Planning Scope in Spoken Sentence Production: The Role of Grammatical Units

Paul H. Allum
Rikkyo University

Linda R. Wheeldon
University of Birmingham

Four experiments investigate the scope of grammatical planning during spoken sentence production in Japanese and English. Experiment 1 shows that sentence latencies vary with length of sentence-initial subject phrase. Exploiting the head-final property of Japanese, Experiments 2 and 3 extend this result by showing that in a 2-phrase subject phrase, sentence latency varies with the length of the sentence-initial phrase rather than that of the whole subject phrase or its head phrase. Experiment 4 confirms this finding in English. The authors' interpretation suggests that these effects derive from grammatical encoding processes. Planning scope varies according to the relation between the 2 phrases composing the subject phrase. A thematically defined *functional phrase* is suggested as defining this scope.

Keywords: speech production, grammatical encoding, planning scope

Supplemental materials: http://dx.doi.org/10.1037/0278-7393.33.4.791/supp

One of the central questions in speech production is how elements in the conceptual formulation of a message are transformed into an utterance that is structured according to the grammatical conventions of the speaker's language. Identification of a scope or scopes used in grammatical planning allows constraints to be placed on theorizing about the processes involved in this transformation. For example, if the representation of the verb in the mental lexicon is not accessed before speech onset, verb subcategorization frames cannot be an essential part of initial grammatical planning (Schriefers, Teruel, & Meinshausen, 1998).

Current experimental evidence provides conflicting results on the scope used in the grammatical planning of utterances. On one hand, there is evidence showing that speech production can be highly incremental, that humans can prepare and produce the words in an utterance one by one (e.g., Griffin, 2001; Meyer, Sleiderink, & Levelt, 1998). Such experiments provide no evidence of effects from grammatical groupings such as the phrase or the clause. On the other hand, there are results that show that phrases or clauses are advance planning units (e.g., Levelt & Maassen, 1981; Schriefers & Teruel, 1999; Schriefers et al., 1998; Smith & Wheeldon, 1999, 2001). However, it is not clear from these latter experiments what the critical units are or to which of the stages identified in mainstream theories of production (e.g., Bock & Levelt, 1994) they apply.

One reason for this lack of clarity is that research to date has focused on head-initial languages, in which the first grammatical phrase is also the subject phrase and the first noun or noun phrase within that subject phrase is the head of the subject phrase. Thus, a conceptual element, typically the agent, is confounded with the sentence-initial grammatical element, the noun phrase. The result is that the critical planning unit could correspond to a major element within the clause, such as the subject phrase, one that represents a major unit in a message-level thematic-role representation, such as the agent (cf. Chang, Bock, & Goldberg, 2003). Alternatively, planning scope could be defined by a smaller grammatical unit, such as the initial noun phrase. In the current experiments we attempt to clarify which grammatical units are preferred in utterance planning and, more important, how the scope is defined.

We use the characteristics of a head-final language, Japanese, to investigate whether a major message-level element such as agent or theme defines the scope of grammatical processing or whether some smaller unit is involved. A two-phrase subject phrase, for example, "The kite above the dog," would in Japanese give the English equivalent of "Dog$_{GEN}$ above$_{GEN}$ kite" (GEN indicates genetive and represents the function of the Japanese particle "no"). This head-final characteristic allows us to place the phrase representing the message-level element agent or theme—that is, the head of the subject phrase—as the second grammatical unit in the sentence, a noun phrase that follows a prepositional phrase,[1] potentially allowing separate effects for an initial phrase not representing a major message-level unit. This is a critical manipulation. If we can show that there is lexical access or

[1] In general, we have used the term *preposition* or *prepositional phrase* for both Japanese and English to avoid constantly switching between *preposition* and *postposition* and to make common reference to both languages easier. Japanese, of course, uses postpositions.

creation of syntactic structure for an initial subordinate phrase that does not represent a major element in the thematic representation whereas there is none for the head phrase that does represent such an element, this would, in our view, support models of speech production that posit direct links between the conceptual or thematically represented message and syntactic ordering processes (e.g., Chang, Bock, & Goldberg, 2003; Chang, Dell, & Bock, 2006).

Our basic reasoning is as follows. If the scope of grammatical planning is a phrase, and such a scope controls lexical access in advance of utterance onset (Martin & Freedman, 2001; Martin, Miller, & Vu, 2004; Smith & Wheeldon, 1999), then there is a need to explain how the conceptual message, which is presumably nonlinear, ensures that lexical access occurs in the correct order. This process is often seen as mediated by means of the thematic marking of lexical concepts affecting the grammatical function marking of the syntactic specification of the word, the lemma—for example, agent to subject—such assignment being affected by conceptual weighting on the lexical concepts in the thematic representation (e.g., Bock, Loebell, & Morey, 1992). Conceptual weighting may result in a particular lexical item being most highly activated, assigned to the role of agent, and consequently assigned to nominative and thus subject. In a configurational language like English (of the subject–verb–object language type), the grammatical system could recognize that the first element of the utterance is available, and the utterance could be initiated prior to access of later lexical items in the clause. Thus, the process of ensuring that lexical items are accessed in the correct order, given that processing is incremental, can possibly be explained within models like that of Bock and Levelt (1994) as resulting from conceptual weighting assigning elements to thematic and thus grammatical roles in the required order (e.g., McDonald, Bock, & Kelly, 1993).

However, we reasoned that such conceptual weighting cannot explain the order of activation of lexical concepts when a clear conceptual weighting factor, such as focus, animacy, or salience, is not involved. Such a situation occurs when the required order of activation and assignment is based on a purely grammatical factor, such as the convention that modifying phrases come before or after the head phrase. In this case, ordering of constituents is unlikely to be susceptible to conceptual influences in the same way that assignment to major thematic roles may be. Thus, if an initial modifying phrase is processed at the grammatical level whereas that for the following head phrase is not, we argue that this suggests an interaction between syntactic ordering information and the thematic representation of the message. For example, for the subject phrase "The dog above the table," the message level can create a thematic representation including a theme ("dog") and a modifier ("above the table"). If this abstract thematic representation interacts with a syntactic sequencing system that can indicate that a modifier comes before the concept it modifies (as in Japanese) and so influence activation and consequent selection of the modifier and its associated lexical concept "table" before that of "dog," then sequencing of lexical access for the initial word could occur without access to the lexical item for the head of the phrase, thus allowing incremental lexical access that is not based on conceptual weighting.

Experimental Evidence of Processing Scope at the Grammatical Level

Speech error evidence from natural speech (e.g., Garrett, 1988, 1993) has been taken to show that in sentence production there is at least a clausal scope for lemma access, the first step in grammatical encoding in lexically driven models (e.g., Bock & Levelt, 1994; Levelt, 1989), and a phrasal scope for positional level processes, access of lexemes and inflectional morphemes, and linearization of the constituent elements, which are seen as a second stage of grammatical encoding. As such models and the evidence from naturally occurring speech errors are thoroughly discussed elsewhere (e.g., Bock & Levelt, 1994), they are not dealt with here.

As a preliminary to the review, we define several types of phrase that are critical to understanding our arguments and the experiments that follow. In line with Schriefers and Teruel (1999), we define the *smallest full phrase* to be that part of a phrase up to and including the head noun but not including any following elements. Thus, in the French phrase "la table verte," "la table" is the smallest full phrase, but in the German phrase "der gruene Tisch," the smallest full phrase is the whole phrase, as the head is in final position. The second type of phrase we term a *simple phrase*, one that does not incorporate any additional or subelements. For example, the phrase "The house" is a simple phrase, whereas "the dog and the cat" is not; it consists of two simple phrases. "La table verte" would be a simple phrase. The third type of phrase we wish to identify is a *verb argument phrase*. This is a phrase that represents a major element in the sentence argument structure, such as subject or object, and includes all elements associated with that thematic or grammatical role. For example, in the sentence "The bird above the tree is a hawk," the first verb argument phrase is "The bird above the tree." Such a phrase may, of course, incorporate a clause—for example, "The foot which is next to the house" (cf. Smith & Wheeldon, 1999). Next we define a *verb argument head phrase* as the simple phrase that specifically represents one of the major verb arguments but not other elements associated with it. In the example above, that would be "The bird." Finally we define a *functional phrase*, one that likely represents a unit in the thematic representation of the utterance but is not necessarily one of the arguments of the verb or the head of a verb argument phrase. For example, "The dog above the table and the chair" is a verb argument phrase consisting of a simple phrase that is a head phrase representing the agent, "the dog," and a functional phrase, "above the table and the chair," representing a descriptor or modifier of the agent, that itself consists of two simple phrases. The head phrase and the functional phrase can be seen as two types of functional phrase, one that represents a major element in the verb argument structure and one that modifies such an element. They would have different assignments in a thematic representation, for example, agent and modifier, and for this reason need to be distinguished. With these definitions in mind, we refer first to experiments that give evidence of a clausal scope and then to those that support a scope smaller than a verb argument phrase. We then look at experiments that provide evidence of a verb argument phrase scope.

Evidence for a Clausal Scope

There is evidence that some aspects of a complete utterance are planned before speech onset. Longer utterances have slower utterance onset latencies (Holmes, 1988). Onset to a sentence with two clauses is longer than to the first clause alone of that sentence (Smith & Wheeldon, 1999). There is also evidence that the clause is a typical planning unit. Planning pauses in speech (Butterworth, 1980) occur more frequently at the end of finite and nonfinite clauses (Ford, 1982; Ford & Holmes, 1978; Holmes, 1988). This evidence does not make clear whether such planning is at a conceptual or grammatical level.

However, other evidence has been interpreted as showing that the clause is a *grammatical* planning unit. Semantic distractors presented at stimulus onset increase onset latency even when related to later elements in a clause (Meyer, 1996). This could suggest that lexical access, a grammatical level process encompassing access of the appropriate lexical concept and its associated syntactic information, its lemma, is completed for the whole clause, though Meyer (1996) does not exclude the possibility of a purely conceptual-level effect for semantic distractors. Latency to naming the action in a simple sentence is the same as that to naming the actor and action irrespective of the order of subject and verb, thus possibly suggesting lexical access for both prior to articulation (Kempen & Huijbers, 1983). In an utterance produced in a reiteration task, one in which participants silently prepare an utterance in advance and then at a specified signal reiterate it, probes semantically related to both early and late items slow onset latency (Dell & O'Seaghdha, 1991, 1992). This again may indicate lemma access for the whole clause. Thus, from experiments, there is evidence supporting some overall planning of the whole utterance before onset, even beyond the initial clause, and more limited evidence supporting a clausal scope that can at best be tentatively interpreted as grammatical but does not firmly exclude a conceptual interpretation.

Evidence for a Scope Smaller Than the Verb Argument Phrase

In contrast, other experimental evidence supports a far more limited degree of forward planning than the whole clause. One key experiment that appears to show that humans produce speech in a purely incremental way is reported by Griffin (2001). By using a three picture display as the stimulus, she had participants produce an utterance such as "The clock and the television are above the needle" while varying the codability and frequency of the second and third nouns. Griffin assumed that codability affects retrieval of the lemma (cf. Griffin, 2001, p. B2), whereas frequency affects retrieval of a word's phonological form (Jescheniak & Levelt, 1994). She found no effect of either codability or frequency of the second word on utterance onset latency or gaze duration to the first picture, though frequency of the first word did have an effect. She also found that once gaze moved to the second or third picture, gaze duration was affected by both the codability and the frequency of each in turn. She concluded that each lexical item was retrieved and encoded in succession. This suggests a purely incremental production process. Support for this is found in other eye-tracking experiments. Meyer et al. (1998) and Van der Meulen (2001) showed that when participants produce a coordinated noun phrase (CNP) to a picture stimulus, their attention moves strictly from left to right and that the left word is completely processed, to the phonological level, before attention moves to the right.

