durations and onset latencies for the first object were affected by word frequency, suggesting that speakers plan up to the level of the phonological form for the first word before moving on to plan the next word (Griffin, 2001; Meyer et al., 1998; Spieler & Griffin, 2006), under the assumption that a frequency effect reflects phonological access (Jescheniak & Levelt, 1994; Jescheniak, Meyer, & Levelt, 2003; however see Caramazza, Costa, Miozzo, & Bi, 2001). Also, Griffin (2001) and Spieler and Griffin (2006) showed that when producing a sentence beginning with a conjoined noun phrase (i.e., "The A and the B are above the C"), participants' gaze durations and onset latencies for the first noun were affected by frequency of the name for the first object, but were unaffected by encodability or frequency for the second object. Encodability, which was the likelihood of participants' using a specific name for a picture, was taken as a measure of the difficulty of accessing a lemma representation for a word from the picture. Thus, the lack of effect of encodability of the second noun on either gaze durations or onset latencies for the sentence suggested that no planning of the second noun occurred prior to onset of the first. In contrast, gaze durations on the second object were related to both encodability and frequency, suggesting that both lemma access and phonological access for the second object occurred after gaze shifted to the second object.

In contrast to the word-by-word view, the phrasal scope of planning hypothesis postulates that speakers access all of the lexical representations of the content words within a phrase prior to articulation. Those taking the phrasal view have typically assumed that the lexical representations which are accessed in phrasal planning are lemma representations, that is, lexical representations that connect lexical semantics to syntactic information but which do not contain phonological information (Allum & Wheeldon, 2007; Martin et al., 2004; Smith & Wheeldon,

---

[1] Although there may be advantages to word selection in planning one word at a time, it might seem difficult for such a process to result in a grammatical sequence of words. However, the recent Chang et al. (2006) model is a word-by-word production model which uses past learning of legal syntactic sequences to ensure grammatical production. For instance, in this model, knowledge that an entity (e.g., child) to be described is definite and knowledge that in English determiners precede nouns will lead to the selection of the determiner "the" prior to selection of the noun. The fact that "the" was selected will provide another influence on the selection of a subsequent noun ("child") due to learned sequencing of determiners and nouns.

[2] A reviewer raised the possibility that the different results from the eye-tracking studies may be a result of subjects knowing that their eye movements were being monitored thus altering their behavior. However, the eye-tracking studies with the critical manipulation of phrase complexity (e.g., Griffin, 2001) differed also in that subjects repeated the same syntactic structure on every trial. It is thus difficult to say with certainty which factor may have lead to different behavior, however, since the initial phrase complexity effect has been found to be reduced by syntactic priming (Smith & Wheeldon, 2001) or when subjects have foreknowledge of the structure to be used (Crowther, 2009), there is at least evidence supporting the notion that differences between the studies may be due to differences in how varied the structures were which subjects produced.

1999). The non-phonological nature of these representations has been assumed because of evidence suggesting that planning at the phonological level has a much smaller scope – namely, one phonological word (Wheeldon & Lahiri, 1997, 2002). However, a number of findings on the production of noun phrases consisting of a determiner and/or adjective and noun have provided evidence of not only lemma access but phonological encoding of all the content words in the phrase, both when grammatical features of the noun specify the phonological form of determiners and adjectives (as in Dutch, Schriefers, 1992, and Spanish, Costa & Caramazza, 2002) and when they do not (as in English, Alario, Costa, & Caramazza, 2002; Damian & Dumay, 2007). For instance, Damian and Dumay (2007) showed that (in English) distracter words phonologically related to the noun speeded onset latencies for noun phrases consisting of determiner + adjective + noun and that onset latencies for phrases consisting of an adjective and noun beginning with the same phoneme (e.g., blue bell) were faster than for those in which the two words began with different phonemes (see also Miozzo & Caramazza, 1999 and Alario & Caramazza, 2002 for similar results in Italian and French, respectively). The results imply that not only lemma but also phonological access was at least initiated for the noun prior to voice onset.

While findings for adjective–noun phrases are consistent with a phrasal scope of planning, the evidence for a phrasal scope for conjoined noun phrases has been more mixed. Data supporting the phrasal planning view were obtained by Smith and Wheeldon (1999) who adopted a moving pictures task. As shown in Fig. 1, on target trials, three objects were shown in a row (Fig. 1a) and then moved such that one object was above the other two (Fig. 1b) or two objects were above the other one (Fig. 1c). Participants described the moving displays, going from left to right across the display, using sentences such as

(1a) Simple–complex: The fork moves above the kite and the dog.
(1b) Complex–simple: The fork and the kite move above the dog.

Even though the target sentences in (1b) were quite similar to those in Griffin (2003), Smith and Wheeldon found results leading to an opposite conclusion from that of Griffin. That is, Smith and Wheeldon showed that speech onset latencies were longer for sentences beginning with a complex noun phrase (with a simple noun phrase in the predicate) such as (1b) than for matched sentences begin-

(a) Three different objects were presented in a row.

(b) Simple-complex: The fork moves above the kite and the dog.

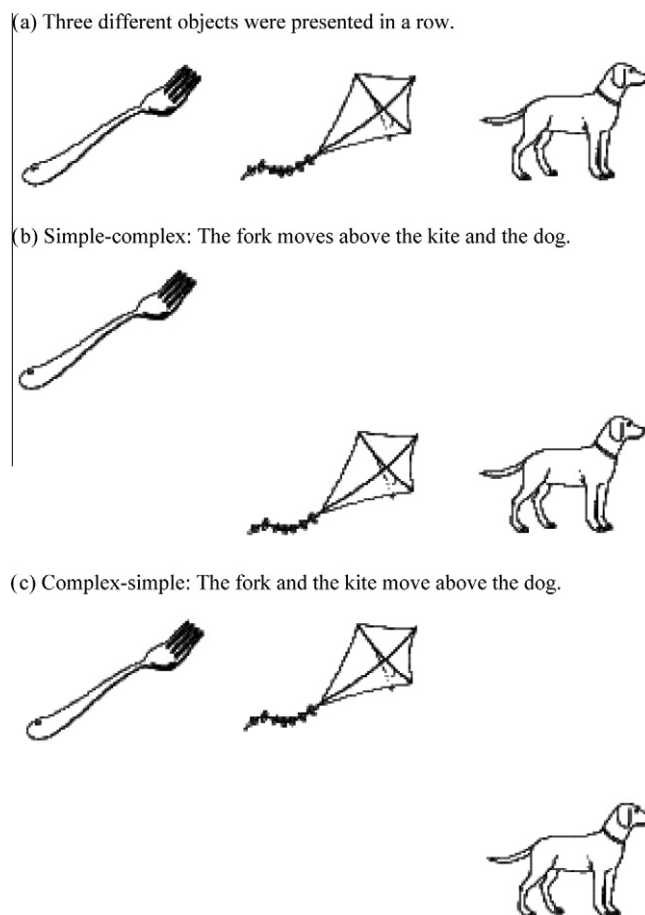(c) Complex-simple: The fork and the kite move above the dog.

Fig. 1. Examples of depictions for simple–complex and complex–simple sentences.

ning with a simple noun phrase (with a complex noun phrase in the predicate) such as (1a). Since the matched sentences began with the same content word, no differences would have been expected if participants were planning only the initial content word. The two sentences were also matched in overall length, syntactic complexity, and the identity of the other content words; therefore, no differences in onset latency would have been predicted if participants were completing lexical access at the sentence or clausal level prior to voice onset. Thus, the longer onset times for the complex–simple sentences compared to the simple–complex sentences support the contention that participants were planning lexical representations for the initial phrase prior to initiating articulation.

Findings from Martin and Freedman (2001) also support the phrasal planning hypothesis. This study followed other work that contrasts patients who show a deficit in retaining lexical-semantic information in short-term memory (STM) versus those who show a deficit in retaining phonological information (Freedman & Martin, 2001; Hanten & Martin, 2000; Martin & Romani, 1994; Martin & Saffran, 1997; Martin, Shelton, & Yaffee, 1994). If speakers plan all of the words in a phrase at a lemma level, then patients with a lexical-semantic retention deficit should have difficulty producing phrases with several content words, either because these lexical-semantic representations are part of a lemma representation, or because such representations need to be maintained for lemma access. However, if planning for production proceeds word-by-word, then such patients should perform similarly regardless of the number of content words in a phrase, given their good ability to produce and remember single words (Martin & He, 2004; Martin et al., 1994). Martin and Freedman (2001) showed that patients with a lexical-semantic STM deficit had difficulty producing phrases with one or more pre-nominal adjectives (e.g., "green leaf" or "open blue book"); however, their production improved when they produced the same information in sentence form (e.g., "the leaf is green" or "the book is open and blue"). That is, if the same content were in separate phrases, their production improved. Patients with a phonological STM deficit performed at a normal level with both types of utterances. One possible interpretation of the normal performance of patients with phonological STM deficits is that planning at the phonological level does not necessarily occur across as many words as planning at the semantic level, as discussed previously (cf., Wheeldon & Lahiri, 1997). Another possibility is that there are separate phonological STM capacities for input and output phonological representations and these patients' phonological retention deficits are on the input side, allowing for normal production (see Martin, Lesch, & Bartha, 1999, for further discussion).

