Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/15/$12.00 http://dx.doi.org/10.1037/xlm0000195

RESEARCH REPORT

The Timing of Verb Selection in Japanese Sentence Production

Shota Momma, L. Robert Slevc, and Colin Phillips
University of Maryland

Many influential models of sentence production (e.g., Bock & Levelt, 1994; Kempen & Hoenkamp, 1987; Levelt, 1989) emphasize the central role of verbs in structural encoding, and thus predict that verbs should be selected early in sentence formulation, possibly even before the phonological encoding of the first constituent (Ferreira, 2000). However, the most direct experimental test of this hypothesis (Schriefers, Teruel, & Meinshausen, 1998) found no evidence for advance verb selection in verb-final (subject-verb and subject-object-verb) utterances in German. The current study, based on a multiword picture-word interference task (Meyer, 1996; Schriefers et al., 1998), demonstrates that in Japanese, a strongly verb-final language, verbs are indeed planned in advance, but selectively before object noun articulation and not before subject noun articulation. This contrasting pattern of advance verb selection may reconcile the motivation for advance verb selection in structural encoding while explaining the previous failures to demonstrate it. Potential mechanisms that might underlie this contrasting pattern of advance verb selection are discussed.

Keywords: sentence production, advance planning, picture-word interference, Japanese

Uttering a sentence requires that a single message be converted into a sequence of words. There is widespread agreement that this conversion occurs incrementally, meaning that speaking can begin before the entire utterance is planned. A strong version of the incrementality hypothesis (e.g., Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; Griffin, 2001; Iwasaki, 2011; Schriefers, Teruel, & Meinshausen, 1998) posits that the sentence plan is, most of the time, developed in the same order that the words are uttered (i.e., the first word is planned first, the second word second, and so on). This strong version of the incrementality hypothesis holds that the sentence plan is developed on a just in time basis, such that words are planned right before they are uttered. This approach might be beneficial for reducing memory demands and avoiding interference between multiple activated lexical items, but it requires that utterance plans be flexibly adjusted to the word order demands of a language, so as to avoid "look-ahead" in planning. The best tests of this strong incrementality arise in situations where words that play a central role in the organization of a sentence are forced to appear late in the utterance, because of the word order constraints of a language. Here we examine such a case, focusing on evidence for look-ahead effects in Japanese, a language whose word-order constraints force verbs to appear in utterance-final position.

Shota Momma, Department of Linguistics, University of Maryland; L. Robert Slevc, Department of Psychology, University of Maryland; Colin Phillips, Department of Linguistics, University of Maryland.

We thank Hiromu Sakai and his lab members, especially Takuya Kubo for his help in data collection and Manami Sato for valuable comments. This work was supported in part by NSF BCS 0848554 awarded to Colin Phillips.

Correspondence concerning this article should be addressed to Shota Momma, 1401 Marie Mount Hall, University of Maryland, College Park, MD 20782. E-mail: shotam@umd.edu

The Timing of Verb Selection in Sentence Production

Linguistic and psycholinguistic evidence suggests that verbs are critical for structural processing. This points to the early encoding of verbs in utterance formulation. Accordingly, many influential models of sentence production (e.g., Bock & Levelt, 1994; Ferreira, 2000; Kempen & Hoenkamp, 1987; Levelt, 1989) adopt the view that the verb's syntactic representation (i.e., lemma; Kempen & Huijbers, 1983) guides structural processes, and thus they predict that a verb's lemma is selected before the relevant structural processes are performed. Most explicitly, Ferreira (2000) argued that selection of a verb's lemma (or the head of any phrase) must be performed before phonological encoding of the first phrase of a sentence is finalized.

Despite the emphasis on verbs' early encoding in models of sentence production, the experimental evidence from tests of this issue is equivocal at best. Some suggestive evidence for advance verb selection comes from Kempen and Huijbers (1983), who found that changing the target verbs between experimental blocks delayed speech onset for both verb-subject and subject-verb (SV) sentences in Dutch, though with a greater delay for verb-subject than SV sentences. This suggests that either verbs' conceptual representation, their lexical representation, or both are at least partially planned in advance, although the results are ambiguous between these three interpretations. In addition, Schnur, Costa and Caramazza (2006) and Schnur (2011), using the multiword picture-word interference task (Meyer, 1996), found that distractors that are phonologically related to the target verb in SV and subject-verb-object (SVO) utterances facilitated speech onset, suggesting that even the phonological representation of a verb is encoded in advance. However, this evidence is hard to interpret because facilitative effects may alter normal planning in favor of advance planning of verbs, because of potential feedback activation from the phonological level to higher levels, and because these findings conflict with existing evidence that suggests a narrow scope of planning at the phonological level (e.g., Griffin, 2001; Meyer, 1996; Meyer, Sleiderink, & Levelt, 1998; Wheeldon & Lahiri, 1997). Nevertheless, the studies by Schnur and colleagues are broadly consistent with models that posit advance selection of verbs.