Others have found a subphrasal scope that is not purely linear but defined by the need to access a particular element within the utterance. Schriefers and Teruel (1999) found for German *determiner adjective noun* phrases such as "der gruene Tisch" (the green table), there was an effect for semantic distractors to the noun at a –150-ms stimulus onset asynchrony (SOA) and for the adjective at a 150-ms SOA, whereas in French "la table verte," where the adjective follows the noun, there was an effect at the early SOA for the noun but none for the adjective at the later SOA. They took this to indicate that lemma access occurred for the smallest full phrase—that is, one that included the head but not necessarily elements following the head—and thus that this was the minimum scope used in lemma access, a grammatical process.

The experiments described in the previous two paragraphs posit a scope that is either the smallest full phrase or the first simple phrase. In either case the scope is less than the verb argument phrase. Results from such experiments are, however, open to query. Participants repeatedly produced one sentence to one fixed pattern of pictures in uniform blocks (there were no fillers or alternative patterns of pictures), and it may be that processing, both visual and linguistic, was strategic rather than typical. Griffin (2001) indicated a typical gaze movement as being very similar to that noted by Meyer et al. (1998) and Van der Meulen (2001). In these experiments, when participants named a pair of objects (e.g., "scooter and hat"), they fixated on them once only, in order, and retrieved the phonology of the first word before attention shifted to the second. Such evidence suggests a purely incremental approach with planning, as evidenced in gaze, progressing only slightly ahead of articulation. Schriefers and Teruel (1999) did not use eye tracking, but the experimental design was similarly repetitious. Experiments that have participants speak short phrases (Meyer et al., 1998; Van der Meulen, 2001) or just non-syntactically-related pairs of words or numbers (Costa, Navarrete, & Alario, 2006; Griffin, 2003; Korvorst, Roelofs, & Levelt, 2006), typically in trials blocked for phrase type, produce incremental, linear processing or processing with a very limited scope (Schriefers & Teruel, 1999).

By contrast, when speakers need to apprehend the scene in order to formulate an utterance, a different pattern of eye movements is observed. Griffin and Bock (2000) had speakers describe a pictured scene in which an agent did something to a patient. Eye movements began with an initial central fixation of approximately 300 ms, followed by the strictly synchronized movement of gaze and speech evidenced in experiments where objects were always named in the same order with the same structure (see also Van der Meulen, 2001). The initial gaze appears to enable some higher level coding of the scene and may well reflect conceptualization processes, creation of the preverbal message, or initial grammatical planning. As this phase is not seen in the experiments cited in the previous paragraph that used eye tracking, it may be that such higher level formulation processes can be bypassed in such experiments. Strategic effects on planning scope resulting from experimental design have been suggested by several researchers (F. Ferreira & Swets, 2002; Korvorst et al., 2006; Meyer, 1997; Meyer, Roelofs, & Levelt, 2003; Schriefers, 1993).

There is evidence, then, that experimental design and utterance context can affect planning scope. The majority of experiments showing highly incremental processing have had designs in which strategy use may have been a factor. In particular, the difference in gaze between experiments that require scene apprehension and those that do not suggest that some higher level processing of the message is bypassed in such experiments. In addition, it has recently been shown that visual attention is not necessarily a reliable indicator of current processing scope. Morgan and Meyer (2005) showed that when utterances are made in response to a visual display, extrafoveal processing of an upcoming picture can result in the processing of its name to the phonological level before gaze has shifted to the picture. In sum, experiments that have shown very limited scope of planning and highly incremental processing do not provide conclusive evidence that normal speech processing uses such a small scope and is typically extremely incremental. Instead, they may show that highly incremental processing is possible when higher level formulation is not required or can be reduced when experimental conditions enable it.

Evidence for a Verb Argument Phrase Scope

By contrast, experiments in which participants produce whole clauses in more varied blocks suggest that the typical planning scope is longer and may cover the first phrase. Schriefers et al. (1998) showed that a distractor semantically related to the verb has no effect on onset relative to an unrelated distractor for intransitive subject–verb sentences or transitive subject–verb–object sentences (Experiment 5). They concluded that the lemma for the verb, the second main argument phrase, was not accessed before speech onset. This suggests access of the lemma, a grammatical level process, for the noun in the subject phrase but not for the verb and questions the view that the verb must be lexically accessed in order for syntactic construction of the clause to take place before utterance onset (e.g., Bock & Levelt, 1994; F. Ferreira, 2000).

F. Ferreira (1991) showed in a recall experiment that increasing the syntactic complexity of the subject phrase while holding other factors constant increases utterance onset latency (see the Experiment 1 comparison of low syntactic complexity with the average of medium and high syntactic complexity conditions). Doing the same for the object phrase does not (Experiment 2) increase the latency. Although Ferreira interpreted her results in terms of a planning scope that covers conversion of a semantic/syntactic plan into a phonological plan, the findings could be interpreted as resulting from higher level planning. Potter and Lombardi (1998) and Lombardi and Potter (1992) suggested that recall involves regeneration of the utterance from a semantic representation. Thus, Ferreira's findings could be interpreted as resulting from syntactic rather than phonological planning of the initial subject phrase.

Levelt and Maassen (1981) showed that utterances using a CNP (e.g., "The circle and the square move up") had slower onset times than those using sentence conjunctives (e.g., "The circle moves up and the square moves up") to describe the same movement. Because the second utterance is lexically and syntactically more complex than the first, they concluded that only a portion of the utterance was processed before onset and that the effect was a result of the greater processing needed for the longer subject phrase in the CNP sentences, which contains two lemmas as opposed to one.

These three experiments (F. Ferreira, 1991; Levelt & Maassen, 1981; Schriefers et al., 1998) provide fairly strong evidence within an experimental setting that the initial verb argument phrase is a planning unit, but there are reservations in relation to each of them. Schriefers et al.'s (1998) result can be interpreted as evidence for purely incremental processing. The single-word subject phrase was processed and produced before the verb lemma was accessed. F. Ferreira (1991) used an experimental paradigm, self-timed recall, that may well engender effects specific to that methodology. In addition, Smith and Wheeldon (1999) have suggested Ferreira's result might be the outcome of differences in conceptual complexity of the utterances rather than grammatical, specifically syntactic processing of the first phrase. It is thus not certain that it results from grammatical level processing of the first phrase. In the case of Levelt and Maassen (1981) there is a potential confounding between length of whole first clause and length of subject phrase, which leaves open the possibility that the effect is based on first clause length rather than subject phrase length.

Smith and Wheeldon (1999) used a paradigm similar to that of Levelt and Maassen (1981) in which participants generated an utterance in response to a moving picture display, but the authors avoided the problems inherent in their experimental design by keeping overall clause length the same while varying subject phrase length. They showed that onset latency to the sentence "The dog and the foot move above the kite" is longer than that to "The dog moves above the foot and the kite." When participants were given a preview of the pictures for the utterance, most of the difference disappeared, suggesting an effect mainly dependent on lexical access. This evidence is thus also consistent with the initial verb argument phrase being a grammatical scope in the planning process.

However, in approximately 70% of all trials, and 100% of those in which pictures appeared on the screen, a noun phrase was followed by the verb "move." (The remaining 30% were fillers in which no pictures appeared on screen.) It is conceivable that the shorter utterance onset latency to the single-noun subject phrase resulted from the fact that the second word, "move," was constantly primed, and that this had an effect in the case of the single-noun subject whereas in the CNP subject it did not. Such could be the case if planning proceeds in an incremental fashion, word by word, as suggested by eye-tracking work referred to above (Griffin, 2001; Meyer et al., 1998; Van der Meulen, 2001), but takes into account the relative phonological lengths (number of syllables) of first and second words (Griffin, 2003).

Griffin (2003) had subjects name two objects while varying the respective length of first and second words. She discovered that when a short word (averaging 1 syllable) preceded a long word (averaging 2.7 syllables), speech onset was slower than when the reverse was true. When the phrase "next to" was inserted between the two words, the effect was lost. She concluded that in order to maintain fluency, speakers are able to assess while preparing to utter the first word whether they will have time during the utterance of that word to prepare the second. If they do not, they buffer the first word briefly while doing sufficient preparation of the second to ensure that no disfluency will occur between first and second words. This

buffering slows down speech onset for the utterance beginning with the shorter word. Given the fact that initial words were very short (averaging 1.2 syllables) in Smith and Wheeldon's (1999) experiments, it could be that in the case of the single-noun subject, participants realized they had the next two syllables (the phrase "moves above") already available and thus enough advance planning to initiate the utterance without losing fluency. In the conjoined noun phrase, the conjunction "and" may not have provided enough time to access the second word after initiation of the first and so caused participants to buffer the first word a few more milliseconds. Thus, their result leaves open the possibility that production is incremental but sensitive to the phonological length of upcoming words in an utterance.

The likelihood of such an explanation, however, is greatly reduced by results reported in Martin, Knight, Tamborello, and Yang (2006). They showed that even when the verb is varied, the same effect occurs. Overall, then, there is fairly strong evidence that the initial verb argument phrase is the preferred scope of normal planning and that planning does not occur in a tightly incremental word-by-word fashion.

Outline of Experiments

Despite the evidence that the initial verb argument phrase is a planning scope, we thought it important to confirm this for Japanese. Japanese is characterized as a *topic comment* language, and topic and subject are different in some respects (cf. Li & Thompson, 1976). We needed to be sure that the topic phrase formed a planning unit in Japanese in order to make valid comparisons with results from English. Moreover, we were employing a different paradigm, using color rather than movement to group items in the stimulus pictures into phrases, which also made replication of the basic effect desirable.

Our first experiment was therefore designed to confirm that a sentence-initial verb argument phrase is indeed a planning unit during speech production and that this holds for Japanese. The verb-final properties of Japanese allow two verb argument phrases that vary from trial to trial (subject and complement) to take the first two positions in the utterance, as the examples below make clear. This allows us to avoid the possible effect of having the second element in the utterance primed and differentially influence the two conditions, a possible confound in Smith and Wheeldon (1999).

(1) *Subject   Complement*
   [Inu wa][tokei to hana] no ue ni arimasu.
   Dog$_{TOPIC}$ clock and flower$_{GEN}$ above$_{LOC}$ is.
   "The dog is above the clock and the flower."

(2) *Subject   Complement*
   [Inu to tokei wa][hana] no ue ni arimasu.
   Dog and clock$_{TOPIC}$ flower$_{GEN}$ above$_{LOC}$ is.
   "The dog and the clock are above the flower."

As the examples show (TOPIC indicates the topic marker "wa" in Japanese; LOC indicates locative, represented by the particle "ni" in Japanese), Japanese subject–object–verb order allows us to have three controlled items in the first three open-class slots while varying the syntactic structure with minimum surface change. Such a design allows us to show that the number of words within a phrase has an effect independent of any effect based on internal

monitoring of the phonological length of upcoming elements (Griffin, 2003). As with Smith and Wheeldon's (1999) design, it further avoids the confounding of initial phrase length and first clause length found in Levelt and Maassen (1981), as there is a single clause with a constant length for both conditions. Similarly, it avoids the additional memory processes necessary in a recall task (F. Ferreira, 1991).

Our following two experiments investigate the main question: What defines the phrasal scope used in grammatical planning? Is it a verb argument phrase that represents a major element within the clause—for example, subject phrase—or is it a smaller element, a functional phrase—for example, a descriptive or modifier phrase? We were particularly interested to identify whether an initial phrase that is not a major element in the clause representing one of the verb arguments forms a planning unit. The head-final characteristic of Japanese, as mentioned, allows us to dissociate the initial simple phrase and the simple phrase representing the theme in the message, as shown in (3).

(3) [Inu no ue no][hana wa] aka desu.
   Dog$_{GEN}$ above$_{GEN}$ flower$_{TOPIC}$ red is.
   "The flower above the dog is red."

