一般社団法人 電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report

The timing of verb selection in English active and passive sentences

Shota MOMMA † Robert SLEVC ‡ and Colin PHILLIPS †

† Department of Linguistics, University of Maryland 1401 Marie Mount Hall, College Park, MD, 20742, USA
‡ Department of Psychology, University of Maryland 1147 Biology/Psychology Building, College Park, MD, 20742, USA

E-mail: † shotam@umd.edu, colin@umd.edu, slevc@umd.edu

Abstract The current study reports the results from an extended picture-word interference task that examines the timing of verb planning in English active and passive utterances. The pattern of semantic interference on verbs suggests that advance planning of verbs occurs selectively before the onset of passive sentences. This pattern of advance verb planning cannot be explained by other factors including the linear position of verbs or limitations in cognitive resources. It is argued that the abstract linguistic relationships between words in a sentence guide the timing of lexical planning above and beyond these factors. This challenges an important premise of strongly incremental models of sentence production: that the production of a word in a sentence needs no advance planning of later coming words.

Keywords Sentence Production, Advance Planning, Incremental production

1. Introduction:

1.1 Incrementality in sentence production

Most language production researchers agree that sentence production is at least to some extent incremental, meaning that a speaker does not have to formulate an entire utterance before starting to speak [1]. This relatively uncontroversial claim is often accompanied by the stronger (but sometimes implicit) claim that the production of a word in a sentence does not require the formulation of the other pieces of the sentence [2-4]. This latter claim, which we term *the independence hypothesis*, implies that the planning of a sentence proceeds in a roughly linear (i.e., left-to-right) fashion such that later coming words are less likely to be included in the initial planning scope than earlier words.

Some experimental evidence suggests that the independence hypothesis is, at least in some circumstances, plausible. For instance, Griffin [2] showed that in an utterance *the A and the B is next to the C*, the codability (how easy it is to choose one name corresponding to an object) of *B* did not affect the speech onset latency. This suggests that even the second noun in a conjoined noun phrase is not planned before the production of the first noun, supporting the strongly incremental, word-by-word view of sentence production. In other words, the scope of advance planning, or the extent to which a speaker plans before speech onset, can be as small as one word.

Although some other studies [4-5] suggest that the scope of advance planning is wider, a recent proposal by Wagner Jescheniak, & Schriefers [6] renders such seemingly conflicting evidence unproblematic for the independence assumption. For instance, one piece of evidence that seemingly conflicts with the independence assumption is Meyer [4]. She showed that a distractor that is semantically related to the second noun in a picture-word interference task with an utterance like *the A is next to the B* delayed speech onset, suggesting the advance planning of B's lemma. This contrasts with Griffin [2], who failed to find any evidence for advance noun planning. However, in a similar paradigm, Wagner et al. [6] demonstrated that the semantic distractor effect on the second noun (the marker of advance planning) disappeared when speech-related task demands were high. This suggests that how much planning a speaker performs is a function of extra-linguistic factors such as available cognitive resources, and thus also suggests that Meyer's [4] evidence (and other evidence pointing to wide scope of advance planning) is not in conflict with Griffin [2] and with the independence assumption.

Given these studies, the independence hypothesis seems plausible. A speaker does not need to plan later coming words in order to articulate the first word of a sentence, and advance planning happens only when a speaker has sufficient available resources. However, one limitation in these (and most other) past studies is that the evidence is almost exclusively concerned with the planning of a noun before the articulation of another noun. This is an unfortunate limitation, because nouns in a sentence usually hold only indirect linguistic relationships with the other nouns in the first place. A stronger test of the

independence assumption would be to examine the planning of a later-coming word that holds a direct linguistic relationship with the to-be-articulated word. One instance of such a strong test involves the advance planning of a verb before the articulation of its argument nouns.