On the other hand, the most direct experimental test of the issue failed to provide evidence for advance verb selection. Schriefers and colleagues (1998) used an extended version of the picture-word interference paradigm in German. In five experiments, participants were presented with action pictures with distractor words. Participants' task was to complete a lead-in fragment by describing a picture as quickly as possible using either a verb-initial (e.g., *laugh—the girl*) or verb-final (e.g., *the girl—laugh*) clause, while ignoring distractors (e.g., *cry*). Schriefers and colleagues reasoned that if verb selection occurs prior to utterance onset then the semantic interference (SI) effect, a relative delay of utterance onset due to semantic relatedness between target and distractor verbs, should obtain even in verb-final utterances. However, distractors that were semantically related to the target verbs elicited SI effects only in verb-initial utterances (although see Hwang & Kaiser, 2014 for evidence from English to the contrary).[1] These findings led some researchers to conclude that verb selection is not necessarily performed in advance (Allum & Wheeldon, 2007; Iwasaki, 2011). This article attempts to reconcile the theoretical motivation for advance verb selection and the previous empirical failure to demonstrate it by examining the timing of verbs' lemma selection at two different points in sentences: before object and before subject noun articulation.

Linguistic Contrasts Between Subjects and Objects

Linguistic analyses suggest that objects are more strongly dependent on verbs than subjects in many respects. First, it has sometimes been argued that a verb and its internal arguments constitute the verb's argument structure, whereas external arguments are not part of such a structure (cf. Marantz, 1984; Kratzer, 1996, 2002). This claim is derived from the observation that the choice of the object, but not the subject, has a significant impact on the meaning of the verb. Second, verbs do not select subjects in the same sense that they select objects—a subject is obligatorily present regardless of the properties of the verb (Chomsky, 1981; see Lasnik, 2003, for a discussion), whereas the presence/absence of an object depends on the subcategorization property of the verb (e.g., Haegeman, 1991). Third, subjects and objects are considered to receive case from different sources. The case-assigner of object nouns is generally considered to be the lexical head V(erb), whereas that of subject nouns is considered to be the functional head I(nflection) in nominative-accusative languages (Chomsky, 1981). Finally, objects have a closer constituency relationship with the verb. In a transitive sentence, the verb and the object noun phrase together form a verb phrase, whereas a subject noun phrase and a verb do not by themselves form a syntactic constituent under most accounts (e.g., government and binding [GB] theory, Chomsky, 1981; head-driven phrase structure grammar [HPSG], Pollard & Sag, 1994; lexical functional grammar [LFG], Bresnan, 2001). These linguistic analyses all suggest that object nouns are more closely associated with verbs than subject nouns, both syntactically and semantically.

The Current Study

Given the difference between subject and object nouns in terms of their dependency on verbs, it is possible that verb selection might be required before object articulation, but not before subject articulation. To test this hypothesis we adapted the extended picture-word interference paradigm (Meyer, 1996; Schriefers et al., 1998) to Japanese, with some modifications. We exploited two properties of Japanese: strict head-finality and liberal argument dropping. Specifically, to probe the timing of verb selection both before subject and object articulation based on speech onset latency, one needs to be able to naturally elicit two types of sentences: one starting with an object noun phrase, another starting with a subject noun phrase, preferably in canonical word order. This condition can naturally be met with Japanese. Japanese allows complete sentences consisting either of SV or object-verb (OV) sequences. The OV structure allows us to test whether verbs are selected prior to object noun phrase articulation. SV and OV utterances are both naturally producible grammatical sentences in Japanese, and they can be closely matched in terms of the length of their initial noun phrase. Therefore, they are well suited for comparing the status of advance verb planning before subjects and objects are uttered.

The three experiments reported here used a similar picture-word interference design and are closely related to each other. Experiment 3 was a phrase-production task that tested the key question of

[1] Hwang and Kaiser (2014) showed a SI effect from the verb in English SVO sentences, suggesting that advance verb selection occurs before subject articulation. This pattern contrasts with the results from Schriefers et al. (1998). However, Huang and Kaiser's choice of unrelated distractors was problematic in multiple respects, making it hard to interpret the results. First, the related and unrelated distractors were not matched in terms of number of repetitions throughout the experiment. Each of the related distractors was presented twice, whereas the unrelated distractors were presented only once, leading to a potential independent advantage for related distractors. Second, the related and unrelated distractors were not matched in terms of lexical frequency. The mean corpus log frequency of the related distractors were 7.74 (SEM = 0.19) whereas that of the unrelated distractors were 7.11 (SEM = 0.24) and this difference was marginally significant in two-tailed t test that does not assume equal variance, even though the number of items was very small (four for the related condition and eight for the unrelated condition). Finally, Hwang and Kaiser used adjectives as unrelated distractors. They argued that same category distractors (i.e., verb distractors) might cause interference and thus they should be avoided, but this is exactly the reason why the same grammatical category distractors should be used for related and unrelated condition alike. Otherwise, the interference effect cannot be reliably identified as a SI effect. It cannot be identified as grammatical category mismatch effect either because adjectives and verbs have very different distributional relations to nouns. Therefore, we consider it premature to take the verb interference effect in Hwang & Kaiser's English experiment as evidence for advance verb planning before subject nouns in English. Indeed, we failed to replicate Hwang and Kaiser's SI effect using balanced related versus unrelated verb distractor in an English SVO sentence production task, with a more powerful design (# of participants = 36; # of items = 24), even though the distractors were shown to be effective in a single verb naming context (Momma, Slevc & Phillips, 2015).

the current study: whether verbs are planned before object articulation but not before subject articulation. Experiment 1 was a single-word production task eliciting only verbs, to verify the effectiveness of the picture-distractor pairs used in Experiment 3. Example stimuli used in Experiments 1 and 3 are illustrated in Figures 1 and 2 below. Experiment 2 was also a single-word production task eliciting only nouns, to verify that any interference effects observed in Experiment 3 were reflections of verb planning rather than noun planning. These three experiments in concert are necessary to attribute the presence/absence of a verb-based SI effect to the presence/absence of advance selection of verb lemmas.