The modifying prepositional phrase precedes the head phrase. Thus first phrase and head phrase within the subject phrase are not the same. This structure allows us to test whether access to a major element in the thematic representation is essential prior to utterance onset. As mentioned, we feel that this has potentially important implications for those aspects of production theories that relate to the ordered access of message-level elements, which is a key issue for incremental models. Our final experiment investigates whether effects found in Japanese also apply to English.

Experiment 1

In Experiment 1, we aimed to confirm that a verb argument phrase, the subject or topic phrase, is indeed a planning unit. The Japanese sentences (subject–complement–verb) are shown in examples (1) and (2) above. The verb occurs in sentence-final position. In both the single and coordinated subject noun phrase sentences (examples [1] and [2], respectively), the three nouns that make up the subject and complement are aligned with only a single syllable particle separating them; either the conjoining particle "to" or the topic marker "wa." The phrasal boundary can be moved from between the first and second nouns to between the second and third nouns with minimal change to the surface structure to effect the change in grammatical role of the second noun from subject to complement.

If planning is purely incremental, and if the planning unit is either the first two words or the first word, irrespective of phrasal grouping, latencies should be the same. If the planning unit is the sentence-initial verb argument phrase, then latencies to sentences like (1) should be faster than latencies to sentences like (2), replicating the effect of Smith and Wheeldon (1999).

Method

Materials. Simple black-and-white drawings of familiar objects were used. The pictures were mainly taken from Snodgrass

and Vandervart's (1980) picture norms, with the remainder being free drawn in a similar style. None of the pictures had, to our knowledge, been used with Japanese participants before. Therefore, an initial norming session was run with 26 participants to establish average naming latencies and error rates. Each participant responded to 150 pictures in a speeded picture-naming task. A total of 80 pictures with the fastest naming latencies and lowest error rates were selected from this pool, of which 48 were used as experimental pictures. These 48 were divided into three sets of pictures, providing the stimuli for three balanced sets of 16 words. The sets were matched for naming latency, standard deviation of naming latency, log frequency, word length, and error rate.[2]

Two sets of 16 trial triplets were constructed by combining the three picture sets in two different ways. Each triplet contained one picture from each set, with the condition that the words for two adjacent items could not share phonological similarity or have any obvious semantic relation. No individual picture appeared in the same combination or in the same position twice. Data for the matched sets are shown in Table 1.

The visual displays used to elicit the experiment sentences were three pictures vertically aligned in the center of a dark screen. Colors were used to signal the required grouping of nouns into subject and complement noun phrases. Colored pictures (light blue) were always to be treated as subjects. Complement pictures were white.

The 32 remaining pictures were used to construct 64 filler items. Fillers consisted of the sentence types shown in (4) and (5) below. The visual display for sentences like (4) consisted of three identical pictures colored green. For sentences like (5), the display showed three blank white squares the same size as the picture squares. In both cases pictures were vertically aligned in the center of a dark screen. A list of trial triplets can be seen in Appendix A.

(4) Subete no e wa onaji desu
    All$_{GEN}$ pictures$_{TOPIC}$ same are.
    "All the pictures are the same"

(5) E wa arimasen
    Picture$_{TOPIC}$ is not.
    "There are no pictures."

We used the "be" verb "aru" in all of our trial sentences. This verb is used with inanimate subjects, whereas the verb "iru" is used with animate subjects. Our experimental triplets included both animate and inanimate nouns. Given the nature of the experimental display and instructions to participants, our Japanese interlocutors stated that "aru" was natural in the experimental utterances.[3]

Design. A counterbalanced design was used to enable main comparisons to be made within subjects and within items. The triplets were considered unrelated items, and the unit used in the items analysis was triplet set. The assignment of triplet set to condition was rotated across participants.

The 32 experimental triplets and 64 fillers were assigned to four blocks of 24 trials. Both assignment of items to block and order of items were pseudorandom. An equal number of items from each trial set and filler type were randomly assigned to each block and then order was randomly assigned to them, with the following restrictions. Experimental items did not appear as the first item in any block, two experimental items of identical type did not appear consecutively, and two experimental items that shared a picture were separated by at least three trials. The order of block presentation was rotated across participants to avoid practice effects on items. In addition to the four experimental blocks, there were two practice blocks of 32 trials. In these blocks all of the pictures used in the experimental blocks appeared once. Experimental items were recombined with different items than in the experimental blocks and occurred in different sentence types.

Apparatus. Participants were tested individually, seated in a sound-attenuating booth facing a 17-in. Eizo Flexscan monitor (Eizo Nanao, Ishikawa, Japan) positioned approximately 90 cm away. A microphone (Model e825s; Sennheiser, Wedemark, Germany) placed in front of the participant was connected to a Nijmegen Experimental Setup (NESU)-2 box (The Max Planck Institute, Nijmegen, the Netherlands). This box was connected to

---

[2] Word frequencies were calculated on the basis of the Nippon Telegraph and Telephone Corporation (NTT) corpus, which is based on all words that have appeared in the *Asahi Newspaper* over 10 years. The kanji and kana tokens for any one word were totaled, and that total was divided by the total number of tokens in the corpus. The log of this figure was used for matching stimulus sets. The word length for Japanese words was based on the number of mora.

[3] Our first two words consisted of 16 inanimate–animate pairings and 11 animate–inanimate pairings along with 5 inanimate–inanimate pairings. Japanese colleagues were not sure which verb would be used when a coordinated noun subject consisted of a mixed-animacy combination. However, on seeing the visual stimuli (cf. Web supplement), they agreed that "aru" was natural. Our instructions (cf. Web supplement) made clear that participants were to say which object was above which; they were told to think of the displays as pictures, and the NESU display itself strongly encouraged this view. Clearly a dog was not really above a flower, but a picture of a dog was above that of a flower. We pointed out the use of "aru" in all sentences. Given our limited stimulus base, we focused on getting very closely balanced items with short words and relatively fast response latencies to bring our stimuli as close as possible to those of Smith and Wheeldon (1999). Because there was no balancing of animacy, we were unable to make statistically valid contrasts. However, whereas the average naming latency for animates used in sentence-initial position was slower than that for inanimates in the same position, latency to single-subject animates in the experimental context was slightly faster than that to inanimates. Further, the average error rate for animates used in sentence-initial position was slightly higher than that for inanimates, but in the experimental context it was slightly lower. We had only three instances of the verb "iru" being used by mistake. We think these data make it unlikely that onset was affected by a conflict between animate nouns and the verb for inanimates, "aru." Finally, as "aru" was used for all sentences, it would not be a clear confound between conditions: Single-noun subjects that were animate also used "aru."

Table 1
Matched Data Sets for Experiment 1

Measure	Set 1	Set 2	Set 3
Naming latency (ms)			
M	791	792	791
SD	163	160	163
No. of mora	2.6	2.8	2.8
Error rate (%)	1.4	1.4	1.4
Log frequency	2.7	2.7	2.6

a Compaq DeskPro II computer running DOS 6.2, on which reaction time was recorded, and to a Sony DTC-55 ES digital audiotape recorder, where the sentence spoken was recorded. The experiment was controlled by the Compaq computer in tandem with the NESU-2 box.

Procedure. Prior to the experiment, participants were informed of what picture displays they would see and what sentence they should make in response to each (sample instructions and display are provided in the Web supplement. There was only one correct sentence pattern per display; alternatives were treated as errors. Participants read the instructions, and the experimenter checked whether they knew the correct utterance by asking them to respond to the sample printed views of the display in the instructions before beginning the experiment. All pictures had very high naming agreement, as the error rate in Table 1 shows. If an alternative name was used in the practice sessions, it was pointed out to the participant. Any names other than these were treated as errors. Participants were tested individually in a sound-attenuating booth. The timing of each trial was as follows. First, a fixation point appeared in the center of the screen for 1,000 ms. The trial triplet then appeared for 3,800 ms. The screen then went blank, and there was a 1.5-s gap before the next trial began. Participants were encouraged to rest between blocks. At the end of the experiment, each participant was asked a number of questions to ascertain whether they were aware of the purpose of the experiment and to check whether they had used any strategies during the experiment. In total, an experiment lasted approximately 40 min.

Errors were recorded by the experimenter, who was seated outside the booth in front of two monitors, one of which displayed the experimental pictures and the other of which displayed progress though the experiment and reaction times. Sound in the booth was monitored from the participant's microphone through earphones. Reaction times were recorded automatically by the voice trigger in the NESU-2 box in combination with software installed on the controlling computer.

Participants. All 24 participants (11 male and 13 female) were native speakers of Japanese. All were students at Rikkyo University and had normal or corrected-to-normal vision. They were paid ¥1,000 (US $8) for their participation. In all of the experiments reported, no participant was run more than once.

Results

Separate analyses were carried out with subjects and items as a random variable, yielding F_1 and F_2 statistics, respectively. The units of analysis in the items analyses were picture triplets. The main analysis had the condition subject phrase length (single noun vs. double noun) as the independent variable and utterance onset latency and error rate as the dependent variables.

Data were excluded from the latency analyses in the following cases. Latencies of less than 300 ms or more than 3 s were regarded as outliers. Errors were also excluded. These were categorized into four types: technical errors, such as voice key failures or false triggers caused by extraneous noise; use of the wrong picture name; use of incorrect syntax; and nonfluent production, such as repair or hesitation. Incorrect syntax was any structure that did not match the target pattern for each display detailed in the instructions to participants. The latter three were concatenated into one factor, production error, for the error analyses. Missing values

Table 2

Mean Sentence Production Latencies and Percentage Error Rates for the Two Conditions of Experiment 1 (Participant Means)

Condition	Latency (ms)		Error rate (%)
	M	SD	
Single-noun phrase	1,017	211	7.5
Double-noun phrase	1,049	220	9.4
Difference	32		1.9

were replaced with the mean for that condition. Outliers and technical errors accounted for 0.6% of the data. Production errors accounted for another 9.6% of the data, making a total loss of 10.2%. One extremely error-prone item (33% loss) was considered unreliable and excluded from the analysis.

Mean production latencies and error rates are given in Table 2. As can be seen, the coordinate subject phrase sentences were initiated more slowly than the single subject phrase sentences. An analysis of variance (ANOVA) on mean sentence production latencies including the variable subject phrase length (single noun vs. double noun) showed that the 32-ms difference in latency was significant by both subjects and items, $F_1(1, 23) = 6.25$, $MSE = 1,716$, $p < .05$; $F_2(1, 30) = 5.00$, $MSE = 3,102$, $p < .05$. Error rates were higher in the slower condition, suggesting there was no speed–accuracy trade-off. A similar analysis of error rates showed no significant effect (F_1 and $F_2 < 1$).

To determine whether participant performance remained consistent across the experiment, we conducted an ANOVA including the additional variable block position (1 to 4). There was no main effect of block position (F_1 and $F_2 < 1$), nor any interaction with subject phrase length, $F_1(3, 69) = 2.37$, $MSE = 5,488$; $F_2(3, 90) = 1.61$, $MSE = 9,105$. Such lack of effect may in part have been a result of the practice blocks, in which all words used appeared once and which gave ample practice with the structures.[4]

Together, these results show that the effects obtained were strictly dependent on the length of the subject phrase. There was no clear evidence of effects from other variables or of strategy use in the data. In the debriefing, participants showed no awareness of experiment aim or design.