Other patient findings from our lab are also consistent with a phrasal scope of planning for conjoined noun phrase production and suggest that this planning occurs at the lemma level. Martin et al. (2004) tested one patient with a semantic retention deficit (ML) and one patient with a phonological retention deficit (EA) on the moving picture materials used by Smith and Wheeldon (1999). Martin et al. reasoned that if the phrasal planning was carried out at either the lemma or phonological level, then a pa-

tient with a STM deficit at that level should have difficulty initiating the sentences beginning with a complex noun phrase. They found that patient ML showed a greatly exaggerated effect of initial noun phrase complexity (a 1027 ms effect compared to a 66 ms effect for controls). Since ML was argued to have reduced lexical-semantic STM capacity, they interpreted this as indicating that both words in the initial phrase must be planned at the lemma level before beginning articulation. Patient EA, with a phonological retention deficit, showed an effect within normal range (58 ms), again suggesting that planning has a smaller scope at the phonological level, or, as suggested previously, different buffers are involved in phonological input and output, and EA's STM deficit was on the input side (Martin et al., 1999).

Allum and Wheeldon (2007) have recently proposed a modification to the phrasal scope account. They conducted their study primarily in Japanese, a head-final language in which heads of phrases typically follow modifying phrases, which contrasts with languages such as English, in which heads typically precede modifying phrases. Taking advantage of the head-final structure of Japanese, they tested whether the scope of planning corresponded to grammatical phrases playing a role such as subject or object, or functional phrases conveying major conceptual information, such as agent, theme, or modifiers. In their study, they varied the complexity of an initial phrase modifying an agent while keeping the complexity of the subject phrase as a whole constant. For instance, they contrasted the production of sentences such as "The dog above the flower and apple is red" with "The dog and the flower above the apple are red," in which the subject phrase contains the same number of nouns, but the head and modifying phrases differ in size. However, in Japanese, the modifying phrase ("above the x") precedes the subject noun(s), so Allum and Wheeldon could investigate whether planning occurred for the modifying phrase, which was the first functional phrase, the head phrase, which was the second functional phrase, or the subject phrase as a whole, which was the first grammatical phrase and contained both the modifying and head phrases. They found that onset latencies were related only to the number of nouns in the modifying phrase, rather than the subject phrase as a whole. They thus argued that the scope of planning corresponded to functional phrases, which are phrases having a major, nonreducible conceptual function, rather than phrases corresponding to a major grammatical unit, such as subject.

1.1. Retrieval fluency hypothesis

As mentioned earlier, the sentences produced by participants in Griffin (2001) were quite similar to those produced by participants in the complex/simple condition of Smith and Wheeldon (1999), but the results of the two studies pointed to opposite conclusions regarding the phrasal vs. single word planning scope hypotheses. One means of reconciling these findings is to note that in Griffin (2001), every sentence produced had the same structure. Under these conditions, participants may have adopted a word-by-word strategy that is unlike what they would use under more naturalistic conditions. That is, because

the syntactic structure produced for the scenes was the same on every trial, and thus readily available, the experiment may have become more like a picture naming task. If this explanation is correct, then the conclusions of Smith and Wheeldon (1999) would appear more valid as they would be more applicable to language produced in more naturalistic conditions where the structure is not fixed across utterances.

Another means of reconciling these findings, however, would be to argue along the lines of Levelt and Meyer (2000) that the language production system proceeds largely in a word-by-word fashion, but advanced planning can occur under limited circumstances. Levelt and Meyer suggested several situations in which speakers may plan several lexical items. The most relevant of which for the current discussion is the use of advanced planning of lexical items in order to maximize fluency. That is, speakers may be able to begin producing an utterance more quickly when they interleave articulation and lexical access, by accessing the lexical representation of the upcoming word while producing the name of the current word, but they run the risk of speaking hesitantly if the upcoming word is difficult to retrieve. Speakers will be able to produce an utterance more fluently if several upcoming words are planned ahead of time, but they will not be able to begin speaking as soon because of the extra time devoted to planning (see also Goldman-Eisler, 1968).

In findings related to Levelt and Meyer's (2000) hypotheses, Griffin (2003) found a reversed word-length effect in utterance onset latencies for two-word productions when subjects were specifically instructed to speak as fluently as possible and try not to hesitate between words. She argued that speakers maximized fluency in this case by delaying production of the first word in the utterance until they had some sense of the availability of the phonological form of the next word. When the first word was monosyllabic, they waited until more information about the second word was retrieved to avoid pauses between the first and second word. When the first word was multisyllabic, they could begin speech earlier as they would have time to retrieve the phonological form of the second word while uttering the first (but see Meyer, Belke, Häcker, & Mortensen, 2007, for a different interpretation). In relation to the Smith and Wheeldon (1999) findings, Griffin (2003) stated:

> The present findings argue against the notion that within a particular situation speakers consistently prepare …a major constituent (see, e.g., Smith & Wheeldon, 1999) before beginning to speak. Speakers may prepare multiple content words before speaking in order to produce a fluent utterance despite short utterance-initial words rather than to obey a requirement of syntactic processing. (p. 608)

The retrieval fluency hypothesis could be used to explain Smith and Wheeldon's (1999) finding of shorter onset latencies for sentences beginning with a simple noun phrase because the second content word of the simple–complex condition was always the verb "moves," while the second content word of sentences with an initial complex noun phrase was a noun that varied from trial to trial. Because "moves" occurred very frequently during the

course of the experiment, it would be easier to retrieve than the nouns; thus, speakers could begin production of the first noun sooner in the simple–complex sentences (where "moves" is always the second content word) than in the complex–simple sentences.[3] A similar argument could be made for the results from Allum and Wheeldon (2007) in Japanese between sentences beginning with two content words in the initial modifier phrase as compared to one, as the first noun in the simple phrase was followed by the word "above", which was repeated on every trial. Such an account would, however, have difficulty explaining the similar increase Allum and Wheeldon observed in onset latencies between sentences with three nouns in the initial phrase as compared to two, since "above" did not immediately follow the first noun in either case.

Although this retrieval fluency proposal assumes some simultaneity in planning two words, according to this account, the scope of planning is not related to syntactic units such as phrases or clauses, but is related solely to the accessibility of the phonological form of the succeeding word in the utterance. It should be noted that this retrieval fluency hypothesis might also account for findings showing phonological access to the noun prior to voice onset in adjective–noun and determiner-adjective–noun phrases (e.g., Damian & Dumay, 2007). That is, many of these studies have used color adjectives in these phrases, which could be retrieved quickly because they are typically short words drawn from a small set and their retrieval would have been well practiced across the course of an experiment. Thus, because speakers could produce the color adjective so easily, they may have waited for some access to the phonology of the noun prior to initiating articulation in order to insure fluency in production of the phrase. Griffin (2003) in fact suggested that a number of classic findings in the speech production literature which have been attributed to advance planning of various grammatical units (e.g., see Kempen & Huijbers, 1983; Lindsley, 1976) might be reinterpreted in terms of speakers' striving to achieve fluency.

Some patient findings noted previously might also be reconsidered in terms of retrieval fluency. For instance, patient ML's much larger latency difference between the complex–simple and simple–complex sentences on the moving-picture task (Martin et al., 2004) might be ex-

---

[3] A reviewer noted that the application of Griffin's (2003) retrieval fluency account to Smith and Wheeldon's (1999) findings would have predicted that Griffin (2001) should have observed an effect of encodability and word frequency of the second noun in the complex-simple sentences for sentence onset latencies, since these factors should influence the time needed to access some phonological information about the second noun. However, such effects were not obtained in Griffin's (2001) study. An account might be offered on the grounds that the same sentence structure was used on every trial in Griffin's (2001) experiment, which may have enabled subjects to proceed in a more strictly word-by-word fashion than was possible in the Smith and Wheeldon study. That is, since subjects knew that "and the" would always follow the first noun, little in the way of cognitive resources would have been needed to plan and articulate these function words, allowing these resources to be used for retrieving the second noun. Griffin (2003) obtained results that fit such an explanation, as the reverse word-length effect in that study was eliminated when subjects produced "next to" between two nouns. With these intervening words repeated on every trial, subjects would then have time enough time and resources to plan the second noun.

plained if one assumed that his lexical-semantic STM deficit made it very difficult for him to maintain one lexical-semantic representation while attempting the retrieval of the next. This difficulty was lessened when the content word was one that had been repeatedly retrieved (i.e., "moves") and thus was in an easily retrievable state.