Experiment 1

Participants

Twenty-four students from Hiroshima University in Japan participated in all three experiments in exchange for 500 yen.

Materials

Twenty-four action pictures were chosen from the UCSD International Picture Naming Project (IPNP) database (Szekely et al., 2004): half corresponding to transitive verbs, and the other half corresponding to intransitive verbs. For each picture two types of distractors were chosen, one semantically related to the target verb and the other unrelated to the target verb. The related distractors were chosen based on two native Japanese speakers' intuitive judgments, and they involved synonym, antonym, or cohyponym relations. Importantly, the related distractor words were all used as unrelated distractors for other pictures, with the effect that the set of distractor words was identical in the related and unrelated conditions. The unrelated distractors were also matched to the target word in terms of transitivity. This ensured that uncontrolled parameters of the distractor words such as frequency, length and orthographic complexity could not differentially affect speech onset latencies in the related and unrelated conditions. Neither the related or unrelated distractors had a systematic phonological relationship to the target words.

Figure 1. Sample picture for transitive verbs and object-verb sentences with the semantically related distractor word superimposed. Target utterance (in Experiment 3): Neko-o (the cat-accusative) naderu (pet). Distractor verb: sasuru (rub). See the online article for the color version of this figure.

Figure 2. Sample picture for intransitive verbs and subject-verb sentences with the semantically related distractor word superimposed. Target utterance (in Experiment 3): inu-ga (the dog-nominative) hoeru (howl). Distractor verb: naku (cry). See the online article for the color version of this figure.

Procedure

Participants were tested in a sound-attenuated, dimly lit room with an experimenter present. In a familiarization phase they first saw each picture and produced the associated target verb, and then practiced the action-naming task using pictures and distractors that were not included in the experimental set. In the experimental session for Experiment 1, participants were instructed to produce one word that describes the action depicted by each picture as soon and as accurately as possible. On each trial a fixation-cross appeared at the center of the screen for 500 ms with a brief click sound (used for calculating speech onset), and then an action picture appeared simultaneously with a written distractor word. The distractor word disappeared after 300 ms, whereas the picture remained for 1,500 ms. A 3,000 ms blank screen separated the trials. Participants saw each picture twice, in the related and unrelated conditions, in different blocks. The ordering of pictures within a block was randomized for each participant, and the presentation of a picture with the related distractor in the first versus second block was counterbalanced across participants.

Analysis

For each trial, speech onset latency was manually measured, specifically, the interval between the click sound and the onset of speech minus 500 ms, using Praat (Boersma & Weenink, 2014). The measurer was blind to the conditions, although he could identify the target utterances. Any trials with disfluencies, audible nonspeech noise before utterance onset, or speech onset of more than 2000 ms were excluded. In addition, for each participant, trials with response times more than 2 *SD*s away from that individual's mean reaction time (RT) were excluded. In total, 12.7% of the data points were excluded (7.3% due to errors and 5.4% due to trimming).[2] Both mean and dispersion (standard deviation) of each condition was analyzed. The analysis of dispersion was conducted here because (a) the mean analysis of Experiment 3 reported below does not yield unequivocal results and this analysis can serve as a point of comparison for the same analysis that was conducted for Experiment 3 and (b) there are good reasons to think that the SI effect, or incongruity effects in any Stroop-like task, might not be best captured by a difference in the sample mean (Heathcote, Popiel & Mewhort, 1991; Scaltritti, Navarrete & Peressotti, 2014).[3]

Results

Mean RTs (of participant averages) in milliseconds as a function of verb type and relatedness in Experiment 1 is shown in Table 1.

A mixed-effects model with maximal random effects structure was constructed, followed by model simplification based on a maximum likelihood ratio test, with the alpha level of 0.1.[4] As a result, the random slope of any factor by items or by subjects was not significant (all ps > 0.15). Also, we tried the forward selection of the model and the resulting model did not change. Thus, we report models with by-subjects and by-items intercepts. The R code corresponding to the LMER model building can be found in Appendix A. This analysis revealed an effect of relatedness ($\beta = 31.03$, $SE = 10.93$, $|t| = 2.84$, $p < .01$). In both models, neither the effect of transitivity nor the interaction between relatedness and transitivity was significant (all ps > 0.85).

To make a comparison between this experiment and Experiment 3, we conducted planned comparisons with Bonferroni correction between related conditions and unrelated conditions for transitive and intransitive verbs. Both transitive ($\beta = 33.55$, $SE = 10.96$, $z = 3.06$, adjusted $p < .01$) and intransitive verbs ($\beta = 28.26$, $SE = 10.77$, $z = 2.62$, adjusted $p < .01$) independently showed a reliable SI effect.