Discussion

In Experiment 1, sentence production latencies increased with the length of the sentence-initial verb argument phrase, suggesting that, minimally, this phrase was planned prior to speech onset. This result replicates the findings of Smith and Wheeldon (1999) in a language with very different syntactic properties. The verb-final property of Japanese allowed us to demonstrate that priming of the verb was not a significant factor in the results obtained by them. Our data also argue against strongly incremental accounts of

---

[4] In the three experiments reported below there was no interaction with the block variable. This may have resulted partly from the appearance of all items in the practice blocks and the fact that there was plenty of practice. It also demonstrates that effects were robust across the course of the experiment. These data are not therefore reported.

grammatical encoding (Griffin, 2001, 2003). The first two words are equivalent in all important aspects in both sentence types, and there is a single syllable particle intervening in both cases. Viewed purely incrementally, there is no difference between the two sentence types. Moreover, they argue against the suggestion made by Schriefers and Teruel (1999, 2000) that the smallest full phrase is a planning unit. The smallest full phrase in both conditions is an initial simple noun phrase (e.g., "dog"). Thus, if access to the head of the first grammatical unit, the simple noun phrase, was all that was required for utterances to be initiated, we would expect no difference in onset to these two sentence types. However, our results leave open the possibility that access to the head(s) of the initial verb argument phrase, the agent(s) or, in our sentences, theme(s) in the conceptual plan, could be requisite before utterance onset and that this could define initial planning scope, or that it could be defined by the whole verb argument phrase or the first functional phrase. We attempt to separate these three elements in Experiment 2.

Experiment 2

Experiment 1 indicates that it is not the first word, smallest full phrase, or simple phrase that defines planning scope. However, whole verb argument phrase, head of verb argument phrase (both noun phrases in the CNP are head phrases), and functional phrase are confounded. It therefore does not allow us to distinguish which of these three might define planning scope. There is evidence that access to the head of the first phrase, but not necessarily beyond, is required prior to speech onset and that this may thus define planning scope (Martin & Freedman, 2001; Martin et al., 2004; Schriefers & Teruel, 1999). There is also evidence that a posthead subordinate clause within the subject phrase is less thoroughly processed in terms of lemma access, as demonstrated by Smith and Wheeldon (1999). Although the exact nature of the difference in degree of processing is not discussed, they showed a greater effect for preview of a second noun within the subject phrase when it was part of a CNP than when it was in a modifying relative clause. It could be that the lemma for the second noun in the CNP is accessed prior to speech onset, whereas only the lexical concept or higher level processing occurs for that in the relative clause. These data suggest that different levels of processing occur within the initial verb argument phrase. As mentioned above, in Japanese the subject phrase "The dog above the table" has the prepositional and noun phrases reversed, giving roughly the phrase "Table$_{GEN}$ above$_{GEN}$ dog" ("Teeburu no ue no inu"). The initial simple phrase, the prepositional phrase, precedes the subject head noun phrase and thus allows us to investigate the role of a functional or subordinate phrase, in contrast to the whole verb argument phrase or the head of that phrase. We can thus separate scope that might depend on access to the head of the verb argument phrase from scope defined by the first phrase, a functional phrase. This allows us to investigate whether access to a major thematic element, the theme, which is the head of a verb argument phrase, plays a role in defining planning scope.

We tested the production of utterances such as those shown in (6) and (7) below. In sentences such as (6), a verb argument phrase consisted of a functional phrase and a head phrase, a modifier and theme, represented in the hierarchically distinct prepositional phrase and noun phrase. In sentences such as (7), a verb argument phrase consisted of two hierarchically equal simple noun phrases performing the role of head(s) of the argument phrase and therefore both representing the theme in the message-level thematic representation (7).

(6) [Inu no ue no] [hana wa] aka desu.[5]
    Dog$_{GEN}$ above$_{GEN}$ flower$_{TOPIC}$ red is.
    "The flower above the dog is red."

(7) [Inu to hana wa] aka desu.
    Dog$_{CONJ}$ clock$_{TOPIC}$ red are.
    "The dog and the clock are red."

If the initial verb argument phrase, the subject phrase, is a planning scope, onset to these two types of phrases should be the same or perhaps longer in the case of sentence (6) due to greater grammatical complexity. Given the head-final nature of Japanese, if access to the lexical item(s) that form the head of the verb argument phrase is required, then we would also expect onset latency to be the same for both types of phrase because the CNP consists of two head words, and the head word in the prepositional-phrase-modified subject phrase (PP) occurs in the second position in the phrase, thus requiring access to both words before onset. If access to only the first phrase that serves one function in the thematic representation, a functional phrase, is required, then onset to the CNP should be longer, as the head phrase(s) of the verb argument phrase have two lexical items that together serve one function as opposed to only one lexical item in the first functional phrase, the prepositional phrase, in the other.

Method

Materials. A total of 42 pictures were selected from the same pool of items described in Experiment 1. Of these, 32 were used as experimental pictures. These were divided to provide stimuli for four balanced sets of 16 words. The sets were matched on the same criteria as those in Experiment 1. Data for the matched sets are shown in Table 3.

The four sets of 16 trial pairs were constructed by combining these two picture sets in four different ways. Two sets of trial pairs were assigned to each condition for each subject. Each trial pair

---

[5] A reviewer pointed out that in Japanese, one usually informs someone of the color of something either by using the adjectival verb "akai" with or without "da" (a "be" verb) (e.g., "Kono kaban wa akai desu," This bag is red) or by adding "the color of" and using the noun form "aka" with a "be" verb (e.g., "Kono kaban no iro wa aka desu," This bag's color is red). "Aka" refers to the color itself, as in the sentence "Aka ga suki desu" (I like red). But it also can be used in sentences like "Shingo wa aka desu" (The signal is red). However, in response to a question like "Ano kaban wa nani iro desu ka?" (What color is that bag?), "Aka desu" (It is red) is a normal reply. The noun form "aka" with the verb "da" (polite form "desu") is frequently used when the information sought is already identified by the context as the color of something. Within our experimental context, participants were instructed that they should identify which object was what color. Thus, informing the color of the appropriate picture was the aim of nearly every utterance. Given this context, and the nature of the NESU display, our Japanese interlocutors preferred "Aka desu." The use of the polite form "desu" was also considered most appropriate by them, given the fact that participants were in an unfamiliar experimental situation with someone senior to them listening.

Table 3
Matched Data Sets for Experiment 2

Measure	Set 1	Set 2
Naming latency (ms)		
M	789	790
SD	168	163
No. of mora	2.7	2.7
Error rate (%)	1.4	1.5
Log frequency	2.6	2.6

comprised one picture from each set, with the condition that the words associated with the two items should not share phonological similarity or have any obvious semantic relation. Each picture appeared four times, resulting in the associated word occurring in the same sentence position twice but in a different phrase type and in a different combination (i.e., once in first position for the CNP type, once for the PP type, and once in second position for each of the phrase types). In other words, any word appeared in first position for, for example, the CNP condition only once. Thus, items were treated as independent. The aim of using the "above" or "below" pattern with the prepositional phrases was to prevent a simple top-down strategy. The visual display for sentences like (6) showed a colored picture above or below a white picture. For sentences like (7), it showed two identically colored pictures. In both cases the two pictures were vertically aligned center screen. Four colors were used for experimental items: green, blue, red, and brown.

The remaining 10 pictures were used to make 40 fillers of sentence type (8), each picture appearing as a pair four times. The visual display for these was two gray pictures. The remaining 40 fillers were sentences like (9). The visual display was two blank squares the same size as the picture squares. In both cases the two pictures were vertically aligned in the center of a dark screen. A list of trial sets can be seen in Appendix B.

(8) Ryoo hoo no inu wa haiiro desu.[6]
  Both$_{\text{PART GEN}}$ dogs$_{\text{TOPIC}}$ gray are.
  "Both dogs are gray."

(9) E wa arimasen
  Picture$_{\text{TOPIC}}$ is not.
  "There are no pictures."

Design. The same counterbalanced design as in Experiment 1 was used. The pairs were treated as unrelated items, and the unit used in the items analysis was pair set. The 64 experimental pairs and 80 fillers were assigned equally to four blocks of 36 trials. Both assignment of items to block and order of items were pseudorandom, following the same restrictions as in Experiment 1. Assignment of experimental pair to condition was rotated across participants, as was block order, to avoid practice effects on items.

In addition to the four experimental blocks, there were two practice blocks. All of the pictures used in the experimental blocks appeared twice. Experimental items were recombined (never with the same items as in the experimental blocks) and used in the same pattern as in one of the trial conditions, half in each, as well as appearing once in type (8) fillers. The filler pictures were recombined four times to make 20 pairs, half of which appeared in trial

pattern (6), half in (7). There were also 10 of type (9), making a total of 76 practice trials, divided into two blocks of 38 trials.

Procedure. The procedure was the same as for Experiment 1 with the following differences in timing. First, a fixation point appeared on the screen for 1,000 ms. The stimulus pair then appeared for 4,000 ms. The trial display time was 200 ms longer than in Experiment 1. This is because the task was considered slightly more difficult: Participants had to apprehend whether the subject was above or below the other object. In addition, in Experiment 1, the error rate had been relatively high, and this could well have been a result of participants feeling too rushed. This is also the reason that the intertrial gap was extended by 500 ms to 2 s.

Participants. All 16 participants (9 male and 7 female) were native speakers of Japanese. All were students at Rikkyo University and had normal or corrected-to-normal vision. They were paid ¥1,000 (US $8) for their participation.

Results

Data trimming followed the same principles as in Experiment 1. Outliers and technical errors accounted for 0.6% of the data. Production errors accounted for another 10.7%, making a total loss of 11.3%. Analyses were carried out on the same basis as for Experiment 1.

Mean production latencies and error rates are given in Table 4. As can be seen, the onset latency was longest for the CNP subject type.

An ANOVA on latencies including the variable subject phrase type (PP vs. CNP) showed that there was a significant main effect for subject phrase type by both subjects and items, $F_1(1, 15) = 11.17$, $MSE = 4,899$, $p < .05$; $F_2(1, 63) = 21.69$, $MSE = 10,095$, $p < .05$. Error rates were the same in both conditions and yielded no significant effect (F_1 and $F_2 < 1$).

Discussion

Sentence production latencies in Experiment 2 increased with the size of the sentence-initial functional phrase: The prepositional phrase served one function in the PP condition, and the head phrase(s) served one function in the CNP condition. This result shows that even when length of verb argument phrase, the subject phrase, is the same as far as open-class words are concerned, sentence production latency differs according to the grouping of elements within the phrase. When both elements are head elements and form one unit (theme) from the point of view of the thematic representation, onset is longer than when there are two units, a head element and a subordinate with different roles in the thematic representation, modifier and theme. When the initial functional phrase consists of two words, onset is longer than when it consists of one. Given the fact that the length of the first phonological word

---

[6] One reviewer commented that the Japanese structure used for this filler was awkward or less natural than two other options, "Inu wa nihikitomo haiiro desu" or "Inu wa ryoohootomo haiiro desu." We were aware of the awkwardness of our structure but decided to keep it because in most utterances a noun phrase started the utterance. We wanted to break up this pattern even if only a little to avoid effects of constant repetition of an initial phrase type.

Table 4

Mean Sentence Production Latencies and Percentage Error Rates for the Two Conditions of Experiment 2 (Participant Means)

	Latency (ms)		Error rate (%)
Condition	M	SD	
PP-modified subject	1,049	231	5.5
CNP subject	1,131	232	5.5
Difference	82		0

Note. PP = prepositional phrase; CNP = coordinated noun phrase.

(Wheeldon & Lahiri, 1997) in either utterance is the same, this effect is likely to have arisen at a higher level than the phonological level.