Both the phrasal scope of planning and retrieval fluency approaches offer an interpretation of the onset latency effect for producing complex vs. simple noun phrases in sentence-initial positions that was reported in Smith and Wheeldon (1999) and Martin et al. (2004). The first two experiments reported in the current study were designed to distinguish between these two accounts. In Experiment 1, we contrasted onset latency differences in the production of sentences like those in Smith and Wheeldon (1999; see 1a and 1b above) with sentences in which participants produced the adjective "yellow" before the second noun in every sentence, as in (2a) and (2b) below. If retrieval fluency is the relevant factor, this manipulation should serve to make the second content word of the complex–simple sentences easier to produce and reduce or eliminate the onset difference between the simple–complex and complex–simple sentences by equating the retrieval fluency of "moves" and "yellow", respectively.

(2a) Complex–simple: The fork and the yellow kite move above the dog.
(2b) Simple–complex: The fork moves above the yellow kite and the dog.

In Experiment 2, we took the opposite approach. Instead of decreasing the retrieval difficulty of the second content word in the complex–simple sentences, we increased the retrieval difficulty of the second content word in the simple–complex sentence (i.e., the verb) by varying the verb across trials. We matched the number of possible verbs in the simple–complex sentences with the number of possible second nouns in the complex–simple sentences. Retrieval times were found to be longer for the verbs than the nouns when produced alone. Thus, if retrieval fluency for the second content word is the issue, these manipulations should eliminate the onset latency difference between the complex–simple and simple–complex sentences.

1.2. Visual grouping hypothesis

In Smith and Wheeldon's (1999) moving-picture displays, two pictures moved together in one direction while the third moved in the opposite direction. For the complex–simple sentences, the initial object to be described moved together with the second one whereas in the simple–complex sentences, the initial object moved by itself. The "common fate" movement of the first two objects in the complex–simple sentences could have caused some difficulty in focusing on the retrieval of the name of one of the two objects. Allum and Wheeldon (2007) eliminated this potential confound caused by the movement of the pictures by presenting subjects with stationary displays, and still found an effect of initial phrase complexity. However, they instead used color to indicate to subjects which

nouns should be grouped together in the initial phrase, which may have again made it more difficult to distinguish the two objects based on perceptual grouping.

Some recent findings lend plausibility to the visual grouping hypothesis. For instance, Morsella and Miozzo (2002) and Navarrete and Costa (2005) found evidence of access to the phonological representation of a distractor object during target object naming when the target and distractor object were superimposed (see Oppermann, Jescheniak, & Schriefers, 2008, for evidence regarding the limitations of visual grouping effects). These findings raise the possibility that the initial phrase complexity effect results from processes occurring during the visual encoding of the scene. Since previous studies manipulating initial phrase complexity have cued subjects to which structure they are to produce by some kind of visual grouping (by movement, proximity, color, alignment, or a combination of these), it is possible that effects attributed to phrasal planning occur at earlier processing level and instead affect the probability of retrieving an incorrect name at the wrong time. Because of the visual association made between the two items in a complex phrase, subjects may be slowed down when producing a conjoined noun phrase because both names are retrieved and interfere with one another, rather than because subjects require extra time to plan both items that are in the same phrase.

Experiments 3 and 4 in the current study tested the possibility that perceptual grouping may give rise to the initial phrase complexity effect. Experiment 3 was a replication of Experiment 1 using a stationary picture display, which would eliminate the strong grouping effect due to movement. Experiment 4 examined whether grouping due to visual alignment might play a role even for stationary displays. This experiment contrasted production of the simple–complex and complex–simple sentences for the stationary display with the production of lists of object names for the same displays. If the complexity effect resulted from visual grouping of pictures then the same pattern of effects should be observed for sentences and lists. Experiment 3 and 4 also differed from Experiments 1 and 2 in that the pictures remained on the screen throughout the participants' utterances. In Experiments 1 and 2, in order to encourage encoding of the conceptual representation of the entire picture prior to utterance onset, the pictures were removed when articulation began. However, such a procedure may have forced participants to plan more of their utterances than they would under normal circumstances.

2. Experiment 1

Experiment 1a was a replication of Smith and Wheeldon's (1999) Experiment 1. Experiment 1b was identical to Experiment 1a, except that the middle object within each three-picture display was colored yellow and participants were asked to describe both the movement of the objects in the animated scene and the color of the middle object. According to the phrasal scope of planning hypothesis, the addition of the adjective should, if anything, increase planning time for the initial phrase of the

complex–simple sentences because an additional content word was added to the initial phrase. According to the retrieval fluency hypothesis, the addition of the highly redundant word "yellow" as the second content word of the complex–simple sentences should make the complex–simple sentences easier to initiate, thereby reducing or eliminating the onset latency difference.

2.1. Method

2.1.1. Participants

Twelve Rice University students were tested in Experiment 1a and twelve in Experiment 1b. They received credit toward course requirements for experiment participation.

2.1.2. Materials

Both Experiment 1a and 1b used a set of black and white line-drawing pictures from Snodgrass and Vanderwart (1980). These pictures were all familiar objects that had one, two, or three syllable names. Ninety pictures were used for experimental and filler trials. Pictures used for experimental trials were combined to make 24 sets of three-picture displays, and those used for filler trials were combined to make 16 sets of three-picture displays. Care was taken to ensure that the three objects within each display were conceptually different (e.g., fork, kite, dog). For each experimental trial, two objects moved together in one direction (up or down) and the other object moved alone in the opposite direction (as in Fig. 1). Four experimental versions of each display were generated by combining initial noun phrase type (simple–complex vs. complex–simple) with the direction of relative movement of the objects from left to right (either up–down or down–up). The crucial difference between Experiment 1a and 1b was that the middle object in each display frame was colored yellow in Experiment 1b.

For the filler trials, the three objects moved in the same direction (up, down, right or left). Participants were asked to produce sentences like "They all moved up/down/right/left." The middle objects in the filler displays were also colored yellow in Experiment 1b. Filler trials were used to increase the variation of the syntactic constructions of the sentences produced during the experiment. There were 96 experimental trials and 128 filler trials. Trials were divided into four blocks of 56 trials and the experiment was completed in one 50-min session. Only one of the four versions of each picture display appeared within each trial block. All four versions of each picture display appeared across the testing session.

2.1.3. Design and procedure

Both Experiment 1a and 1b adopted the same design and procedures. The experiments were run on PsyScope 1.2.5 program for Macintosh (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were tested individually in a quiet and comfortable environment. Before running the experiment, a test was administered for each participant to ensure the appropriate sensitivity setting for the voice key. Participants first watched a set of 10 moving pictures movies and discussed the correct description of each movie with the experimenter. Participants then completed a set of 12 practice trials to ensure they understood the types of sentence constructions to use in the production task. In both the practice and experimental trials, participants saw a rectangular fixation point in the center of the screen for 500 ms followed by the moving-picture displays. They were asked to describe the upward or downward movement of the objects starting from the left side of the display. Each moving-picture display was removed 500 ms after participants began their utterances in order to encourage participants to conceptually encode the scene before beginning to speak (Smith & Wheeldon, 1999). Voice onset latency for the sentence was the dependent measure and was calculated from the beginning of the presentation of the object display to the beginning of participants' speech. The experimenter coded each trial for accuracy of the participants' responses and technical errors.

2.2. Results

2.2.1. Experiment 1a

Response latencies shorter than 300 ms and longer than 3s were considered outliers and were excluded from data analysis. This resulted in the loss of 0.9% of the data. Unexpected noises that triggered the voice key before participants began their response to the pictures were considered technical errors. There were two types of response errors: (1) participants misnamed an object in the display, and (2) participants did not produce the expected syntactic structure for the sentence. Trials with technical errors and response errors constituted of 0.4% and 1.7% of the data, respectively. These data points were also excluded from analyses.