A repeated-measures analysis of variance analysis in which each cell represented a participant's (for F1) or an item's (for F2) standard deviation (see Heathcote et al., 1991 for an analogous method of analysis in the Stroop task) revealed that the standard deviations in the related and unrelated conditions were significantly different, $F1(23) = 21.36$, $p < .01$; $F2(23) = 7.71$, $p = .01$, without any interaction with verb type ($p > .5$). This suggests that the semantic relatedness of the distractor has an effect on the dispersion of the RT distribution. This SI effect on dispersion is visualized in Figures 3–4 by means of vincentile plots (e.g., Vincent, 1912; Staub, 2010), using only five bins due to the relatively small number of trials per condition per participant. These plots suggest that the SI effect is only reliably identified in the right tail of the distribution (i.e., for slow responses), which is normally reflected by the τ parameter of an ex-Gaussian distribution, and has the following relation to the sample standard deviation: $S = \sqrt{\sigma^2+\tau^2}$. We however cannot identify whether the significant effect of relatedness resides in σ or τ from the current experiment, and it is possible that the SI effect might reside in both parameters.

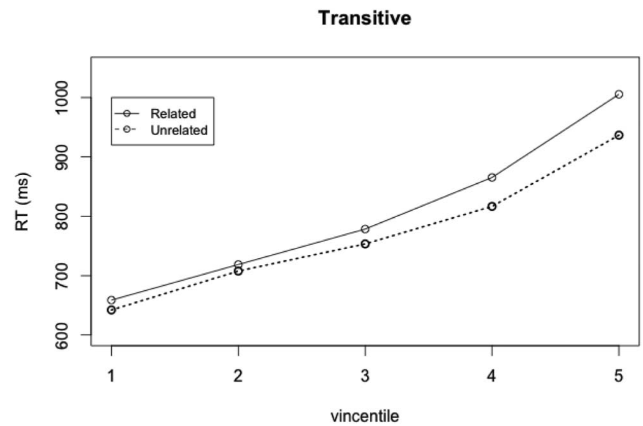

Figure 3. Vincentile plot for the transitive condition in Experiment 1, with 20% percentile increments; that is, Vincentile 1 corresponds to the fastest 20% of trials per participant, Vincentile 2 corresponds to the next fastest 20% of trials per participant, and so on. RT = reaction time.

Discussion

The aim of Experiment 1 was to test whether the set of verb distractors would reliably elicit a SI effect. The results clearly show that the verb distractors were indeed effective in eliciting an SI effect regardless of transitivity, both in mean and standard deviation measures (for a discussion of how these measures relate to the ex-Gaussian parameters, see Heathcote et al., 1991). Hence, the same verb-distractor pairs were suitable for testing sentence production in Experiment 3.

Table 1
Mean (SD) Reaction Times Based on Participant Means in Milliseconds as a Function of Verb Type and Relatedness in Experiment 1

	Relatedness	
Verb type	Related	Unrelated
Intransitive	786 (89)	760 (77)
Transitive	794 (84)	761 (82)

[2] Analysis without trimming yielded the same results.

[3] We have not conducted a full ex-Gaussian distributional analysis because we did not have enough trials per condition to obtain a stable estimate of the ex-Gaussian parameters. Standard deviation can capture a simple measure of the distributional parameter that is not at all (σ) or not strongly (τ) reflected by sample mean.

[4] Barr, Levy, Scheepers, and Tily. (2013) argued against this data-driven approach for selecting random effect structures in confirmatory hypothesis testing. For this experiment, including the maximal random effects structure following Barr et al. (2013) did not change the statistical pattern (but see Experiment 3 for diverging results and discussion).

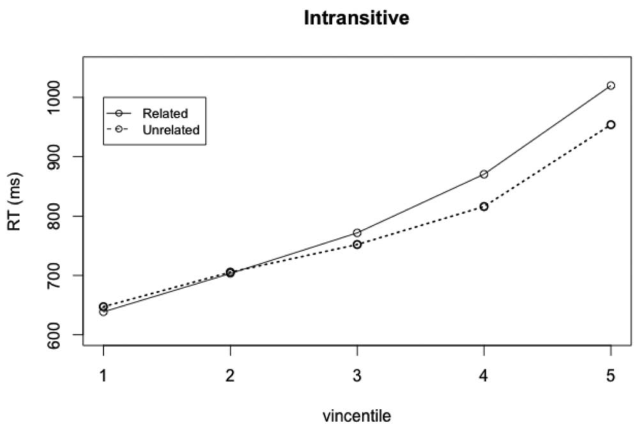

Figure 4. Vincentile plot for the intransitive condition in Experiment 1, with 20% percentile increments; that is, Vincentile 1 corresponds to the fastest 20% of trials per participant, Vincentile 2 corresponds to the next fastest 20% of trials per participant, and so on. RT = reaction time.

Experiment 2

Participants

The same set of participants as in Experiment 1 participated in Experiment 2, except that one participant was excluded due to recording failure. The order of Experiments 1 and 2 was counterbalanced across participants.