However, although Experiment 2 shows that it is not the whole subject phrase alone that determines scope of initial planning, three explanations other than the different combination of phrase types might be suggested. The first is that with the CNP phrase, buffering is used to ensure fluency, whereas with the PP phrase it is not necessary. The PP subject type has the initial word followed by three very high frequency items, the particle "no," the noun "ue" or "shita," and the particle "no," which together function as a closed-class preposition. The CNP has only one particle, "to," between the first and second words. This could have differentially affected speech onset. Our initial word length averaged 2.7 mora, which, with the addition of the linking particle "to" for the CNP phrases, would give participants 3.7 mora to prepare the next word, as opposed to 6.7 mora for the PP phrases. It is conceivable that if buffering of short initial words is used to ensure downstream fluency, then that was more likely to have occurred with the CNP phrase type than with the PP phrase type, owing to the difference in number of closed-class items intervening between first and second words. Closed-class items incur minimal preparation time but give extra time during their articulation for a following item to be prepared (cf. Griffin, 2003). Experiment 1 argues against this, as the intervening particle between first and second words was either "wa" for the single-noun subject or "to" for the CNP subject. Thus, the type of buffering referred to above could not account for that effect. However, that effect also involved different phrases for the two conditions. We therefore felt any such buffering explanation still needed to be firmly ruled out.

A second possibility concerns the status of the middle word of the Japanese prepositional phrase "no ue/shita no." It might be claimed that the arguably open-class "ue" (above) or "shita" (below)[7] acts as a second word that is more highly primed than the second word in the CNP. Such an explanation is unlikely but not impossible. Third, and most important, an apparently counterintuitive result—that is, a subordinate phrase being more thoroughly processed at the grammatical level, for example, in terms of either lexical access or creation of syntactic structure, than the head phrase it modifies—requires further confirmation. Experiment 3 was run to provide a stronger test of the effect of the first, subordinate phrase.

Experiment 3

Experiment 3 was designed to determine clearly whether an initial subordinate, functional phrase within a verb argument

phrase, a subject or topic phrase, is a planning unit. To do this, we used a topic phrase similar to that in Experiment 2 that consisted of a subordinate prepositional phrase, representing a modifier in the thematic representation, and a head noun phrase, representing the theme in the thematic representation. However, this time, the phrase consisted of four simple phrases grouped into two functional units. The length of the first unit, the prepositional phrase, was varied relative to that of the second, the head phrase, while keeping the overall subject phrase length the same. The following sentences were used.

(10) *Prepositional phrase   Head noun phrase*
[Inu no ue no] [hana to zubon to ringo wa] aka desu.
Dog_GEN above_GEN flower and trousers and apple_TOPIC red are.
"The flower and apple and trousers above the dog are red."

(11) *Prepositional phrase   Head noun phrase*
[Inu to hana no ue no] [zubon to ringo wa] aka desu
Dog and flower_GEN above_GEN trousers and apple_TOPIC red are.
"The trousers and apple above the dog and flower are red."

(12) *Prepositional phrase   Head noun phrase*
[Inu to hana to zubon no ue no] [ringo wa] aka desu
Dog and flower and trousers_GEN above_GEN apple_TOPIC red is.
"The apple above the dog and flower and trousers is red."

If speech onset increases in relation to the length of the first functional phrase, then we would expect onset to increase between sentences (10) and (11) and between sentences (11) and (12) because the functional phrase increases in length by one word in each case. If a buffer is necessary, as suggested by Griffin (2003), and the particle "to" used in the conjoined noun phrase is not sufficient for that purpose, or if the second debatably open-class word "ue" was primed in the prepositional phrases in Experiment 2, thus providing an explanation for the faster onset, then we would expect an increase between sentences (10) and (11) but not between sentences (11) and (12). Any increase in onset latency from a two-phrase to a three-phrase initial phrase is highly unlikely to result from buffering (Griffin, 2003).

Method

Materials. A set of 82 pictures were selected from the same pool of items used in Experiment 1. A subset of 64 of these were used as experimental pictures. They were made into four sets of 16 pictures balanced on the same criteria as those in Experiment 1. These were combined three ways, with the same restrictions as in Experiment 1, to make three sets of 16 trial quadruplets. Thus, each picture was used three times in one complete subject set of 48 quadruplets, but each quadruplet was unique. No picture appeared in the same position more than once. Data for the matched sets are shown in Table 5.

---

[7] Unlike English prepositions, Japanese postpositions can be formed by use of a particle ("no"), a word that is sometimes classed as a noun ("ue"), and a following particle ("no"). Particles are closed class, and the postpositions thus formed must be a closed class. But it is difficult to classify "ue" itself as closed class in the sense in which a particle in Japanese or a preposition in English may be. Though not an exactly equivalent phrase, a useful comparison might be with the phrase "on top of" in English.

Table 5
Matched Data Sets for Experiment 3

Measure	Set 1	Set 2	Set 3	Set 4
Naming latency (ms)				
M	774	775	774	773
SD	158	155	158	160
No. of mora	2.8	2.8	2.8	2.9
Error rate (%)	1.4	1.4	1.4	1.2
Log frequency	2.8	2.8	2.8	2.8

The visual display for type (10) sentences was a single colored picture above or below three horizontally aligned white pictures; for type (11), it was two same-colored pictures above or below two white pictures; for type (12), it was three same-colored pictures above or below one white picture. Four colors were used: red, blue, green, and brown. Pictures were centered and horizontally aligned approximately 1 cm above and below the screen center.

There were 72 fillers. They consisted of one of the two following sentence types.

(13) Subete no usagi wa aka desu
All$_{GEN}$ rabbits$_{TOPIC}$ red are.
"All the rabbits are red."

(14) E wa arimasen
Picture$_{TOPIC}$ is not.
"There are no pictures."

The 18 remaining pictures were used to make the first type of filler. They were used twice, accounting for 36 of the fillers. Blank pictures (empty white squares of the same size as the picture squares) were used for the second type, accounting for the other 36 fillers. For filler type (13), four vertically and horizontally centered, horizontally aligned same-colored pictures appeared. The four colors were red, blue, green, and brown. For filler type (14), four vertically and horizontally centered, horizontally aligned blank (white-colored) pictures appeared.

Design. The 48 experimental quadruplets and 72 fillers were divided equally into four blocks with pseudorandom assignment to block under the same restrictions as in Experiment 1. In addition, an experimental item was never preceded by a filler using the same color.

In addition to the four experimental blocks, there were two practice blocks. All of the pictures used in the experimental blocks appeared twice. Experimental items were recombined (never with the same items as in the experimental blocks) and used in the same pattern as in one of the trial conditions. The filler pictures appeared once each in type (13) sentences. Twelve fillers of type (14) were also used to make a total of 62 practice items, divided into two blocks of 31 items.

Procedure. The procedure was the same as for Experiment 1 with the following differences in timing. First, a fixation point appeared on the screen for 1,000 ms. The quadruplet then appeared for 5,500 ms in one of the three configurations. The display time was increased from Experiments 1 and 2 because participants had to process four pictures, as opposed to two or three, and decide whether the subject was above or below for the experimental items. Then there was a 1.5-s gap before the next trial began.

Participants. All 36 participants (16 male and 20 female) were native speakers of Japanese. All were students at Rikkyo Univer-

sity and had normal or corrected-to-normal vision. They were paid ¥1,000 (US $8) for their participation.

Results

One participant was far slower than the others, taking 1,951 ms to respond (against an average of 1,297 ms; $SD = 283$ ms). One item sustained losses of just over 30%. Data from these two sources were considered unreliable, and they were removed from the analysis. Data trimming was the same as for the previous two experiments. Outliers and technical errors accounted for 0.6% of the data. Production errors accounted for another 10.7% of the data, making a total loss of 11.3%. The units of analysis in the items analysis were picture quadruplets.

Mean production latencies and error rates are given in Table 6. As can be seen, the onset latency increased as the length of the prepositional phrase increased.

An ANOVA on latencies featuring the variable preposition phrase length (single word, double word, or triple word) showed that there was a significant main effect for phrase length by both subjects and items, $F_1(2, 68) = 30.03$, $MSE = 3,086$, $p < .05$; $F_2(2, 92) = 18.70$, $MSE = 6,824$, $p < .05$. Post hoc Bonferroni pairwise comparisons showed that all differences were significant in both subject and item analyses ($p < .05$).

Error rates were higher in the slower condition; however, a similar analysis showed no effect by subjects or by items, $F_1(2, 68) = 1.03$, $MSE = 0.02$; $F_2(2, 92) = 1.02$, $MSE = 0.02$.

Discussion

In Experiment 3, we found that sentence onset latencies increased in relation to the length of the first functional phrase when the length of the first verb argument phrase, the subject phrase, was kept constant. This finding clearly indicates that the first functional phrase within the subject phrase undergoes more thorough high-level processing than the second phrase, the head phrase, and that such processing is not a function of the first phrase being the head phrase. The likely locus, as suggested earlier, is either lexical access or creation of syntactic structure. Detailed discussion of these possibilities is left to the General Discussion.

Experiment 3, then, offers robust confirmation of the result from Experiment 2. All previous experiments that have shown the first verb argument phrase as being more thoroughly processed have either consisted only of a head phrase or had the head phrase in the initial position. This result supports the results of Experiments 1

Table 6
Mean Sentence Production Latencies and Percentage Error Rates for the Three Prepositional Phrase (PP) Conditions of Experiment 3 (Participant Means)

Condition	Latency (ms)		Difference (ms)	Error rate (%)	Difference (%)
	M	SD			
Single-word PP	1,220	242		9.8	
			57		1.7
Double-word PP	1,277	232		11.5	
			48		0.9
Triple-word PP	1,325	254		10.6	

and 2 in showing that it is not the lowest level grammatical unit, noun phrase, but a higher level grammatical unit that serves as a planning unit. It thus reinforces the interpretation that higher level planning scope is related to a certain level within the grammatical hierarchy, one that we have suggested represents a function in the thematic representation.

The result of this experiment also reinforces the conclusion drawn from Experiments 1 and 2 that sentence production latency is not determined simply by buffering requirements (Griffin, 2003). It is not at all likely, given evidence on degree of advance phonological planning (e.g., Meyer, 1996; Wheeldon & Lahiri, 1997), that the difference in onset from a two-noun prepositional phrase to a three-noun prepositional phrase could result from any phonological processing. Thus, the latency increase must be a result of higher level processing.

Experiment 4

Although there does not seem to be a strong prima facie reason why the above conclusions should be specific to Japanese, it is also true that there are substantial differences between Japanese and English that might be relevant. For example, as mentioned earlier, Japanese is a topic comment language, and topic and subject can play different roles (Li & Thompson, 1976). The prepositional phrase in Japanese is created using the "no" particle and the debatably open-class category noun "ue" (above). It is not certain what effects such linguistic differences might have. In the Japanese experiment, there are three words, "no ue no," between the first and second trial words for the prepositional phrase but only one word, "to," between them in the CNP. Although buffering effects seem to be ruled out by the previous experiments, there might be some advantages at higher levels (e.g., access to the second lemma, "ue") that speed response to initial prepositional phrases. Further, we wished to see whether there was support for our basic supposition that initial message-level processes are similar across languages, as we will suggest that their scope may be conceptually defined. Whether in Japanese or English, the initial functional unit is similar. Therefore, if functional unit is a key factor in defining scope, it should function similarly across languages. Experiment 4, therefore, replicated Experiment 2 in English. It allowed us to check that our results were not likely to be a result of unspecified differences between Japanese and English. In English the difference between the two phrases is minimal in terms of number of words or number of syllables as well as in terms of word category, open versus closed. The following sentence structures were tested.

(15) The dog above the flower is red.

(16) The dog and the flower are red.

The prediction is that in English, as in Japanese, the onset to the CNP subject phrase will be slower than that to the PP phrase because the initial functional phrase is longer in the former than in the latter.

Method

Materials. A set of 42 simple black-and-white drawings of familiar objects was used. These were mainly taken from Snodgrass and Vandervart's (1980) picture norms, with the re-

mainder being free drawn in a similar style. All pictures had been extensively tested in a simple picture-naming paradigm (cf. Wheeldon, 1989; Wheeldon & Monsell, 1992). The naming latencies and percentage error rates used in the selection and matching of items were based on these data. Word frequencies were calculated by averaging the orthographic token and stem frequency count for noun uses in Kučera and Francis (1967), the orthographic token count from Hofland and Johansson (1982), and the same count summed with the count for any orthographic tokens that could be considered the stem noun plus suffix.