Table 1 shows the onset latencies to the moving-picture displays as a function of sentence type. The mean onset latency for the complex–simple sentences was 76 ms longer than that for the simple–complex sentences. This 76 ms difference was significant both by participants and by items ($t_1(11) = 4.67$, $SE = 16.27$, $p < 0.001$; $t_2(45) = 4.71$, $SE = 16.13$, $p < 0.001$). The percent error rates were 4.0% in the complex–simple condition and 3.0% in the simple–complex condition, a difference that was far from significance, ($t_1(11) = 0.861$, $SE = 1.16$, $p = 0.40$). This pattern replicates the findings of Smith and Wheeldon (1999) and the magnitude of the reaction time effect was nearly identical to that reported in their Experiment 1 (77 ms).

2.2.2. Experiment 1b

Outliers, technical errors, and response errors of both experimental and filler trials resulted in the loss of 2.6%, 0.6% and 2.9% of the data, respectively. As shown in Table 1,

Table 1
Mean onset latencies (ms) as a function of sentence type in Experiments 1a and 1b.

Sentence type	Experiment 1a (uncolored version)	Experiment 1b (yellow middle object)
Simple–complex	1056	997
Complex–simple	1132	1069
Difference	76	72

participants took 72 ms longer to describe moving-picture displays with a complex–simple format than with a simple–complex format, a difference that was significant both by participants and by item ($t_1(11) = 3.49$, $SE = 20.63$, $p < 0.005$; $t_2(45) = 2.68$, $SE = 26.87$, $p < 0.01$). The error rates were 7.7% in the complex–simple condition and 5.0% in the simple–complex condition, a difference that failed to reach significance ($t_1(11) = 1.39$, $SE = 1.94$, $p = 0.191$).

2.2.3. Across-experiment analysis

In order to compare the effects from Experiment 1a and 1b, the reaction time data were analyzed using experiment as a between-participants variable. The analysis of variance (ANOVA) showed a main effect of sentence type (complex–simple vs. simple–complex) ($F(1, 22) = 31.56$, $MSE = 2374$, $p < 0.001$). There was no main effect of experiment ($F(1, 22) = 0.447$, $MSE = 99,892$, $p = 0.511$) nor any interaction between sentence type and experiment ($F(1, 22) = 0.018$, $MSE = 2374$, $p = 0.896$).

2.3. Discussion

The results from Experiment 1 did not support the retrieval fluency hypothesis regarding the source of the onset latency difference between the complex–simple and simple–complex sentences types, which predicted the complexity effect should disappear when retrievability of the second content word was equalized across the two sentence types. Instead, the results are consistent with the phrasal scope of planning hypothesis because the longer onset latencies for complex–simple relative to simple–complex sentences persisted despite the incorporation of the redundant second content word "yellow." Thus, these findings in English add to the findings in Japanese reported by Allum and Wheeldon (2007) in showing that initial phrase complexity, in terms of number of content words, rather than retrieval fluency, is the source of longer onset latencies for more complex phrases.

Interestingly, the addition of the adjective in Experiment 1b did not increase the onset latency difference, which might have been expected if participants took longer to retrieve three content words in the initial phrase than to retrieve two, as was required in Experiment 1a. This finding can be explained by assuming that the nouns and adjective are retrieved in parallel. If so, then time to onset would be dependent on the time to access the most slowly retrieved word (Kempen & Huijbers, 1983; Schriefers, de Ruiter, & Steigerwald, 1999). Given the redundancy of the word "yellow," it is likely that this word is more quickly retrieved than the nouns. Thus, onset latencies would depend on the retrieval times for the two nouns – which were equivalent across Experiments 1a and 1b.

3. Experiment 2

Another means of equating the retrieval difficulty of the second content word in the simple–complex and complex–simple sentences is to increase the retrieval difficulty of the verb in the simple–complex sentences. In Experiment 2, we again used the moving-picture task, but modified it such that five different actions were depicted (i.e., bump, follow, jump over, lead, move). Also, the three objects in each display were chosen from a set of six objects. Thus, the predictability of the second content word in the complex–simple sentences (1 out of 5 remaining objects) was equivalent to that of the second content word in the simple–complex sentences (1 out of 5 actions).

If ease of retrieval of the second content word is the determining factor in the onset latency difference between the simple–complex and complex–simple sentence types, then this manipulation should serve to eliminate the onset latency difference between the two sentence types. To verify that the verbs were at least as difficult to retrieve as the nouns, a test was carried out with a subset of the participants that required them to name only the verb in the display or to name only the middle object in the display. As indicated below in the methods section, naming times for the verbs were significantly longer than those for the nouns. Thus, according to the retrieval fluency hypothesis, one would predict that the onset latency pattern would reverse with longer times for the simple–complex than complex–simple sentences, given the greater difficulty of accessing the verbs than the nouns. In contrast, the phrasal scope of planning hypothesis predicts that the disadvantage for the complex–simple sentences should persist.

3.1. Method

3.1.1. Participants

Twenty-two Rice University students participated in this experiment. They received credits towards course requirements.

3.1.2. Materials

Six black and white line-drawing pictures (brush, fence, gate, jacket, pencil, and football) selected from the Snodgrass and Vanderwart (1980) and Philadelphia Naming Task (Roach, Schwartz, Martin, Grewal, & Brecher, 1996) materials were used in this task. These pictures were objects with one- or two-syllable names that were closely matched in lexical frequency (from the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995): lexical frequency averaged 37.3, and ranged from 34 to 46). These six pictures were combined to make 16 sets of three-picture displays, within which each picture appeared in each screen position two or three times. Each three-picture display was used to derive two types of experimental trials by varying the relative movements of objects in the picture to correspond to the complex–simple and simple–complex sentence formats. These 32 displays were then combined with the five kinds of animation (bump, jump over, lead, follow, and move) to make a total of 160 experimental trials for the task. The nature of the animation is described below. Note that to simplify these descriptions, only those animated scenes that instantiated the simple–complex sentence format are described.

(1) Bumps: The left-most object moves to the right while the other two objects on the right remain stationary. The right two objects jiggle slightly after the left-most object touches the middle object.

(2) Jumps over: The left-most object moves in an arc over the other two objects on the right which remain stationary.

(3) Leads: The object on the left moves in a diagonal direction to the upper or lower left and the other two objects follow at 300 ms delay.

(4) Follows: The rightmost two objects move in a diagonal direction to the upper right or lower right and the left-most object follows at 300 ms delay.

(5) Moves: The object on the left moves in one direction (up or down) while the objects on the right move in the opposite direction.

Filler trials used the same set of 16 three-picture displays as the experimental trials. The displayed objects in filler trials all moved either in the same direction (up, down, right, or left) or in different directions (for example, the left-most object moved up, the middle object moved right, and the right-most object moved down, etc.). Participants were asked to produce, "They all moved up/down/right/left" or, "They all moved apart", as appropriate. There were two sessions for each participant with 5–7 days between sessions. In each session, a total of 160 experimental trials together with 160 filler trials were divided into two blocks with 80 experimental and 80 filler trials each. Only one version for each three-object display appeared within one block and all versions of each three-picture display appeared across the complete session. In the second session, the same experimental and filler trials were used but in a different order of blocks.

3.1.3. Design and procedure

The design and procedures were similar to those of Experiment 1a except that additional practice was introduced to familiarize the participants with using the appropriate verbs to describe the different kinds of animated motion. Participants were first shown examples of the five types of movement and told what verbs to use. A practice test was then administered to determine their accuracy in producing the verbs. Object pictures other than those used in the experimental trials were selected to make 22 combinations of three-picture displays. The animated scenes using these objects were constructed by varying the different kinds of movements (bump, jump over, lead, follow, and move) and the complexity of sentence type (simple–complex and complex–simple). The participants were told to describe only the type of movement based on the animations they saw on the screen. This practice test was repeated if participants made any wrong responses. Participants were allowed to progress to the experimental trials only when they got all practice trials correct.

Eighty displays were used for the noun and verb naming task, using each of the 16 sets of objects and five types of animated actions (bump, jump, lead, follow, and move) with half depicting the simple–complex and half the complex–simple sentence types. These 80 displays were used in both the verb-naming and object-naming task. In the verb-naming task, participants were asked to describe only the animation in the scene, whereas they were to name only the middle object during the object-naming task. Twenty verb-naming trials and 20 object-naming trials were administered at the end of each experimental block. In order to verify that varying the verb had served to equate retrieval difficulty for the nouns and verbs across sentence types, a sub-test was also conducted with 12 of the 22 participants in which they were asked to name just the middle noun or the verb. This sub-test was conducted at the end of each experimental block during the experimental sessions so that these latencies could be compared at a point when the participants were quite familiar with the nouns and verbs being used.