Materials

Twenty-four new pictures from the IPNP database were chosen to elicit the bare target nouns that were used also in Experiment 3, that is, without associated actions. No actions were depicted in the pictures, unlike in Experiment 1. The distractor words, as well as the pairing between target nouns and distractor words, were kept constant as in Experiments 1 and 3.

Procedure and Analysis

The procedure was identical to Experiment 1, except that participants were instructed to describe the object depicted in the pictures in one word that is, a noun. The same R code was used as in Experiment 1, except that the factor transitivity is now changed to noun type.

Results

Speech onset latency did not significantly differ across conditions. No effect of noun type, relatedness, or interaction was found ($ps > 0.35$). Mean RTs (of participant averages) in milliseconds as a function of noun type and relatedness in Experiment 2 are shown in Table 2 below. A standard deviation analysis as in Experiment 1 revealed no effect of noun type, relatedness, or interaction between them (all $ps > 0.15$).

Discussion

The results of Experiment 2 suggest that there was no systematic relationship between the target nouns and distractor verbs that affected utterance onset latency. This is important for ensuring that the nouns that were used in Experiment 3 have no accidental relationship to the distractor verb in a way that could have altered the pattern of SI effects.

Experiment 3

Participants

The same set of participants participated in Experiment 3 after completing Experiments 1 and 2, following a short break.

Materials

The same set of picture-distractor pairs was used with identical presentation parameters as in Experiment 1. The target nouns preceding the target verbs were matched in mean length (number of moras: 2.33 for the transitive condition and 2.5 for the intransitive condition). The transitive condition targeted 11 inanimate and 1 animate nouns, and the intransitive condition targeted eight inanimate and four animate nouns. As participants saw the same set of pictures and distractors in Experiment 1, they had seen each picture four times by the end of Experiment 3.[5]

Procedure and Analysis

The same procedure was followed as in Experiment 1 and 2, except that participants were instructed to describe the action depicted in the pictures in sentential forms using two words, that is, a noun, inflected with a case particle, followed by a verb. This elicited an SV sentence for the intransitive action pictures, and an OV sentence for the transitive action pictures. The same analysis procedure was adopted as in Experiment 1 and 2. In total, 11.2% of the trials were excluded (6.3% due to errors, and 4.9% due to trimming).

Results

Mean RTs (of participant averages) in milliseconds as a function of sentence type and relatedness in Experiment 3 is shown in Table 3 below.

The same mixed effects model analysis procedure was applied as in Experiment 1. After model simplification based on maximum likelihood ratio tests, the model with both by-subjects and by-items intercepts revealed a marginally significant interaction ($\beta = 25.20$, $SE = 13.08$, $|t| = 1.93$, $p = .054$).[6] The same R code was used as in Experiment 1, except that the factor Transitivity is now changed to Sentence Type.

There were no main effects of relatedness or sentence type ($ps > 0.7$). As in Experiment 1 we conducted planned comparisons with Bonferroni correction between related and unrelated conditions for OV and SV sentences. These analyses showed that verb-related distractors significantly delayed sentence onsets in

[5] This number of repetitions is lower than in many other PWI studies, which often involve 10 or more repetitions per picture. (e.g., Schriefers et al., 1998; Schriefers, Meyer, & Levelt, 1990)

[6] A model with maximal random effect structure (Barr et al., 2013) yielded nonsignificant interaction ($p = .14$). However, see the standard deviation analysis and vincentile analysis below.

Table 2
Mean (SD) Reaction Times Based on Participant Means in Milliseconds as a Function of Noun Type and Relatedness in Experiment 2

	Relatedness	
Noun type	Related	Unrelated
Subj. Nouns	652 (77)	660 (79)
Obj. Nouns	642 (81)	647 (81)

OV sentences ($\beta = 26.53$, $SE = 9.11$, $z = 2.91$, adj. $p < .01$), but not in the SV condition ($p > .95$).

Also, a post hoc comparison of noun frequencies between OV and SV conditions was conducted to ensure that the differential effect of relatedness on OV and SV conditions was not attributable to the properties of the nouns. Based on the Tsukuba Web Corpus (http://corpus.tsukuba.ac.jp/), the mean log frequencies of the nouns did not differ between the transitive (M = 10.80; SEM = 0.39) and intransitive (M = 9.63; SEM = 0.58) conditions, although there was a trend toward a difference in a two-tailed t test ($p = .11$). A further test using individual items found no correlation between noun frequency and the amplitude of the effect of semantic relatedness ($r = .07$, $p > .7$).

As in the verb-only experiment (Experiment 1), we also analyzed the difference in the standard deviation of the related versus unrelated conditions. This analysis can be thought of as a coarse measure of the dispersion of an ex-Gaussian distribution that reflects both σ and τ parameters. This analysis revealed a fully significant interaction between relatedness and sentence type on standard deviations (i.e., dispersion, $F1[1, 23] = 6.07$, $p = .02$; $F2[1, 22] = 11.41$, $p < .01$), with no significant main effects (ps >0.13). Planned comparisons revealed that in the OV conditions, the dispersion was larger in the related condition ($M = 112$; $SD = 60$)[7] than in the unrelated condition ($M = 90$, $SD = 27$; $t1(23) = 2.44$, adj. $p < .05$; $t2(11) = 2.73$, adj. $p < .05$) whereas in the SV conditions, the standard deviations were not different between the related condition ($M = 105$, $SD = 31$) and the unrelated condition ($M = 112$, $SD = 35$; $t1(23) = 1.28$, adj. $p > .4$; $t2(11) = 2.32$, adj. $p > .08$), indicating that the SI effect on dispersion was only present in the OV utterances. This pattern corroborates the central tendency analysis we conducted above. This semantic effect on dispersion is visualized by means of vincentile plots in Figures 5 and 6.