1.2 The timing of verb selection in sentence production

Models of sentence production [7-8] assume that verbs do have a special syntactic and/or semantic role in generating sentences. Most explicitly, Ferreira [8], based on *lexicalized tree-adjoining grammar* [9], argues that the lemma representation of verbs is necessary for completing grammatical processing of the first argument NP. This means that a verb's lexico-syntactic information (or *lemma* [10]) needs to be planned before a speaker starts the articulation of the first noun in a sentence. This view strongly conflicts with the independence assumption, as grammatical processing of the pre-verbal arguments needs verb planning to be completed. This view essentially abandons the independence hypothesis, and thus we consider this a version of a non-independence hypothesis.

However, a series of extended picture-word interference studies by Schriefers, Teruel & Meinhausen [11] suggested that even verbs are not obligatorily processed when they occur in sentence final position in German. They showed that semantic interference from verb-related distractors within German subordinate SV and SOV clauses failed to obtain. This is a strong test of the independence assumption, and it suggests that articulation of the subject does not need the verb to be planned in advance. Thus, the independence hypothesis seems to hold even in the case of a verb and (one of) its arguments. A speaker does not have to plan even the syntactic and semantic core of a sentence, i.e., the verb, before producing its dependents, i.e., the arguments.

It is still arguable, however, that Schriefers et al.'s [11] test may not be strong enough. This is because there are some linguistic reasons to think that the external argument (or underived subject) is not in a true sense an argument of the verb [12]. That is, external arguments are often considered to hold a weaker relationship than internal arguments to verbs. Thus, an even stronger test of the independence assumption would be to examine the advance planning of verbs before the articulation of their *internal* arguments. Momma, Slevc & Phillips [13] conducted such a test based on a modified version of the extended picture word interference paradigm as in Schriefers et al [11], and indeed found that semantic interference from associates of sentence-final verbs in Japanese SV and OV utterance obtained selectively in OV utterance. This suggests that advance verb planning does occur before the articulation of internal arguments but not before that of external arguments, and that it is premature to accept Schriefers et al. [11] as a conclusive demonstration of the independence hypothesis.

2. Current Study

The current study extends Momma et al. [13] in three important respects. First, it is unclear from Momma et al. [13] what triggered the selective advance verb planning in OV utterances. It is possible, for instance, that the object noun and the verb form a syntactic unit, i.e., a verb phrase, and that this shared phrasal membership resulted in simultaneous planning of object noun lemmas and verb lemmas. Alternatively, it is also possible that the accusative case marker on object nouns needs licensing from verbs, or that internal arguments need some syntactic licensing from verbs. The results from Momma et al. are ambiguous between these three possibilities, and it is important to evaluate these alternative explanations in order to make the non-independence hypothesis more specific, and thus more readily testable.

Second, Momma et al. [13] did not match the lexical content between SV and OV sentences. This might have made the processing of one sentence type easier or harder than the other, potentially weakening the effectiveness of the comparison between these sentence types.

Finally, the position of verbs in Momma et al. [13] was linearly close to the sentence onset, i.e., they were the second word of the sentence. Although linear position cannot explain why Momma et al. [13] found evidence of advance verb planning only in OV sentences, it is still possible that linear position conditions advance planning. For instance, it is in principle possible that a speaker plans a verb in advance only when it is linearly close to utterance onset and it is included in the first phrase (VP) of an utterance. The non-independence hypothesis should predict that the advance verb planning effect is not affected by the linear position of the verb, and thus stronger support for the non-independence hypothesis can be obtained if more distant verbs are planned before speech onset.