3.2. Results

In the sub-test comparing the naming of single nouns and verbs, 1.8% of the data were considered outliers, and 2.8% of the data were response errors. The reaction times for these trials were excluded from the analysis. The mean onset latency for naming the verbs (975 ms) was significantly longer than that for the nouns (696 ms) ($t_1(11) = 9.90$, $SE = 28.18$, $p < 0.001$, $t_2(15) = 17.90$, $SE = 15.59$, $p < 0.001$). The error rates were 3.4% for verb naming and 2.2% for object naming, a difference that failed to reach significance ($t_1(11) = 1.71$, $SE = 0.70$, $p = 0.115$; $t_2(15) = 0.95$, $SE = 1.26$, $p = 0.357$).

The same criteria for outliers and coding errors as in Experiment 1 were adopted for this experiment. For the main experiment on sentence type, the outliers excluded from data analysis comprised 3.2% of the data across both experimental and filler trials. In addition, response errors resulted in the loss of 4.9% of the data and were also excluded from reaction time analysis. Table 2 shows the onset latencies to moving pictures as a function of sentence type. An ANOVA was conducted using onset latencies as dependent measures and the following factors: the complexity of sentence types (simple–complex vs. complex–simple) and verb (bump, jump, lead, follow, and move).

The results indicate that the main effect of sentence type reached significance both by participant ($F_1(1, 21) = 31.58$, $MSE = 42.81$, $p < 0.001$) and by items ($F_2(1, 15) = 33.76$, $MSE = 40.05$, $p < 0.001$). It took 52 ms longer for the participants to produce complex–simple sentences than to produce simple–complex sentences. The main effect of type of verb was also significant by participants and by items ($F_1(4, 84) = 23.07$, $p < 0.001$; $F_2(4, 60) = 65.17$, $p < 0.001$). Fig. 2 shows the voice onset latencies as a function of types of verbs. Post hoc contrasts further indicated that the voice onset latencies of sentences with the verb "follow" were significantly longer than the pooled voice onset latencies of sentences with the other four types of verbs (bump, jump, lead and move) both by participants ($t_1(21) = 10.02$, $p < 0.001$) and by items ($t_2(15) = 14.57$, $p < 0.001$). No other comparisons among the verbs were significant. Crucially, the interaction

Table 2
Mean onset latencies (ms) as a function of sentence type in Experiment 2.

Simple–complex	1244
Complex–simple	1296
Difference	52
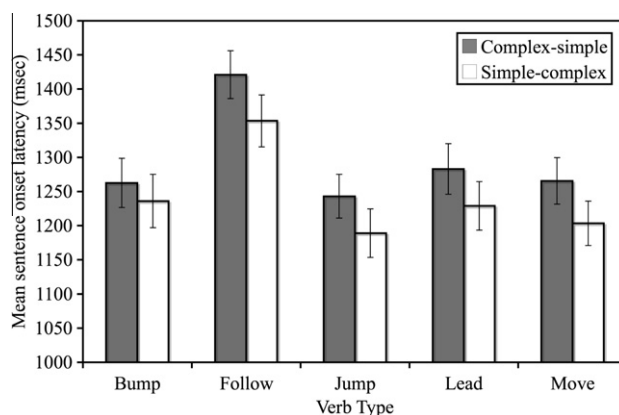
Fig. 2. Voice onset latencies (ms) for difference sentence types (simple–complex, complex–simple) in Experiment 2 as a function of the verb type in the moving pictures task. Error bars show the 95% confidence interval around the mean.

effect between these two factors, complexity of sentence type and the verb type, was not significant (F_1 (4, 84) = 0.75, p = 0.56; F_2 (4, 60) = 0.78, p = 0.54), indicating that the phrase complexity effect was of similar magnitude irrespective of differences in the ease of identifying the different types of movement. There was no significant effect on accuracy as a function of the complexity of sentence type by participants or by items (t_1(21) = 0.045, p = 0.964; t_2(15) = 0.051, p = 0.960), as the percent error rate of experimental trials was 4.9% for both the simple–complex and complex–simple conditions.

Although the onset latency difference was substantial (52 ms), it was somewhat reduced relative to that in Experiment 1a (76 ms) and 1b (72 ms), even though onset latencies for both conditions of Experiment 2 were increased compared to Experiments 1a and b. However, no interaction was found between experiment and phrase complexity across Experiments 1a and 2 (F_1 (1, 33) = 1.93, MSE = 350,132, p = 0.174).

3.3. Discussion

Despite the participants' longer times to retrieve verbs describing the animated displays than to retrieve nouns, the disadvantage in onset latencies for the complex–simple sentences relative to the simple–complex sentences persisted, consistent with the phrasal planning hypothesis. Thus, the second experiment also provides support for the phrasal scope of planning hypothesis and adds further evidence against the retrieval fluency account of the effect of initial noun phrase complexity.

The variation in sentence onset latencies for the different verbs is also of some interest. A case can be made that these differences derive mainly, if not entirely, from the relative difficulty in identifying the action and in switching focus from the verb to the nouns rather than to variation in the retrieval of lexical representations for the verbs. The main source of evidence for this claim is the long onset latency for sentences with the verb "follow" compared to those with other verbs. In the case of "follow", the objects on the right start to move first, which would capture attention, but the utterance has to begin with the objects on the

left. For the other verbs, the objects on the left begin movement, either alone (as in bumps, leads, and jumps over) or simultaneously with movement of the objects on the right (as for moves). Thus, attention is more likely to already be on the objects which have to be named first. In contrast to the evident role of these conceptual/attentional factors, word frequency, a lexical variable, appeared to have no relation to these onset latencies. For instance, "bump" has the lowest frequency (11, CELEX database, Baayen, Piepenbrock, & Gulikers, 1995) and "move" has the highest (435) but onset latencies were similar, whereas "follow" and "lead" have the same frequency (308), but "follow" had by far the longest onset latency. If our reasoning about the role of conceptual/attentional factors is correct, then the independence of the verb effect and the initial phrase complexity effect results because they reflect different stages of processing – conceptual encoding and attention switching for the verb effect and lexical retrieval for the phrase complexity effect. As such, the findings would argue against theories that hypothesize that verbs are retrieved first in sentence production (e.g., Ferreira, 2000). We should acknowledge, however, that the experiment was not designed to assess the role of conceptual or lexical factors related to the verb on onset latencies. The large variation in latencies due to conceptual/attentional factors may have masked any contribution from lexical factors. Moreover, if frequency reflects phonological access, as some have claimed (Jescheniak & Levelt, 1994), then the lack of a frequency effect would only rule out phonological access, but not lemma access. Clearly, further work would be needed to ascertain whether the action was processed only at a conceptual level or at both conceptual and lexical levels prior to voice onset.

4. Experiment 3

One possible criticism of the design of Experiments 1 and 2 is that the "common fate" movement of the pictures in the complex noun phrase may have resulted in a visual grouping effect which influenced participants' utterance planning. Exactly why such grouping would slow onset

latencies for the complex–simple sentences relative to the complex–simple sentences is not entirely clear. One possibility, though, is that it may be difficult to ignore the second object of the grouped pair, resulting in interference from the second object that slows lexical retrieval for the first object to be named relative to the case where the first object is isolated by itself. As mentioned earlier, grouping by color may have influenced the findings in Japanese reported by Allum and Wheeldon (2007).

Another possible criticism of the design of Experiments 1 and 2 relates to the fact that the displays disappeared at utterance onset. This was done in order to encourage participants to encode the message level representation of the display prior to initiating speech. In naturalistic production, one might assume that speakers have a complete idea in mind before beginning to speak. However, an opposite argument might be made. That is, the disappearance of the display may have encouraged participants to lexically encode more than they would under more naturalistic situations in order to aid in their memory for the display once it disappeared.

In order to address these concerns, Experiments 3a and b were carried out to replicate Experiments 1a and b, but using stationary displays that remained in view throughout the utterance. The sentences were the same as in Experiment 1, except that participants would say "is above/below" or "are above/below" instead of "moves above/below" or "move above/below."

4.1. Method

4.1.1. Participants

Twelve Rice University students were tested in Experiment 3a and twelve in Experiment 3b. They received credit toward course requirements for experiment participation.