Of these drawings, 32 were used as experimental pictures. These were divided into two sets of 16 pictures whose associated words were matched on the same criteria as those in Experiment 1. Data for the matched sets are shown in Table 7.

The four sets of 16 trial pairs were constructed by combining these two sets in two different ways in exactly the same way as in Experiment 2. There was one change in terms of display. In the Japanese experiment, it had been thought that varying the picture position of the initial noun in the visual display might reduce the possibility of strategic processing, but it was found that responses when the initial word was stimulated by the lower picture were slower than when it was stimulated by the higher picture. In either position, response to the PP pattern was still faster than to the CNP; however, position might be considered a confounding factor. In the English experiment, therefore, the picture representing the first word in the PP pattern was always on top. Except for this change, the construction of the experiment was kept the same as Experiment 2. Thus, the visual display for the PP pattern for sentences like (15) showed a colored picture above a white picture. For sentences like (16), it showed two identically colored pictures. In both cases the two pictures were vertically aligned center screen. Four colors were used for experimental items: green, blue, red, and brown.

The remaining 10 pictures were used to make 40 fillers of type (17), each picture appearing as a pair four times. The visual display for these was two gray pictures. The remaining 40 fillers were sentences like (18). The visual display was two blank squares the same size as the picture squares. In both cases the two pictures were vertically aligned in the center of a dark screen.

(17) Both dogs are gray.

(18) There are no pictures.

Design. The same counterbalanced design as in Experiment 2 was used. The pairs were treated as unrelated items, and the unit used in the items analysis was pair set. The construction of the

Table 7
Matched Data Sets for Experiment 4

Measure	Set 1	Set 2
Naming latency (ms)		
M	566	567
SD	180	181
No. of syllables	1.6	1.7
Error rate (%)	0.2	0.5
Log frequency	2.3	2.3

experiment, including the two practice blocks, was exactly the same as for Experiment 2.

Apparatus. Participants were tested individually, seated in a sound-attenuating booth facing a Gateway 2000 15-in. monitor positioned approximately 90 cm away. A Sennheiser e825s microphone placed in front of the participant was connected to a NESU-2 box. This box was connected to a Gateway G6-266 computer running DOS 6.2, on which reaction time was recorded, and to a Sony DTC-55 ES digital audiotape recorder, where the sentence spoken was recorded. The experimenter was seated outside the booth in front of two monitors, one of which displayed in real time the same pictures as the participant was viewing and the other of which displayed progress though the experiment and reaction times. Sound in the booth was monitored from the participant's microphone through earphones. The experiment was controlled by the Gateway computer in tandem with the NESU-2 box.

Procedure. The procedure and timing were the same as for Experiment 2.

Participants. All 16 participants (6 male and 10 female) were native speakers of English. All were students at the University of Birmingham and had normal or corrected-to-normal vision. They were paid £4.50 or 45 course credits for their participation.

Results

As with the Japanese experiments, items faster than 300 ms or slower than 3,000 ms were treated as outliers. Together with technical errors they accounted for 2.6% of the data. Response errors accounted for another 9.5% of the data, making a total loss of 12.1%. The units of analysis in the items analysis were picture pairs.

Mean production latencies and error rates are given in Table 8. As can be seen, the onset latency was once again longest for the CNP subject type.

An ANOVA on latencies featuring the variable subject phrase type (PP vs. CNP) showed that there was a significant main effect for subject phrase type by both subjects and items, $F_1(1, 15) = 31.38$, $MSE = 3,026$, $p < .05$; $F_2(1, 63) = 65.12$, $MSE = 5,833$, $p < .05$. Error rates were higher in the slower condition. A similar ANOVA on percentage error rates showed a significant effect by items only, $F_1(1, 15) = 3.05$, $MSE = 0.00$; $F_2(1, 63) = 5.63$, $MSE = 0.01$, $p < .05$.

Table 8

Mean Sentence Production Latencies and Percentage Error Rates for the Two Conditions of Experiment 4 (Participant Means)

Condition	Latency (ms)		Error rate (%)
	M	*SD*	
PP-modified subject	936	213	7.4
CNP subject	1,045	265	11.5
Difference	109		4.1

Note. PP = prepositional phrase; CNP = coordinated noun phrase.

Discussion

The results were the same as those for Japanese in Experiment 2, thus giving initial indication that language differences are not likely to be the cause of that result. In both experiments, the prepositional function words follow the first lemma, and in both cases the initial verb argument phrase is slightly longer in the PP condition than in the CNP condition. Despite the increased length, the onset to the verb argument phrase that includes the prepositional phrase is significantly faster. These results give initial indication that scope is similarly defined across the two languages.

General Discussion

The four experiments described here provide evidence concerning the preferred scope of grammatical planning in speech production. Experiment 1 demonstrated that onset latency was related to the size of the first verb argument phrase, the subject phrase. However, Experiments 2 and 3 showed that when a subject phrase comprised two smaller units, which we have described as two types of functional unit (i.e., a functional phrase and a verb argument head phrase), sentence production latencies were related to the size of the first of these rather than simply to the size of the subject phrase as a whole. This effect occurred whether the sentence-initial functional phrase was the head of the subject phrase (as in English, Experiment 4) or a subordinate prepositional phrase (as in Japanese, Experiments 2 and 3).

The pattern of results we have reported suggests that a head phrase, when it falls in second position in the initial verb argument phrase, is not as thoroughly processed as the initial functional phrase, even if that is a subordinate phrase. Moreover, it is unlikely that this effect could be explained in terms of lower level phonological buffering. In Experiment 1 there was no difference in the phonological length of the first two words and intervening topic particle. Nevertheless, onset latencies varied systematically with length of the initial phrase. In Experiment 3, the effect was seen to be cumulative and to occur when a two-noun initial functional phrase was increased to a three-noun phrase. There is currently no evidence that phonological planning extends this far ahead.

These findings extend those of Smith and Wheeldon (1999, 2001). First, their claim that the processing of the first verb argument phrase is more thorough than for subsequent phrases is strongly supported. However, the experiments here provide a major extension. A phrase that is a smaller unit than the verb argument phrase is more thoroughly processed than that phrase as a whole. The four experiments demonstrate that the subject phrase alone is not the only higher level planning unit but rather there is more thorough high level processing for the first phrase, even if it is not the head of the verb argument phrase. The implications are discussed later, but first we attempt to offer a clearer definition of this type of phrase and examine how such a scope might fit with previous experiments.

We have used the term *functional phrase* to refer to a certain level within the grammatical hierarchy. Such phrases are neither a large phrase representing a major element, such as subject or object phrase (which we have termed *verb argument phrases*), nor the smallest grammatical phrase (e.g., simple noun phrase). We have defined this level of phrase within the grammatical hierarchy in terms of function. In other words, the phrase represents a

conceptual function such as modifier, theme, or agent. For example, in the sentence "The flower above the dog is red," the phrase "The flower" is the theme, and the phrase "above the dog" acts as a modifier. Exchanges between phrases in this type of relation, where a prepositional phrase modifies a head noun phrase, are quite common (e.g., Berg, 1987; Vigliocco & Hartsuiker, 2002), and it has been suggested that each phrase represents a separate function within the process of functional assignment, as evidenced by *phrasal exchanges* (Vigliocco & Zorzi, 1999). The level within the grammatical hierarchy that our results indicate to be a preferred planning scope may well be equivalent to the smallest functional role, even though that may not be one that plays a role in relation to the verb (i.e., one of the main roles in the clause directly designated by verb subcategorization). In other words, "The flower above the dog" has the function of theme within the sentence, but that function is composed of two distinct functions, and it is these nonreducible functions that are preferred planning units in utterance production.

How would such an analysis sit with earlier online production work indicating a phrasal scope? There is no conflict with the result found by Schriefers et al. (1998), as they showed that the lemma for a single-noun subject was accessed prior to onset but not that for the following verb phrase. Their experiment did not distinguish between head phrase and subsidiary phrase, having only a one-word subject that can be seen as a minimal functional unit. It is also consistent with Levelt and Maassen's (1981) result, as they found that onset to a clause with a CNP subject was longer than that to a clause with a single-noun-phrase subject. Further, it is consistent with Smith and Wheeldon's (1999) finding that the first phrase is processed more thoroughly. This conclusion was based on the fact that a sentence with a CNP subject/single-noun-phrase object has a longer onset than one with a single-noun subject/coordinated-noun object when the overall clause length is kept constant. It is also consistent with their finding that the second functional element within the subject phrase is not as thoroughly processed as the first; onset to a verb argument phrase, a subject phrase, consisting of a CNP is slower than that to a phrase consisting of a single noun modified by a relative clause. The clause serves a different function than the head phrase.

In sum, a functional unit account is consistent with much previous work in the field. In relation to experiments that show more incremental processing, the result could be in agreement with Schriefers and Teruel (1999). Recall that they found an effect for semantic distractors to the noun at an SOA of –150 ms and adjective at 150 ms for German determiner–adjective–noun phrases such as "der gruene Tisch" (the green table) but found only an effect for the noun, not the adjective, for the French phrase "la table verte," where the adjective follows the noun. If we assume that scope of planning extends to the head of the first functional phrase and may not include words beyond that, then their result is not in conflict with our own. We did not have any postmodification of the final word in the first functional phrase. Our result could also fit with Martin and Freedman's (2001) *lexical head theory*, which, similar to Schriefers and Teruel (1999), suggests that the head of a phrase must be accessed before phonological processing occurs and thus before speech onset. This would require modification of their assumption that *lexical head* refers to the head of a verb argument phrase (cf. Martin et al., 2004, p. 640), but their data do not contradict our claim, as they tested only simple noun

phrases that had prenominal adjective modification. Thus, the first phrase was the verb argument head phrase and also a single functional unit in the conceptual plan in the utterances they tested.

Given the results from previous work (Martin & Freedman, 2001; Martin et al., 2004; Smith & Wheeldon, 1999), which indicate that the phrasal scope is a scope of lemma access, our preferred interpretation of the data is that a functional phrase defines one scope of grammatical planning, that of lexical access. Moreover, Allum and Wheeldon (2007) provide direct evidence that the scope of lexical retrieval matches the planning scope demonstrated in the experiments above. However, other possibilities need to be considered, and we address them in the following section.

One alternative interpretation is based on the finding made by V. S. Ferreira (1996) that when alternative syntactic forms are available to encode a message, onset is quicker and less error prone. In our comparison of a verb argument phrase consisting of CNP subject with one consisting of a PP subject (Experiment 2), more alternative structures could have been available for the latter, and thus latencies may have been quicker. For example, the following alternative utterances are possible to describe the display for the PP utterances (6): "Akai hana wa inu no ue ni arimasu" (Red flower$_{TOPIC}$ dog above$_{LOC}$ is); "Inu no ue ni akai hana wa arimasu" (Dog above$_{LOC}$ red flower$_{TOPIC}$ is). In the CNP case, the only alternative is not syntactic but the swapping of the order of the two nouns. This interpretation of the effect we found does not seem plausible to us. Participants were informed as to what structure they should use. The alternative structures suggested above never appeared in the experiments reported here, nor do they fit with the aim of the utterance as described to participants. There were only five syntax errors in total in Experiment 2, three for the PP subject and two for the CNP subject. These involved applying one of the structures specified in the experiment instructions to an inappropriate display. Insofar as this error difference may be considered significant, it goes in the opposite direction to what V. S. Ferreira (1996) would predict, which would be fewer errors with the flexible structure utterance. There were 23 dysfluency errors (hesitation, false start, midphrase pause) for the PP condition and 24 for the CNP condition. Our analysis of error rates showed that there was no effect on error rate of phrase type, contrary to what Ferreira would predict (i.e., a greater error rate when there are fewer syntactic options). In addition, because the experimental instructions make our condition closer to Ferreira's order-constraining condition, it should be no easier to produce the PP utterances as only the specified structure was allowed, and therefore no advantage could be taken of any syntactic flexibility that might exist.