The current study addresses these three points, using an extended picture word interference task similar to

Schriefers et al., [11] and Momma et al., [13] to examine the timing of verb planning in English active and passive sentence production. In English passive sentences, the first noun phrase does not receive accusative case, nor it is part of the verb phase in surface syntactic structure. Thus, if selective advance verb planning is due to some abstract linguistic dependency between internal arguments and verbs, it is predicted that verbs should be planned before the articulation of subject nouns in passive utterances. On the contrary, if selective advance verb planning is due to surface constituency or to a case dependency, it is predicted that verbs should not be planned before subject nouns in active or passive utterances, replicating Schriefers et al. [11]. Thus, the current experiment is useful for testing and constraining the non-independence hypothesis. Furthermore, the position of the verb in passive utterances in the current study is farther away in linear distance from the sentence onset. Thus this provides a stronger test than Momma et al. [13] of the independence assumption's implication that later-coming words should be less likely to be planned early. Also, it is important to note that Wagner et al.'s [6] explanation that scope of planning is extended with more available cognitive resources predicts that advance verb planning in passive sentences is less likely to occur than advance verb planning in active sentences. This is because passive utterances are by every metric (e.g., utterance complexity, utterance length, structural frequency, etc.) more cognitively demanding, and because the position of verb is further away from sentence onset in passive than in active sentences. Thus, Wagner et al.'s [6] explanation should predict that advance verb planning should be less likely to occur in passive utterances than in active utterances.

3. Method

3.1. Participants

A total of 38 undergraduate students at University of Maryland participated for course credit. The data from 4 participants were excluded due to excessive error rate (>40%).

3.2. Materials

Twenty-six transitive action pictures were used as picture stimuli. In these depicted actions, there were 13 unique identifiable characters that were the participants of those actions, e.g., doctor, boxer, swimmer, etc. Each character appeared four times total, twice as an agent and twice as a patient, each time in a different action context. This means that, as a set, the noun preceding the verbs in active and passive utterances were exactly matched. This avoids any potential confounds of lexical frequency, codability, visual salience, etc. of the nouns preceding verbs on advance verb planning pattern, unlike in Momma et al. [13]. The agent and patient entities appeared on the left side and right side of the picture equally often. This prevented participants from relying on positional information to identify the agent and the patient in an event. For each action picture, semantically related distractor verbs were chosen according to native English speakers' intuitive judgments. 10 of the semantic distractors were from the verbs used as a target in the other pictures. Related distractor words were used as the unrelated distractors for other pictures, such that the set of distractor words was identical in the related and unrelated conditions. Examples of the picture stimuli are presented in Figure 1 below.

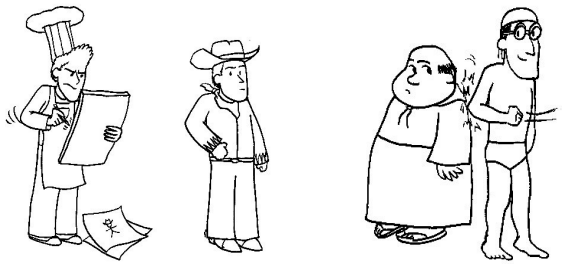

Figure 1: Two example picture stimuli used in this experiment. The left picture elicited *the chef is sketching the cowboy* in the active condition and *the cowboy is being sketched by the chef* in the passive condition. The right picture elicited *the swimmer is nudging the monk* in the active conditions and *the monk is being nudged by the swimmer* in the passive conditions.

3.3. Procedure

Participants were tested in a sound-attenuated room with an experimenter present. Before the main part of the experiment, they first performed a picture naming task where they named all characters used in the current experiment with a different set of pictures (following a familiarization phrase with written target nouns on top). They also performed an action naming task (following a familiarization phrase with written target verbs on top) where they named all actions with the verbs used in the

current experiment, with the same set of pictures and distractors used in the main experiment (with the same presentation parameters and the same number of repetitions). Following these tasks they then participated in the main experimental task. As they had already practiced naming the characters and actions used in the current experiment in the preceding two tasks, there were no familiarization or practice trials. In the experimental session, participants were instructed to name the action depicted on the picture in sentential form, as if the action were happening right now, i.e., in progressive form. In addition, for the 15 participants who were randomly assigned to the passive utterance condition, they were instructed to mention first the character that corresponded to the 'target' of the action in the sentence. For the other 17 participants who were in the active condition, they were instructed to mention first the character that corresponded to the 'doer' of the action.