4.1.2. Materials

Both Experiment 3a and 3b used a set of black and white line-drawing pictures from the International Picture Naming Project (Szekely et al., 2004) database. All pictures were familiar objects with one- or two-syllable names. 144 pictures were used for experimental and filler trials. The 144 pictures were divided into 48 experimental sets of three pictures and combined into another 48 filler sets of three pictures. For each experimental trial, two objects would appear next to each other with another picture diagonally displaced from the picture closest to it. The distance from the center of the middle picture to the center of the pictures of either side was the same whether the picture was on the same level vertically or not. As in Experiment 1, the difference between Experiment 3a and 3b was that the middle picture was colored yellow in 3b.

Four experimental versions of each display were generated by combining initial phrase type (simple–complex vs. complex–simple) with the direction of displacement of the objects from left to right (above-below or below-above). For the filler trials, all three pictures appeared in a row at the top, bottom, right, or left of the screen. Participants were asked to produce sentences like "The fork, kite, and dog are all at the top/bottom/left/right." The middle objects in the filler displays were colored in yellow in Exper-

iment 3b. There were 96 experimental trials and 48 filler trials. Participants completed the experiment in one 40-min session. There were two versions of both Experiment 3a and 3b. In one version, half the triplets were presented in both a simple/above-complex/below and complex/below-simple/above format and the other half were presented in both a simple/below-complex/above and complex/above-simple/below format. In the other version, the displacement of initial and final noun phrases were reversed. In total, participants were presented with both a simple–complex and complex–simple version of each experimental picture triplet once.

4.1.3. Design and procedure

Both Experiment 3a and 3b adopted the design and procedure of Experiment 1 except for the following differences. Participants were presented two examples of both the experimental and filler sentences with accompanying sentences. Participants then performed 24 practice trials (3 of each experimental and filler version) and feedback was given by the experimenter if an error was made. Participants were then presented with each experimental picture and its name individually and asked to name them. Participants were instructed to use the name presented when naming the picture later if possible, but using another acceptable name would not be counted as an error. During this naming portion, the voice key was adjusted for each participant.

In both practice and experimental trials, participants saw a "+++" display presented for 500 ms followed by the picture display. They were instructed to describe the configuration of the items starting from the left side of the display, or the top of the display in the case of the filler sentences in which items were arranged vertically at the left or right of the screen. The display remained on the screen throughout the participant's utterance. The experimenter coded each of the participants' productions for accuracy.

4.2. Results

4.2.1. Experiment 3a

Response latencies shorter than 300 ms and longer than 2500 ms were considered technical errors and outliers, respectively, and were excluded from the analysis. Response errors were coded the same as in the earlier experiments. Outliers, technical errors, and response errors constituted 0.2%, 0.4%, and 5.0% of the data, respectively. Table 3 shows the onset latencies to the picture display as a function of sentence type. The mean onset latency for the complex–simple sentences was 33 ms longer than

Table 3
Mean onset latencies (ms) as a function of sentence type in Experiments 3a and 3b.

Sentence type	Experiment (uncolored version)	Experiment 3b (yellow middle object)
Simple–complex	1076	1068
Complex–simple	1109	1115
Difference	33	47

for the simple–complex sentences, a difference which was significant both by participants and by items ($t_1(11) = 2.29$, $SE = 14.06$, $p = 0.042$; $t_2(47) = 2.53$, $SE = 13.67$, $p = 0.015$). Response error rates were 6.3% for the complex–simple and 3.7% for the simple–complex, which was marginally significant by participants ($t_1(11) = 2.01$, $SE = 1.29$, $p = 0.06$) and significant by items ($t_2(47) = 2.18$, $SE = 1.19$, $p = 0.03$).

4.2.2. Experiment 3b

Outliers, technical errors, and response errors constituted 0.26%, 1.4%, and 7.7% of the data, respectively. As shown in Table 3, participants' onset latencies were 47 ms longer in the complex–simple structure than in the simple–complex structure, which was significant by both participants and items ($t_1(11) = 2.91$, $SE = 16.94$, $p = 0.013$; $t_2(47) = 3.28$, $SE = 12.91$, $p = 0.002$). Response error rates were 9.6% for the complex–simple and 5.9% for the simple–complex, with the difference not significant by participants ($t_1(11) = 1.97$, $SE = 1.88$, $p = 0.07$), but significant by items ($t_2(47) = 2.02$, $SE = 1.83$, $p < 0.05$).

4.2.3. Across-experiment analysis

An ANOVA comparing sentence type and experimental version was conducted. There was a main effect of sentence type (complex–simple vs. simple–complex) ($F(1, 22) = 12.48$, $MSE = 1454$, $p = 0.002$). There was no main effect of experiment ($F(1, 22) < 0.01$, $MSE = 65,828$, $p = 0.993$) nor any interaction between sentence type and experiment ($F(1, 22) = 0.37$, $MSE = 1454$, $p = 0.549$).

4.3. Discussion

Experiment 3 replicated the results of Experiment 1 using stationary picture displays that remained on the screen throughout the participants' utterances. These results indicate that the complexity effect on utterance onsets is not due to visual grouping due to common fate movement nor to the participants being forced to plan further ahead in their utterances than they normally would. The failure to find a significant reduction in the complexity effect in Experiment 3b relative to 3a provides further evidence that phrasal scope rather than retrieval fluency is the source of the initial phrase complexity effect.

However, it should also be noted that there was a substantial reduction in the complexity effect in Experiment 3a (33 ms) compared to Experiment 1a (76 ms) and also to what was reported previously by Smith and Wheeldon (1999) (77–92 ms). One possibility is that some of the complexity effect results from visual grouping due to movement of the pictures, with the portion of that effect being removed by having the pictures remain stationary. Another difference between Experiment 3a and the previous experiments was that the fillers were of the form "The fork, kite, and dog are all at the top," rather than "They all moved up," as in the previous experiments. It is possible that there was also some priming of the complex–simple syntactic structure when preceded by a filler trial with several nouns in the initial phrase. A similar reduction of the complexity effect was found by Allum and Wheeldon (2007, see Experiment 1) when their filler

items were more structurally similar to sentences beginning with a complex noun phrase.

5. Experiment 4

While Experiment 3 ruled out the possibility that the initial phrase complexity effect was due entirely to grouping through movement, it remains possible that some visual aspect of the stationary displays contributed to the effect. A serious issue relating to this line of research is what effect properties of the visual displays themselves may ultimately have on naming latencies. In all studies contrasting the production of complex versus simple phrases, some form of grouping must be used to indicate to the subjects which objects should be grouped together in a phrase and which should appear in different phrases. Effects of initial phrase complexity have been found using a range of visual cues, including grouping by movement, proximity, horizontal/vertical alignment, and color. Despite finding similar effects across different visual cues, there is the possibility that initial phrase complexity effects result from properties of the visual displays themselves, rather than from linguistic processes. Experiment 4 was intended to address this issue. In Experiment 4a, participants produced sentences with the same syntactic structures as in Experiment 3a. In Experiment 4b, participants simply named the pictures individually from left to right. If the visual configuration of the pictures does make a contribution to the complexity effect, then such an effect should also be found when participants simply named the pictures. In contrast, if the effect depends on phrasal planning, then the effect should only be obtained in Experiment 4a. Experiment 4 thus provides a strong test of whether phrasal planning is the source of the onset latency differences between the complex–simple and simple–complex sentences describing these displays.

5.1. Method

5.1.1. Participants

Fourteen Rice University students participated in Experiment 4a and twelve in Experiment 4b. Two participants were removed from the analysis in Experiment 4a due to excessive errors (over 15% of responses). Participants received credit toward course requirements for experiment participation.

5.1.2. Materials

Both Experiment 4a and 4b used a set of black and white line-drawing pictures from the International Picture Naming Project (Szekely et al., 2004) database. All pictures were familiar objects with one- or two-syllable names. A set of 192 pictures was used for experimental trials and another set of 48 pictures was used for filler trials. The 192 experimental pictures were divided into 64 experimental sets of three pictures and the 48 filler pictures were combined into 32 sets of three pictures. Participants were presented with the same displays as in Experiment 3, and produced the same syntactic structures in Experiment 4a as in Experiment 3a. Participants in Experiment

4b simply named the pictures individually from left to right.

There were 128 experimental trials and 64 filler trials. Trials were divided into two blocks of 96 trials and the experiment was completed in one 45-min session. Only one of the four versions of each picture display occurred in each block. As in Experiment 3, each participant saw only two versions, a simple–complex and complex–simple version with contrasting placement of the initial and final noun phrases. There were four versions of both Experiment 4a and 4b. The differences between the first and second versions were the same as in Experiment 3. The third and fourth version corresponded to the first and second, except that the first and second pictures were flipped in position. Whether a picture triplet was seen in a complex–simple or simple–complex version in the first or second block was counterbalanced.