Table 3
Mean (SD) Reaction Times Based on Participant Means in Milliseconds as a Function of Sentence Type and Relatedness in Experiment 3

	Relatedness	
Sentence type	Related	Unrelated
SV	766 (85)	764 (89)
OV	764 (103)	736 (80)

Note. SV = subject-verb; OV = object-verb.

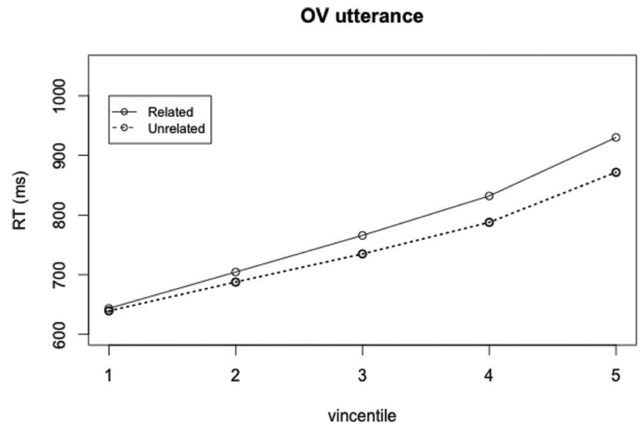

Figure 5. Vincentile plot for the OV condition in Experiment 3, with 20% percentile increments; that is, Vincentile 1 corresponds to the fastest 20% of trials per participant, Vincentile 2 corresponds to the next fastest 20% of trials per participant, and so on. RT = reaction time.

Discussion

Experiment 3 showed a SI effect in object-initial sentences but not in subject-initial sentences. Specifically, distractors that were semantically related to the sentence-final verb interfered with the utterance of the sentence-initial noun, but only when the sentence-initial noun was a direct object. Note that the same set of picture-distractor pairs was used as in Experiment 1, where SI effects were reliably obtained for transitive and intransitive verbs alike. Therefore, we can be confident that the contrast observed in Experiment 3 is a consequence of the sentence production task, and not simply a result of poor selection of picture-distractor pairs. Note also that previous research tends to report more consistent SI for intransitive than transitive verbs (Schnur et al., 2002; Tabossi & Collina, 2002). Therefore, it is unlikely that the observed contrast between SV and OV conditions is due to a smaller amplitude of SI in the SV conditions.

The contrast between SV and OV sentences is consistent with our hypothesis that verbs are selected before the onset of objects but not before subjects, reflecting the closer dependency of objects on verbs. The contrasting pattern of SI effects casts in a new light the previous failure to find an SI effect in verb-final utterances in German (Schriefers et al., 1998). Specifically, it seems premature to draw the conclusion that a producer can dispense with verb information when processing all of a verb's arguments in verb-final utterances. Certainly, our results do not show that speakers retrieve the verb's syntactic information before articulating the object in every utterance. However, they do show that Schriefers et al.'s (1998) findings are, in fact, consistent with the claim that verb choices guide structural processes in sentence planning. That is, the absence of an SI effect in Schriefers et al.'s data and in the SV condition here, combined with the presence of an SI effect in the OV condition here, likely reflects the selective nature of the advance planning mechanism for verbs.

[7] This larger SD value (60) was due to one participant showing an especially large standard deviation in RT. Removing this participant lowers the standard deviation to 33 but does not change the statistical results.

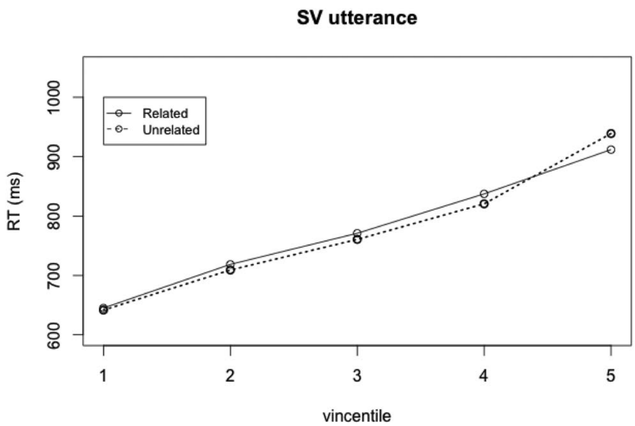

Figure 6. Vincentile plot for the SV conditions in Experiment 3, with 20% percentile increments; that is, Vincentile 1 corresponds to the fastest 20% of trials per participant, Vincentile 2 corresponds to the next fastest 20% of trials per participant, and so on. RT = reaction time.