On each trial, a fixation-cross appeared at the center of the screen for 750 ms, and then a distractor word was presented visually at the center of the screen. 150 milliseconds after the onset of distractor presentation, a picture stimulus was presented on the screen simultaneously with a brief click sound that served as a picture onset marker. The visual distractor disappeared after 500 milliseconds. The picture stimulus remained on the screen for 1500 milliseconds. A 3600 millisecond black screen separated the trials. The ordering of pictures within a block was pseudo-randomized for each participant, and the order of presentation of a picture with different types of distractors within a block was counterbalanced across participants.

For each trial, the onset latency of noun production was manually measured, using Praat [14]. The measurer was blind to the conditions, although he could identify the target utterances. Any trials with audible hesitation (e.g., fillers), or with non-target words were excluded. Disfluency was tolerated, as long as the utterance finished before the trial ended. In addition, for each participant, trials with response times (RT) more than two standard deviations away from that individual's mean RT were excluded. In total, 16.7% of data points were excluded. Response times were log-transformed for statistical analyses.

4. Results

Using R [15] and *lme4* [16], a mixed effects model with maximal random effects structure was constructed (without random slopes for interaction terms, which caused convergence failure), followed by model simplification based on maximum likelihood ratio tests. As a result, the by-items random slope of Utterance Type was significant ($p > 0.01$). Also, the by-subjects random slope was marginally significant ($p = 0.09$). Thus, we report the model with by-subject intercept, by-item intercept, by-subject slope of Relatedness, and by-item slope of Utterance Type. All the p-values reported here were derived via MCMC. The interaction between Relatedness and Utterance Type was significant ($b = 0.05$, $SE = 0.02$, $|t| = 2.04$, $p < 0.05$). The main effect of Utterance Type was also significant ($b = 0.14$, $SE = 0.06$, $|t| = 2.28$, $p < 0.05$). The main effect of Relatedness was not significant ($p > 0.9$).

A planned comparison between the related and unrelated conditions in active and passive utterances was conducted. This analysis revealed that related verb distractors delayed utterance onset in passive utterances ($b = 0.05$, $SE = 0.02$, $z = 3.04$, $p > 0.01$) but not in active utterances ($b = 0.005$, $SE = 0.02$, $|z| = 0.33$, $p > 0.9$).

The results are summarized in Table 1 below. Semantic interference effects for each utterance type are plotted in Figure 2 below.

	Relatedness	
Utterance type	Related	Unrelated
Active	1162 [178]	1169 [151]
Passive	1344 [228]	1264[178]

Table 1: Mean RT by condition calculated over subject means [sd].

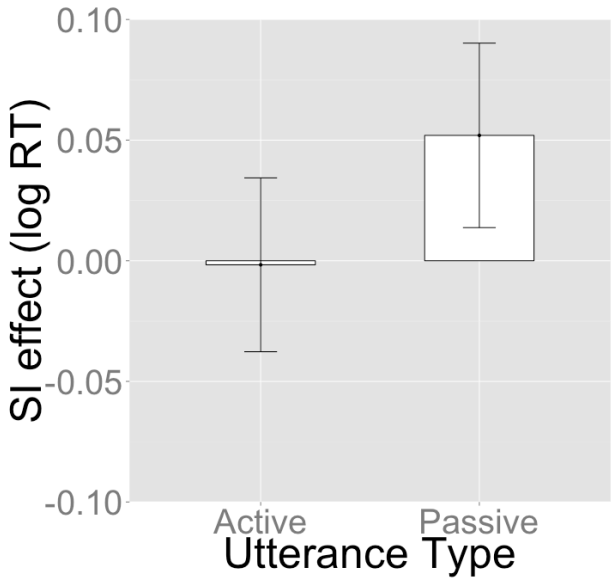

Figure 2: Semantic interference (SI) effect from verb

distractors in active and passive utterances estimated from the statistical model. The error bars represent 95% confidence intervals.