5.1.3. Design and procedure

Experiment 4 used the design and procedure of Experiment 3 with the following exceptions. Participants first named each of the 240 pictures individually. They were instructed to use the name presented when naming the picture later if possible, but using another acceptable name would not be counted as an error. Participants were then presented with examples of each display and a practice session. Participants in 4a were instructed to use the same syntactic structures as in Experiment 3, including the format of the filler trials, while participants in 4b were instructed to name the items from left to right (or top to bottom, in the case of the vertically aligned filler trials).

At the beginning of each trial, an electronic beep sounded at the onset of the display. Response latencies were measured from the onset of the beep to the onset of the picture name (i.e., excluding "the") on a recorded sound file of the session. Once the participants finished an utterance, they would press a key to proceed to the next trial. Once they pressed the key, the display would disappear and be followed by a 2.5 s blank screen before the onset of the next display.

5.2. Results

5.2.1. Experiment 4a

No technical errors occurred, since onset latencies were extracted from the sound file. The criteria for excluding outliers and coding response errors were the same as in Experiment 3. Outliers and response errors constituted 3.6% and 7.2% of the data, respectively. As shown in Table 4, participants' onset latencies were 63 ms longer in the com-

Table 4
Mean onset latencies (ms) as a function of sentence type in Experiment 4a and 4b.

Production type	Experiment 4a (sentence)	Experiment 4b (list)
Simple–complex	1243	1251
complex–simple	1306	1255
Difference	63	4

plex–simple structure than in the simple–complex structure, which was significant by both participants and by items ($t_1(11) = 4.08$, $SE = 15.37$, $p < 0.001$; $t_2(47) = 4.68$, $SE = 14.70$, $p < 0.001$). Response error rates were 8.7% for the complex–simple and 5.7% for the simple–complex, which was significant both by subjects ($t_1(11) = 2.54$, $SE = 1.18$, $p = 0.03$) and by items ($t_2(127) = 2.55$, $SE = 1.18$, $p = 0.01$).

5.2.2. Experiment 4b

Outliers and response errors constituted 1.8% and 4.5% of the data, respectively. As shown in Table 4, participants' onset latencies were 4 ms longer in the complex–simple structure than in the simple–complex structure, which was significant neither by participants nor by items ($t_1(11) = 0.09$, $SE = 17.46$, $p = 0.933$; $t_2(47) = 0.27$, $SE = 13.52$, $p = 0.787$). Response error rates were 5.2% for the complex–simple and 3.8% for the simple–complex, which was not a significant difference by subjects ($t_1(11) = 1.14$, $SE = 1.23$, $p = 0.278$) but was marginally significant by items ($t_2(127) = 1.78$, $SE = 0.79$, $p = 0.07$).

5.2.3. Across-experiment analysis

An ANOVA comparing sentence type and experimental version was conducted. There was a main effect of sentence type (complex–simple vs. simple–complex) ($F(1, 22) = 7.63$, $MSE = 1623$, $p = 0.011$) and a significant interaction of sentence type and experiment $F(1, 22) = 6.94$, $MSE = 1623$, $p = 0.015$. There was no main effect of experiment ($F(1, 22) = 0.11$, $MSE = 42,334$, $p = 0.741$).

5.3. Discussion

When participants produced a sentence to describe a picture display, as in Experiment 4a, there was an effect of the complexity of the initial noun phrase. In contrast, in Experiment 4b when participants simply named the pictures, there was no difference in onset latencies for the displays corresponding to the complex–simple and simple–complex conditions in Experiment 4a. This pattern provides strong evidence that the complexity effect is due to phrasal planning during sentence production rather than to any visual aspect of these stationary displays.

The size of the complexity effect in Experiment 4a (63 ms) was also more similar to the size of the effect in Experiment 1a (76 ms), and greater than the reduced effect that was found in Experiment 3a (33 ms). Since the form of the experimental and filler sentences were the same in Experiments 3a and 4a, this may indicate that the reduced complexity effect found in Experiment 3a was due to random variation in the size of the effect. However, it should also be noted that the overall onset latencies were much shorter in Experiment 3a (1093 ms) compared to Experiment 4a (1275 ms), presumably because a larger set of pictures was used in Experiment 4a (192 pictures) than in Experiment 3a (144 pictures), with some of the pictures having longer naming latencies when produced in isolation. It is thus possible that the complexity effect was larger in Experiment 4a due to greater difficulty in word retrieval for the larger picture set.

6. General discussion

The present results provide support for the phrase as a default planning unit in sentence production. The experiments reported here, like those reported by Smith and Wheeldon (1999), demonstrated longer onset latencies for sentences beginning with a complex noun phrase than for matched sentences beginning with a simple noun phrase. Smith and Wheeldon had interpreted their results as indicating that participants had planned both nouns in the initial phrase prior to speech onset. The present experiments ruled out alternative explanations of their findings consistent with word-by-word planning that have to do with retrieval fluency (Experiments 1 and 2) and visual grouping (Experiments 3 and 4). The results thus support the contention of Smith and Wheeldon that it is the number of nouns in the initial phrase that is the relevant factor in causing the onset latency difference and, consequently, the present findings support a phrasal scope of planning.

As discussed in the introduction, Allum and Wheeldon (2007) have recently provided additional evidence supporting a phrasal scope of planning, examining the production of sentences by Japanese speakers to describe stationary pictures with sentences such as "The flower above the clock and the dog is red." Since all of their experiments used stationary pictures, their findings also provide evidence against grouping on the basis of movement as a source of their onset latency effects. It was the case, however, that grouping on the basis of color could have affected their results, as color was used to indicate the nouns that should be grouped together in the initial phrase. Our findings in Experiment 4, however, suggest that grouping on the basis of a visual characteristic is unlikely to be the source of their effects.

While there are now several findings in the literature supporting a phrasal scope of planning, there are other findings that support a word-by-word hypothesis regarding planning in language production – primarily results from studies using eye-tracking methodology. Some of these findings come from studies in which participants produced simple sentences such as "The turtle squirted the rabbit" to describe a picture (Griffin & Bock, 2000). For such sentences, findings indicating planning of only a single content word at a time would be consistent with a phrasal scope, given that there is only one content word in each noun phrase. Other findings consistent with word-by-word planning, however, have come from studies in which participants produced conjoined noun phrases either alone (e.g., "ball and hat") (Meyer et al., 1998) or as part of a larger utterance (e.g., "The star and the screw are above the glass,") (Griffin, 2001). However, as noted earlier, the Griffin (2001) study employed only one utterance form, as was the same case for the Meyer et al. study. In Experiments 1–3 reported here, two different target syntactic structures (simple–complex vs. complex–simple) were crossed with variation in terms of whether "above" or "below" was correct. Also, there were a large number of fillers that had structures different from the target structures. Thus, because of the lack of variation in structures in the Griffin (2001) and Meyer et al. (1998) studies, partic-

ipants may have adopted some strategy in dealing with these displays that is different from that engaged in normal speech production. As discussed by Allum and Wheeldon (2007), eye-tracking studies of language production in which production is less formulaic have revealed an initial scan of the entire scene taking about 300 ms, which precedes sequential fixation of the objects to be named. It is possible that this initial scan supports not only conceptual encoding of the scene but also the initial stages of grammatical encoding, including lemma access for words in initial phrases. In studies in which the form of the utterance is fixed or utterances of the same type are presented in blocks, this initial scan is missing (e.g., Griffin, 2001; van der Meulen & Meyer, 2000).

In addition to the possible role of this initial scan, other studies indicate that lexical information is not retrieved entirely sequentially even in the case of the simple naming of multiple objects. Morgan and Meyer (2005a,b: see also Meyer, Ouellet, & Häcker, 2008) reported results from an eye-tracking experiment indicating that participants were processing the name of a second object up to the level of phonological retrieval in parallel with the first while fixating the first object to be named. They used a design in which participants named an array of three objects. During the saccade from the first to the second object, the second object was switched to a new object. They found that, relative to an unrelated condition, shorter onset latencies for naming the second object were obtained when its name was a homophone of the original object (e.g., the animal bat switching to baseball bat). Thus, these results argue against the conclusion that gaze durations on an object are solely related to the name retrieval for that object. However, so far the evidence of name retrieval for non-foveated objects has only been obtained in the naming of multiple objects rather than in phrase or sentence production. As discussed earlier, several researchers have argued for parallel access to the content words within a phrase (Kempen & Huijbers, 1983; Schriefers et al., 1999) and thus evidence for parallel retrieval from eye-tracking would also be expected during sentence production, particularly for words within the same phrase, at least under conditions in which sentence form varies from trial to trial.