A reviewer suggested that these results could reflect a potential difference in the relative "thematic fit" (cf. McRae, Spivey-Knowlton, & Tanenhaus, 1998) between the target nouns and target verb (*dog-NOM howl*/*cat-ACC pet*) versus target nouns and related distractor verbs (*dog-NOM cry*/*cat-ACC rub*). To address this concern, we conducted a post hoc norming study ($n = 11$) where participants rated the thematic fit between the target nouns and each type of verb (e.g., *dog-NOM howl* vs. *dog-NOM cry* for SV, *cat-ACC pet* vs. *cat-ACC rub* for OV utterances) on a 7-point Likert scale, similar to McRae et al. (1998). A mixed-effects analysis analogous to those used in the main experiments was conducted (this model included random intercepts for participants and items as well as random slope of relatedness for participants and items) and revealed a marginally significant interaction between verb type (target verb vs. related distractor verb) and sentence type (SV vs. OV; $\beta = -0.87$, $SE = 0.50$, $|t| = 1.74$, $p = .10$) as well as significant main effects of verb type ($\beta = 3.14$, $SE = 0.39$, $|t| = 6.02$, $p < .01$) and sentence type ($\beta = 1.26$, $SE = 0.43$, $|t| = 2.94$, $p < .01$). This marginally significant interaction partially supports the reviewer's concern. However, we find no evidence for a relationship between the magnitude of SI (in Experiment 3) and thematic fit difference (between noun-target verb and noun-distractor verb) in an item based correlation analysis ($n = 24$, $r = -0.14$, $p = .66$). Thus, it is unlikely that the differing SI effects for SV versus OV sentences can be attributed entirely to the difference in thematic fit.

General Discussion

The results of the current study indicate that SI from distractors related to noninitial verbs—a marker of advance verb selection—is obtained selectively before object noun articulation (Experiment 3). This selectivity is not likely to be due to the ineffectiveness of the semantic distractors for intransitive verbs (Experiment 1) or to an accidental relation between the preceding nouns and the distractors (Experiment 2). The SI effects reflected by sample means in each experiment are shown in Figure 7 below, and the same pattern was observed for the SI effect reflected by sample dispersion. These findings may reconcile the apparent conflict between the need for verb selection to determine the syntactic/semantic properties of preverbal arguments and the lack of verb-related SI effects before subject noun articulation in Schriefers et al. (1998). Based on this pattern of results, we argue that advance verb selection is selectively performed before object nouns, but not before subject nouns.

Before we outline the potential mechanism underlying the selective nature of advance verb planning, it is worth noting that the current result at first seems to conflict with previous findings in the scope of planning literature. In particular, the results of Experiment 3 might be regarded as in conflict with a previous finding by Meyer (1996), where the lemma of the second noun in a sentence like *"the dog is next to the baby"* appeared to show evidence of advance planning. That is, it appeared that object nouns, positioned after the predicate, were planned before articulation of the subject noun phrase (*the dog*). If a word that appears after a predicate is already planned, this invites the conclusion that the predicate is also planned. However, this conclusion derives from the assumption that the scope of planning at the lemma level is already linear. This assumption is, as far as we know, not warranted. In fact, the production model of Bock and Levelt (1994) assumes that linearization does not occur until the positional processing stage, which is subsequent to lemma selection. Also, it is important to note that speakers repeated the same structure (either *A and B* or *A is next to B*) throughout the experimental session in Meyer (1996), potentially expanding the scope of planning. The current experiment did not allow such formulaic production. Thus, it may not be appropriate to directly compare Meyer (1996) and the current experiments.

In addition, the absence of a SI effect in the SV conditions in Experiment 3 appears to be at odds with findings by Kempen and Huijbers (1983) and Schnur et al. (2002), which suggest that a verb's lemma selection occurs even before subject noun articulation. One crucial difference between the current experiment and the previous experiments is that their target utterances had less

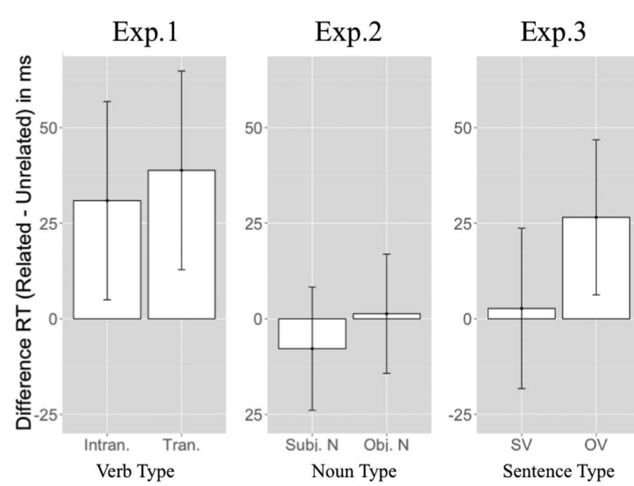

Figure 7. Semantic interference effect by verb type (Exp. 1), noun type (Exp. 2), and sentence type (Exp. 3), estimated from the statistical model. Error bars represent 95% confidence intervals. N = noun; Subj. = subject; Obj. = object; SV = subject-verb; OV = over-job; RT = reaction time.

variation in the preverbal noun phrases (four variants in Kempen and Huijbers (1983), two variants in Schnur et al. (2002), and 24 variants in the current experiments). Also, Schnur and colleagues' preverbal noun was a pronoun (s*he* or *he*), which is particularly easy to process. Thus, although our results and Schriefers et al.'s (1998) results suggest that a verb's lemma is not necessarily retrieved in advance of the articulation of subject nouns, it might under some circumstances be retrieved in advance of that point, in situations where there is less response variability and hence less processing difficulty in the preceding words (cf. the retrieval fluency hypothesis; Griffin, 2003).