Two further possibilities come from a series of experiments investigating agreement error rates and onset latencies for phrases consisting of a noun phrase modified by a prepositional phrase, or conjoined noun phrases (Pearlmutter & Solomon, 2006; Solomon, 2004; Solomon & Pearlmutter, 2004). The authors proposed that a close conceptual relation between lexical items, which they termed *semantic integration*, could result in processing of such lexical items in closer temporal proximity than those that are not semantically integrated. For example, in a phrase such as "The drawing of the flowers," which is closely semantically integrated, activation of the two lexical items "drawing" and "flowers" will largely overlap, whereas in the phrase "The drawing with the flowers,"

which is not closely semantically integrated, the two lexical items will likely be activated more in sequence. Their proposal was supported mainly by a higher occurrence of agreement errors when two lexical items are semantically integrated than when they are not. However, they also found that onset to integrated phrases is marginally faster (Solomon, 2004). Given the nature of our display, it would seem that neither of our utterance patterns is closely semantically integrated and that neither is more closely semantically integrated than the other. Solomon (2004) cited "The shelf above the sink" as a phrase in which the two items are not closely semantically integrated. It seems unlikely that conjunction would make items semantically closer or more distant. Thus, degree of semantic integration, which was found to marginally speed onset (Solomon, 2004), would not be likely to differentially affect onset to our two initial phrase types. In addition, the speeded onset in their case was largely dependent on the contribution of one condition, the unpreferred condition (Solomon, 2004), which is not related to our conditions.

A second possibility is that the longer onset latency to the CNP condition is not a result of the number of lexical items it encompasses and consequent longer advance planning but rather derives from the difficulty in resolving a conflict that occurs when two message elements share a message function, in our case theme, or subject in the case of utterances beginning with a conjoined noun phrase. It is suggested that in such cases, it may be difficult to decide which item to lexicalize, and thus there is a delay (Pearlmutter & Solomon, 2006; Solomon, 2004; Solomon & Pearlmutter, 2004).

The above suggestion results from a second manipulation they carried out to influence the comparative ease of assignment of the first lexical item to the first syntactic slot (Solomon, 2004; Solomon & Pearlmutter, 2006). For the picture displays used in their experiment, a preferred and unpreferred description were established, both of which consisted of a prepositional phrase modifying a noun phrase but each requiring a different word order. For example, for a picture of an apple with a spot on it, the preferred description was "the spot on the apple" and the unpreferred description was "the apple with the spot." A neutral condition, in which lexical order was flexible, was constituted by a conjoined noun phrase, "the apple and the spot" or "the spot and the apple." In this case, flexibility was defined by degree of flexibility in word order. Participants were shown the linking word (e.g., "with," "on," "and") either before or after the picture stimulus appeared, with the aim of biasing them to move from a preferred to an unpreferred pattern in some cases. These cases should be characterized by greater difficulty in assigning the first lexical item to the first syntactic slot due to competition from the preferred item. As evidence of competition, one relevant finding in Pearlmutter and Solomon (2006) is that at negative SOAs the flexible condition, the CNP, was slowest. They took this to be evidence of increased competition between lemmas for the first slot in this phrase type.

Although difficulty resulting from two lemmas competing for the same function assignment might cause a slower onset, a pre-lemma-access effect, difficulty in deciding order of activation for two lexical concepts is an equally likely cause, especially in incremental models. When the linking word is a preposition, it immediately requires a fixed order of activation. This alone might make this condition faster at negative SOAs. As soon as the display appears, the order of activation is decided, whereas in the flexible condition it has to be decided after appearance of the display and without guidance. In contrast, when the linking word was shown at positive SOAs, after the onset of picture display, onset to the conjoined pattern, the flexible pattern, was as fast as to the preferred prepositional phrase pattern for unintegrated items up to an SOA of 1,500 ms, a result that is in direct conflict with ours. The positive SOA data, then, seem to suggest that there is no competition for assignment to function of the two initial words in the conjoined noun phrase condition, as such competition is not a factor in the prepositional phrase phrases and onset to the conjoined noun phrases is as fast as that to the prepositional phrases.

Unfortunately, Pearlmutter and Solomon (2006) made no simple comparison of preferred, unpreferred, and conjoined noun phrases in terms of onset latency, one in which the SOA was not manipulated by the display of a linking word, nor did they provide any data on the frequency with which each alternative structure would be produced to the displays involved. It is thus very difficult to calculate what processing took place even at the minimum positive SOA of 500 ms between picture onset and onset of linking word. There is no baseline.

We feel, therefore, that the latency data from Solomon (2004) and Pearlmutter and Solomon (2006) do not yet provide strong evidence of competition when two words are assigned to the same role or function. In the case of the negative SOA result, the data could as easily be interpreted as a result of difficulty in ordering of activation at the level of lexical concepts when there is no conceptual weighting to indicate the order. Equally important, we do not think that our conjoined noun condition is the same as their flexible condition insofar as our participants knew in what order the conjoined nouns had to be produced, from top to bottom in the visual display, and thus conflict for the first slot in the phrase is far less likely to have been a factor. It was clear to participants what item needed to be lexicalized first or assigned to the first slot. The same is true of Smith and Wheeldon's (1999) experiment. More work is needed, then, to demonstrate that when two words have the same role or function, this leads to competition and thus delay in utterance onset, for example, under conditions in which ordering decisions are not likely a factor.

A third possibility is that an experimental design factor may have caused the effect. Specifically, the CNP phrases were stimulated by two pictures of the same color whereas the PP phrases used one colored and one white picture. This could have made the CNP items less visually distinct and therefore harder to lexicalize. We do not think the pictures in the CNP condition were any less visually distinct than those in the PP phrases. However, we did run a check experiment in which we had participants produce a CNP structure to displays that consisted of two pictures of the same color or a colored and a white picture. The factor that indicated a CNP was required was orientation rather than color. PP phrases and fillers were vertically aligned and used gray and white for colors, whereas CNP phrases were horizontally aligned and used the colors we used in our experiments. The independent variable was color (same or different) for the CNP phrases. We found that there was no difference in onset to the CNP display types. We therefore think that same or different color for the two lexical items was unlikely a factor in our results.

We now look more closely at how the processing scope we define could fit into models of speech production. Does this scope operate at the level of conceptual or grammatical encoding? If the

latter, does it apply to phrase structure building or instead to lexical access, as we have suggested above? Let us first consider the processes of lexical access. Word exchange errors often occur within a clausal scope, and this has been used to argue that all lemmas for at least a clause are activated prior to utterance onset (e.g., Garrett, 1988). However, Smith and Wheeldon (1999) provided conflicting experimental evidence that shows that processing is more thorough for the first verb argument phrase than for subsequent phrases. They claimed that this indicates lemma access for items within the initial verb argument phrase and not those further downstream. However, they also showed that words further downstream within the initial verb argument phrase may be processed to a lesser degree, for example, at the level of lexical concepts. Some such difference is suggested by their finding (Experiment 5) of a bigger effect of preview to a second noun within the first functional phrase (a CNP) than to one in a postnominal modifying relative clause, which is a second functional unit. In our experiments, a similar line of reasoning would suggest that in Japanese PP phrases, only the lemma for the word in the initial subordinate prepositional phrase is accessed.

Such a conclusion, however, requires that a linear ordering process that is a purely grammatical convention affects the activation of elements within the conceptual message level. The prepositional phrase has the function in the conceptual message of modifying or defining the theme. However, the order of modifier phrase and head phrase is grammatical rather than conceptual; it differs according to the grammatical conventions of each language rather than to any conceptual formulation of the message. Thus, as mentioned in the introduction, purely conceptual weighting could not account for the lexical concept associated with the modifier in these sentences being initially more activated than the lexical concept representing the head and therefore initially selected for processing. Because, as we have already argued, such ordering appears to have no conceptual basis, it would imply that a grammatically based ordering process, one that indicates that a modifier comes before the thing it modifies, must be in direct contact with the conceptual message level in order to ensure lemmas become available in the correct order. The grammatical ordering process is working not on the ordering of two activated lemmas, as only one is activated, but on the lexical concepts or thematic roles associated with those lemmas. Thus, if the lemma for the head of the verb argument phrase is not activated, the ordering within the verb argument phrase must be taking place at a pre-lemma-access level.

Although the assumption in Bock and Levelt (1994) is that lemmas for a whole clause are accessed and assigned to grammatical functions stipulated by the verb argument structure, it might be possible even within their model for incremental access of lexical items to be explained by the effect of conceptual weighting on those message elements that are one of the arguments of the verb and thus an element whose prominence in the message could affect the grammatical role to which it is assigned. We have also suggested that because a modifying phrase is not susceptible to such weighting, the only way for Bock and Levelt's (1994) model to order lexical items into subordinate elements within the phrase would be through the creation of a syntactic structure. However, in their model such structure is created after lemma access. Such a model strongly suggests that access to the head would be necessary to create the appropriate syntactic structure prior to linear ordering of the items. It is thus difficult for such a model to account for our

findings if our effect is a result of scope of lexical access, as we suggest.

However, it is possible for models that allow conceptual and grammatical processes to interact without access to lexical items to account for activation of a lemma in an initial modifying phrase before the head of that phrase. In Chang et al.'s (2006) model, lexical concepts are assigned to thematic roles. Role assignment is affected by features of scene perception, such as "most saliently changed," and mediated through the XYZ roles feature (Chang et al., 2006, p. 241) that automatically links such conceptual features to a role defined as X, Y, or Z, each of which typically subsumes a more precise role (e.g., patient, theme, experiencer, or figure in the case of Y roles). Event semantics then give weighting to each of these roles. This weighting affects activation level of abstract conceptual elements such as thematic roles and as a consequence affects sequencing. Thus, if an element assigned to the Y role gets high activation from the event semantics, it is likely to be selected for lexicalization first, and a passive sentence is likely to result. One key point for us is that this ordering takes place prior to lexical access. It thus offers a way to explain how conceptual weighting leads to incremental access of lexical elements in the correct order.

Up to this point, however, we have referred only to thematic roles that serve as arguments of the verb. Although Chang et al. (2006) did not specify how a modifying element such as we used in our experiments would be handled, they did show how elements such as articles or prepositions might be activated in the correct order. It would thus seem possible that when activation spreads from the event semantics representation to the thematic representation, the thematic role could be linked to a modifying element at the thematic level with the function of describing or identifying the theme, just as verb argument roles, such as agent, are linked, for example, to an element such as definiteness to indicate a definite article. This would then allow the sequencing system to favor the modifying phrase over the head phrase, as the sequencing system has learned the grammatical ordering of elements. Thus, greater activation would go to the modifying element, resulting in initial selection of the lexical item associated with it. Chang et al.'s (2006) model has the potential to account for higher activation of a thematic element that is not associated with a lexical item representing the head of a verb argument phrase and thus lexical access of that element prior to lexical access of the head.

Thus, our findings, if indicative of a planning process that controls scope of lexical access, as we claim is likely, strongly favor a model such as that described by Chang et al. (2006), in which an abstract thematic representation of the message interacts directly with an independent grammatical process that can affect weighting of elements within the thematic representation and, through this interaction, affect the ordering of activation and access of lexical items. Conversely, our findings are difficult to instantiate in a model that supports creation of a syntactic hierarchy following a lexical access process driven solely by conceptual weighting, in which linearization takes place after the creation of that hierarchy.