The prior discussion implies that different strategies regarding the scope of planning may be engendered by different experimental conditions. Of course, a phrasal scope of planning has to at least be an option as, in some languages, grammatical features of the head noun (such as gender and number) determine the form of preceding determiners and adjectives. A study by Ferreira and Swets (2002) provided some evidence for strategic control of the scope of planning, as they found that in the production of arithmetic sums in a carrier phrase such as "the sum is," speakers showed greater evidence of a smaller scope when under a response deadline. Importantly, this study did not show evidence of planning below the level of the phrase under a response deadline. A recent study of Damian and Dumay (2007) found evidence for phrasal planning at the phonological level for adjective–noun phrases irrespective of a response deadline. That is, even though responses sped up with the deadline, the same evidence for phonological

planning of both the adjective and noun were obtained (i.e., facilitation in onset latencies for adjective–noun phrases with a distracter word phonologically related to the noun and for phrases with the adjective and noun beginning with the same phoneme). Thus, the degree and conditions under which strategic control may be exercised has only begun to be addressed.

With regard to the retrieval fluency hypothesis, we should acknowledge that our findings do not rule out the possibility that some of the findings taken as supporting phrasal planning might, in fact, be due to speakers' trying to ensure fluency. As discussed in the introduction, findings like those of Damian and Dumay (2007) indicating phonological retrieval of the noun in adjective–noun phrases might be due the adjectives' being short and easy to encode, leading to a delay in their production until some phonology for the noun was obtained, along the lines argued by Griffin (2003). The issue could be addressed by using adjectives that are longer or more difficult to retrieve. Of course, even if evidence for phonological retrieval of the noun disappears under these conditions, it would be important to determine if lemma access has occurred.

Clearly, there are a number of unresolved issues with regard to the scope of planning in speech production. The present results rule out a particular word-by-word planning interpretation of the findings of Smith and Wheeldon (1999) and artifactual effects of the visual displays in Smith and Wheeldon (1999) and Allum and Wheeldon's (2007) studies, and consequently, provide support for a phrasal scope of planning in interpreting their results. Moreover, we would argue that there are currently no strong findings supporting word-by-word planning of multi-word phrases as a default planning scope, as the existing experiments supporting those claims used only one utterance format (e.g., Griffin, 2001; Meyer et al., 1998). In addition, studies which have shown flexibility in the scope of planning have not established that the scope can be reduced below the level of the functional phrase.

We have argued for a phrasal scope as the preferred planning scope in online language production, with word-by-word planning within a phrase only occurring when the experimental paradigm is such that the subject need only retrieve lexical items but need not perform other processing that would normally occur. Since it does appear that people can be flexible in the extent to which they plan ahead, it seems necessary to address the question of why a phrasal scope would be preferred. One possibility for this preference may relate to the fact that the phrase forms a basic unit at both the conceptual and grammatical levels (see also Allum & Wheeldon, 2007). In terms of conception, the items within a phrase are seen as being the same or similar in some way (i.e., performing the same action, occupying the same region of space, sharing some specific property, etc.). In terms of grammar, the items within a phrase would have the same thematic and grammatical roles, which would be overtly reflected in different ways in different languages (i.e., having the same case markings in languages that have them, both contributing to the inflection of the verb in languages in which that occurs, etc.). Because the phrase is a meaningful unit at both the conceptual and grammatical levels, it may then be natural

for the planning of lexical representations to occur over this scope. Further research that examines additional phrasal and sentence structures in a variety of paradigms (and languages) will provide the basis for determining whether a sub-phrasal scope at the lemma level can be established under any circumstances in which speakers must prepare sentence structures that vary from trial to trial.

Returning to the general issue of scope of planning in cognitive domains, the present findings on language production are consonant with those in motor planning and problem solving, where evidence for the advance planning of multiple units has been obtained (Catrambone, 1998; Rosenbaum, Carlson, & Gilmore, 2001). Thus, it appears to be the case that even in the language domain, in which structure and units must be selected from a large pool of possibilities, planning occurs beyond the minimal incremental level.

Acknowledgements

References

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

Alario, F.-X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implications for models of lexical access. *Language and Cognitive Processes, 17*, 299–319.

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 791–810.

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

Bock, K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*, 1–47.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). Specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1430–1450.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355–376.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*, 234–272.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers, 25*, 257–271.

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*, 178–198.

Crowther, J. E. (2009). *Mechanisms and scope of planning in language production*. Unpublished doctoral dissertation, Rice University.

Damian, M. F., & Dumay, N. (2007). Time pressure and phonological advance planning in production. *Journal of Memory and Language, 57*, 195–209.

Ferreira, F. (2000). Syntax in language production: An approach use tree-adjoining grammars. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). London: Psychology Press.

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*, 57–84.

Ford, M. (1982). Sentence planning units: Implications for the speaker's representation of meaningful relations underlying sentences. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 797–827). Cambridge, MA: MIT Press.

Freedman, M. L., & Martin, R. C. (2001). Dissociable components of short-term memory and their relation to long-term learning. *Cognitive Neuropsychology, 18*, 193–226.

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

Griffin, Z. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*, B1–14.

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

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

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

Griffin, Z. M., & Ferreira, V. S. (2006). Properties of spoken language production. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 21–59). London: Elsevier.

Hanten, G., & Martin, R. C. (2000). Contributions of phonological and semantic short-term memory to sentence processing: Evidence from two cases of closed head injury in children. *Journal of Memory and Language, 43*, 335–361.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824–843.

Jescheniak, J. D., Meyer, A. S., & Levelt, W. J. M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 432–438.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*, 185–209.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology, 12*, 433–452.

Lindsley, J. R. (1976). Producing simple utterances: Details of the planning process. *Journal of Psycholinguistic Research, 5*, 331–354.

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*, 641–682.

Martin, R. C., & Freedman, M. L. (2001). Short-term retention of lexical-semantic representations: Implications for speech production. *Memory, 9*, 261–280.

Martin, R. C., & He, T. (2004). Semantic short-term memory deficit and language processing: A replication. *Brain and Language, 89*, 76–82.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language, 41*, 3–29.

Martin, R. C., Miller, M., & Vu, H. (2004). Lexical-semantic retention and speech production: Further evidence from normal and brain-damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology, 21*, 625–644.

Martin, R. C., & Romani, C. (1994). Verbal working memory and sentence comprehension: A multiple-components view. *Neuropsychology, 8*, 506–523.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33*, 83–111.

Meyer, A. S., Belke, E., Häcker, C., & Mortensen, L. (2007). Use of word length information in utterance planning. *Journal of Memory and Language, 57*, 210–231.

Meyer, A. S., Ouellet, M., & Häcker, C. (2008). Parallel processing of objects in a naming task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 982–987.

Meyer, A. S., Sleiderink, A., & Levelt, W. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*, 25–33.

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

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

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

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*, 555–563.

Navarrete, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language, 53*, 359–377.

Oppermann, F., Jescheniak, J. D., & Schriefers, H. (2008). Conceptual coherence affects phonological activation of context objects during object naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 587–601.

Roach, A., Schwartz, M., Martin, N., Grewal, R., Brecher, A. (1996). The Philadelphia naming task: Scoring and rationale. *Clinical Aphasiology* (Vol. 24, pp. 121–134). Austin, TX: Pro-Ed.

Rosenbaum, D. A. (2010). *Human motor control* (2nd ed.). San Diego, CA: Academic Press/Elsevier.

Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). The acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology, 52*, 453–470.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchial control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 86–102.

Schriefers, H. (1992). Lexical access in the production of noun phrases. *Cognition, 45*, 33–54.

Schriefers, H., de Ruiter, J. P., & Steigerwald, M. (1999). Parallelism in the production of noun phrases: Experiments and reaction time models. *Journal of Experimental Psychology, 25*, 702–720.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73*, 205–246.

Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production—An online study. *Cognition, 78*, 123–164.

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

Spieler, D. H., & Griffin, Z. M. (2006). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes, 21*, 291–321.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistics studies. *Journal of Memory and Language, 51*, 247–250.

van der Meulen, F. F., Meyer, A. S. (2000). Coordination of eye gaze and speech in sentence production. Poster presented at the 41st annual meeting of the psychonomic society, New Orleans, LA, November.

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

Wheeldon, L. R., & Lahiri, A. (2002). The minimal unit of phonological encoding: Prosodic or lexical. *Cognition, 85*, B31–B41.