5. Discussion & Conclusion

The current experiment was designed to test a critical claim of the Independence Hypothesis that a word in a sentence can be produced without planning later-coming words in advance. The result does not support the Independence Hypothesis, and indeed shows the opposite pattern of what that hypothesis predicts. Instead, the result supports a version of the non-independence hypothesis, specifically the claim that the abstract linguistic relationship between internal arguments and verbs selectively triggers advance verb planning. A robust semantic interference effect on verbs was obtained in passive utterances and not in active utterances. This pattern corroborates the results by Momma et al. [13], and strongly contrasts with the prediction of the independence assumption, as the factors that are considered relevant for the advance planning under the independence hypothesis (i.e., linear distance and available cognitive resources) make the opposite prediction. The timing of verb retrieval is instead predicted by the presence/absence of a specific linguistic relation (namely, the relation between internal arguments and verbs). Thus, the current results support the non-independence hypothesis, challenging an important premise of strongly incremental models of sentence production.

References

[1] F. Ferreira and B. Swets, How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums, Journal of Memory and Language, vol. 46, no.1, pp. 57-84, 2002.

[2] Z. Griffin, Gaze durations during speech reflect word selection and phonological encoding, Cognition, vol. 82, B1-B14, 2001.

[3] G. Kempen, and E. Hoenkamp, An incremental procedural grammar for sentence formulation. Cognitive science, vol. 11, no. 2, pp. 201-258, 1987.

[4] A. Meyer, Lexical access in phrase and sentence production: Results from picture–word interference experiments. Journal of Memory and Language, vol. 35, no. 4, pp. 477-496, 1996.

[5] M. Smith and L. Wheeldon, High level processing scope in spoken sentence production. Cognition, vol. 73, no. 3, pp. 205-246, 1999.

[6] V. Wagner, J. Jescheniak, J, and H. Schriefers, On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. Journal of Experimental Psychology: Learning, Memory, and Cognition, vol. 36, no. 2, pp. 323-340, 2010.

[7] K. Bock and W. Levelt, Language production: Grammatical encoding, in Handbook of Psycholinguistics, ed. M. Gernsbacher, pp. 945-984, Academic Press, London, 1994.

[8] F. Ferreira. Syntax in language production: An approach using tree-adjoining grammars. Aspects of language production, L. Wheeldon, ed., pp. 291-330, Psychology Press, London, 2000.

[9] A. Joshi, An introduction to tree adjoining grammars, Mathematics of language, vol. 1, pp. 87-115, 1987.

[10] G. Kempen and P. Huijbers, The lexicalization process in sentence production and naming: Indirect election of words. Cognition, vol. 14, pp. 185-209. 1983.

[11] H. Schriefers, E. Teruel, & R. Meinshausen, Producing simple sentences: Results from picture–word interference experiments. Journal of Memory and Language, vol. 39, no. 4, pp. 609-632, 1998.

[12] A. Kratzer, Severing the external argument from its verb. In Phrase structure and the lexicon, pp. 109-137, Springer, Netherlands, 1996

[13] S. Momma, R. Slevc, & C. Phillips, “Advance selection of verbs in head-final language production, ” Poster at the 26th annual CUNY Conference on Human Sentence Processing, Columbia, SC, March. 2013.

[14] P. Boersma, Praat, a system for doing phonetics by computer. Glot International vol. 5, pp. 341-345, 2002

[15] R Core Team, R: A language and environment for statistical computing.

[16] Foundation for Statistical Computing, Vienna, Austria.

[17] D. Bates, M. Maechler, and B. Bolker, lme4: Linear mixed-effects models using S4 classes. R package version version 1.0-6, 2014