A second possibility to explain our effect is that it occurs post lexical access at the stage of creation of constituent structure. In Bock and Levelt (1994) this is described as "creation of a control hierarchy for phrasal constituents that manages the order of word production and captures dependencies among syntactic functions" (pp. 947–948). This would mean, for example, that from Experi-

ment 2 we could conclude that construction of a more complex constituent structure for the two lemmas in the CNP phrase, along with their assignment to their respective slots, takes longer than construction of a simpler constituent structure and assignment for one in the PP condition. This implies syntactic processing costs for constituent structure creation. That such costs exist has been shown by Smith and Wheeldon (2001). Their experiments indicate that this syntactic process is separate from and incurs costs that are additional to both conceptual or visual processing and lexical access. Further evidence that it is independent of higher level processing was revealed by Wheeldon and Smith (2003), who demonstrated that this process has a much shorter lived persistence effect than that of higher processing levels. Further, Smith and Wheeldon (2001) showed in a syntactic priming experiment that the scope of such a syntactic process is the initial verb argument phrase. However, as in their earlier experiment (Smith & Wheeldon, 1999), there is not a clear distinction between the different levels of phrase, and comparative effects for functional phrase, head phrase, or whole verb argument phrase are not assessed.

It is therefore possible that after accessing the lemmas for the initial verb argument phrase, constituent structure is created for the first functional unit within that phrase, with processing costs reflecting the syntactic complexity of the unit or number of lemmas to be assigned. This possibility cannot be completely ruled out. However, in the PP condition in Japanese, this account would require that a specifically grammatical component, the lemma of the head noun, which dominates the verb argument phrase and its syntactic structure, be accessed but the syntactic structure associated with it not be immediately used. Only the syntactic structure for the lemma associated with the subordinate phrase within the initial verb argument phrase would be created, which seems counterintuitive. Why would the lemma for the head phrase be accessed but creation of associated syntactic structure wait? As mentioned above, given the results from previous work (Martin & Freedman, 2001; Martin et al., 2004; Smith & Wheeldon, 1999), which indicate that the verb argument phrase is a scope of lemma access, we think our first interpretation is much more likely.

Finally, there is the question of how a functional unit is distinguished within the production system. Because the verb argument phrase and the first grammatical phrase within the subject phrase are noun phrases in English (Experiment 4) and in Japanese (Experiment 1), it cannot be that the grammatical category of the phrase per se identifies the functional phrase. We have already indicated that it cannot be identified as the smallest complete phrase or as a phrase representing a verb argument. One possibility is that it is conceptually defined, as suggested by our use of the term *functional unit*. We have argued that our results fit better with a model in which there is a direct link between the conceptual message-level representation and the syntactic and ordering functions within the production system, one that is not mediated though lexical access. We have suggested a development of the Chang et al. (2006) model that could allow such a system to account for our results. It is likely, therefore, that the conceptual process involved in defining what a functional role is in the thematic structure of the sentence could also affect speakers' decisions about what scope of grammatical encoding is preferable. It may be that speakers are ready to initiate utterances when they have lexicalized the smallest functional unit within the conceptual message. Although they need to have an overall conceptual representation of what they want to say, speakers are comfortable with planning grammatical production in minimal whole units of meaning, functional phrases.

References

Allum, P. H., & Wheeldon, L. (2007). *The relationship between scope of lexical access and the thematic representation of an utterance.* Manuscript in preparation.

Berg, T. (1987). The case against accommodation: Evidence from German speech error data. *Journal of Memory and Language, 26,* 277–299.

Bock, J. K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). New York: Academic Press.

Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99,* 150–171.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production: Vol. 1. Speech and talk* (pp. 155–176). London: Academic Press.

Chang, F., Bock, K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition, 90,* 24–49.

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

Costa, A., Navarrete, E., & Alario, F.-X. (2006). Accessing object names when producing complex noun phrases: Implications for models of lexical access. *Cognition, 18,* 3–23.

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review, 98,* 604–614.

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

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

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

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35,* 724–755.

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.

Ford, M., & Holmes, V. M. (1978). Planning units and syntax in sentence production. *Cognition, 6,* 35–53.

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

Garrett, M. F. (1993). Errors and their relevance for models of language production. In G. Blanken, J. Dittmann, H. Grimm, & J. C. Marshall (Eds.), *Linguistic disorders and pathologies* (pp. 72–92). Berlin, Germany: Walter de Gruyter.

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

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

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

Hofland, K., & Johansson, S. (1982). *Word frequency in British and American English.* Harlow, England: Longman.

Holmes, V. M. (1988). Hesitations and sentence planning. *Cognition, 3,* 323–361.

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, 20,* 824–843.

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

Korvorst, M., Roelofs, A., & Levelt, W. J. M. (2006). Incrementality in naming and reading complex numerals: Evidence from eyetracking. *Quarterly Journal of Experimental Psychology, 59,* 296–311.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present day American-English.* Providence, RI: Brown University Press.

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

Levelt, W. J. M., & Maassen, B. (1981). Lexical search and order of mention in sentence production. In W. Klein & W. J. M. Levelt (Eds.), *Crossing the linguistic boundaries* (pp. 221–252). Dordrecht, the Netherlands: Reidel.

Li, C. N., & Thompson, S. A. (1976). Subject and topic: A new typology of language. In C. N. Li (Ed.), *Subject and topic* (pp. 457–489). New York: Academic Press.

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language, 31,* 713–733.

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., Knight, M., Tamborello, F., & Yang, C. (2006). *Noun phrase complexity vs. word retrieval fluency in sentence production.* Manuscript submitted for publication.

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.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology, 25,* 188–230.

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

Meyer, A. S. (1997). Conceptual influences on grammatical planning units. *Language and Cognitive Processes, 12,* 859–864.

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

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

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

Pearlmutter, N. J., & Solomon, E. S. (2006, August). *Semantic integration and the time course of planning complex noun phrases.* Poster presented at the Third International Workshop on Language Production, Chicago.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall sentences. *Journal of Memory and Language, 38,* 265–282.

Schriefers, H. (1993). Syntactic processes in the production of noun phrases. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 841–850.

Schriefers, H., & Teruel, E. (1999). The production of noun phrases: A cross-linguistic comparison of French and German. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st Annual Conference of the Cognitive Science Society* (pp. 637–642). Mahwah, NJ: Erlbaum.

Schriefers, H., & Teruel, E. (2000). Grammatical gender in noun phrase production: The gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1368–1377.

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

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

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

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

Solomon, E. S. (2004). *The pizza with the yummy toppings: Semantic integration in language production.* Unpublished doctoral dissertation, Northeastern University, Boston.

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology, 49,* 1–46.

Van der Meulen, F. (2001). *Moving eyes and naming objects* (MPI Series in Psycholinguistics No. 17). Nijmegen, the Netherlands: Max Planck Institute for Psycholinguistics.

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin, 128,* 442–472.

Vigliocco, G., & Zorzi, M. (1999). Contact points between lexical retrieval and sentence production. *Behavioral and Brain Sciences, 22,* 58–59.

Wheeldon, L. (1989). *Priming of spoken word production.* Unpublished doctoral dissertation, Cambridge University, Cambridge, England.

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

Wheeldon, L., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 44*(A), 723–761.

Wheeldon, L., & Smith, M. C. (2003). Phrase structure priming: A short-lived effect. *Language and Cognitive Processes, 18,* 431–442.

Appendix A

Experimental Triplets for Experiment 1

Set 1

pig	ぶた	cloud	くも	ear	みみ
koala	コアラ	candle	ろうそく	microphone	マイク
elephant	ぞう	glasses	めがね	gorilla	ごりら
cat	ねこ	fan	うちわ	umbrella	かさ
snake	へび	vase	かびん	eggplant	なす
penguin	ペンギン	book	ほん	glass	コップ
camel	らくだ	tent	テント	kangaroo	カンガルー
mouse	ねずみ	bridge	はし	water sprite	かっぱ
shoe	くつ	ant	アリ	tie	ネクタイ
ribbon	リボン	tortoise	かめ	tooth	は
chair	いす	giraffe	きりん	chestnut	くり
desk	つくえ	whale	くじら	kettle	やかん
trousers	ずぼん	frog	かえる	grape	ぶどう
mushroom	きのこ	dragonfly	とんぼ	cross	じゅうじか
mirror	かがみ	dog	いぬ	tree	き
nail	くぎ	lion	ライオン	key	かぎ

Set 2

ear	みみ	koala	コアラ	glasses	めがね
microphone	マイク	shoe	くつ	dog	いぬ
gorilla	ごりら	ribbon	リボン	frog	かえる
umbrella	かさ	snake	へび	tent	テント
eggplant	なす	penguin	ペンギン	fan	うちわ
glass	コップ	camel	らくだ	bridge	はし
kangaroo	カンガルー	chair	いす	lion	ライオン
water sprite	かっぱ	trousers	ずぼん	giraffe	きりん
tie	ネクタイ	nail	くぎ	tortoise	かめ
tooth	は	cat	ねこ	cloud	くも
chestnut	くり	pig	ぶた	book	ほん
kettle	やかん	mouse	ねずみ	vase	かびん
grape	ぶどう	mushroom	きのこ	whale	くじら
cross	じゅうじか	mirror	かがみ	ant	アリ
tree	き	elephant	ぞう	candle	ろうそく
key	かぎ	desk	つくえ	dragonfly	とんぼ

(*Appendixes continue*)

Appendix B

Experimental Pair Sets for Experiment 2

Set 1				Set 2			
pig	ぶた	cloud	くも	flower vase	かびん	pig	ぶた
koala	コアラ	trousers	ずぼん	ribbon	リボン	koala	コアラ
dog	いぬ	flower vase	かびん	tent	テント	dog	いぬ
mouse	ねずみ	ribbon	リボン	fan	うちわ	mouse	ねずみ
whale	くじら	tent	テント	kettle	やかん	whale	くじら
snake	へび	fan	うちわ	nail	くぎ	snake	へび
giraffe	きりん	kettle	やかん	trousers	ずぼん	giraffe	きりん
penguin	ペンギン	nail	くぎ	cloud	くも	penguin	ペンギン
umbrella	かさ	cat	ねこ	ant	アリ	umbrella	かさ
glass	コップ	elephant	ぞう	lion	ライオン	glass	コップ
cross	じゅうじか	ant	アリ	tortoise	かめ	cross	じゅうじか
bridge	はし	lion	ライオン	frog	かえる	bridge	はし
chair	いす	tortoise	かめ	camel	らくだ	chair	いす
glasses	めがね	frog	かえる	dragonfly	とんぼ	glasses	めがね
desk	つくえ	camel	らくだ	cat	ねこ	desk	つくえ
mirror	かがみ	dragonfly	とんぼ	elephant	ぞう	mirror	かがみ

Set 3				Set 4			
pig	ぶた	tent	テント	nail	くぎ	pig	ぶた
koala	コアラ	fan	うちわ	kettle	やかん	koala	コアラ
dog	いぬ	kettle	やかん	cloud	くも	dog	いぬ
mouse	ねずみ	nail	くぎ	trousers	ずぼん	mouse	ねずみ
whale	くじら	trousers	ずぼん	ribbon	リボン	whale	くじら
snake	へび	cloud	くも	flower vase	かびん	snake	へび
giraffe	きりん	ribbon	リボン	fan	うちわ	giraffe	きりん
penguin	ペンギン	flower vase	かびん	tent	テント	penguin	ペンギン
umbrella	かさ	lion	ライオン	camel	らくだ	umbrella	かさ
glass	コップ	camel	らくだ	cat	ねこ	glass	コップ
cross	じゅうじか	frog	かえる	elephant	ぞう	cross	じゅうじか
bridge	はし	tortoise	かめ	dragonfly	とんぼ	bridge	はし
chair	いす	dragonfly	とんぼ	frog	かえる	chair	いす
glasses	めがね	cat	ねこ	tortoise	かめ	glasses	めがね
desk	つくえ	elephant	ぞう	ant	アリ	desk	つくえ
mirror	かがみ	ant	アリ	lion	ライオン	mirror	かがみ