There are at least three classes of possible mechanisms that could explain the selective pattern of advance verb planning observed in Experiment 3. The first possibility, which in a broad sense falls under the lexicalist view of sentence production, is that encoding some structural dependency relations between verbs and their arguments requires the advance selection of verbs. One such potential dependency relation is the assignment of case to verbs' arguments. Under this account, the selective pattern of advance verb selection we observed in Experiment 3 can be explained by a difference in how grammatical case is assigned between subject nouns and object nouns. As noted in the Introduction, accusative case is usually assigned by the lexical head V(erb), whereas nominative case is usually assigned by the inflectional head I(nflection). Thus, the necessity to assign accusative case via a verb head may require that the verb's lemma be selected before the encoding of the object noun is completed. Alternatively, it is possible that the building of structural representations itself is dependent on verbs (e.g., Ferreira, 2000). This is a stronger lexicalist position than the first alternative, in the respect that even the most elementary structural process of building or retrieving phrase structure is considered to be dependent upon verbs under this account. In other words, the current experiments may be interpreted as evidence of partial lexical guidance in phrase structure building: structure building for the subject noun position might not be dependent on verbs, whereas structure building for the object noun position would be crucially dependent on verbs.

The second possibility, which falls under a nonlexicalist view of sentence production, is that the selective advance selection of verbs observed here is due to syntactic constituency determining the scope of lemma selection. This possibility is broadly consistent with the proposal that syntactic phrases define the default scope of lemma planning (e.g., Smith & Wheeldon, 1999). This idea implies that phrase structure representations can be built prior to lemma selection independently from lexical items (unlike the lexicalist accounts described above), and that such structural representations control the dynamics of lemma selection processes. Although we are not aware of specific claims that the verb phrase defines the scope of planning, our data are consistent with the idea that the syntactic phrase is a determinant of the scope of planning at the lemma level. In this view, the dynamics of lemma selection are controlled by such higher-level syntactic schemes.

A final possibility is that it is the structure of semantic/conceptual representations, rather than syntactic phrases, which is the source of the contrasting pattern that we observed in the current study. Specifically, it is possible that internal arguments and verbs constitute an integrated unit that is planned together at the lemma level or higher (e.g., the message level). This is consistent with the linguistic analysis of Kratzer (1996, 2002), although it is unclear whether semantic representations in linguistic theory should be equated to message level representations in production models. Kratzer argued that external arguments should be excluded from the argument structure of predicates and suggested instead that inflectional heads introduce external arguments. If this is the case, then advance verb selection may be due to the semantic necessity to compute a predicate and its internal argument in tandem, or at least temporally closely, either at the conceptual or the lemma level.

There are of course some limitations in the current study. For instance, aside from these theoretically motivated interpretations discussed above, we acknowledge that the SV and OV utterances might differ in how much thematic processing participants must perform before starting to speak. In particular, OV utterances might require deeper thematic processing because they might have forced participants to omit the agent. However, participants in our experiment were probably not actually forced to produce OV sentences. Given the structure of the experiments, it was more natural for Japanese speakers to omit the subject than to explicitly mention the subject in the OV conditions because the agent entities in most OV pictures were underspecified (and Japanese disprefers the use of overt pronouns) and hence hard to name without explicit instruction. This is reflected in the RT data: OV utterances had slightly *shorter* response latencies than the SV sentences, suggesting that they were, if anything, less costly than SV utterances.[8] This difference cannot be attributed to the lexical frequency of subject versus object nouns, as the noun-only experiment showed closely matched average latencies for S nouns and O nouns. Thus, it is unlikely that the contrastive pattern of SI in SV and OV utterances can entirely be attributed to the difference in thematic processing depth.

Conclusion

The current study used an extended picture-word interference paradigm to probe the time course of verb planning, using Japanese as a test case due to its strongly verb-final word order. The results suggest that verb selection occurs selectively before articulation of the object noun, but not the subject noun in Japanese. This may reconcile past failures to find evidence for advance verb selection (Schriefers et al., 1998) with models that posit a central role for verbs in structural processes (Levelt, 1989; Kempen & Hoenkamp, 1987; Bock & Levelt, 1994; Ferreira, 2000).

[8] One might think that, if the verb is planned in advance in the OV condition, then the RTs in the OV sentences should generally be longer. However, it should be noted that this RT may largely reflect the speed of noun processing given the action picture, so it seems that naming a patient noun given a transitive action picture was easier than naming an agent noun given an intransitive picture. Also, it should be noted that verb planning might not be necessarily costly. This is because the retrieved verb lemma might facilitate the planning of object nouns via semantic priming, or what one might call the resonance between verbs and object nouns. This possible facilitative effect could counteract the cost of advance verb planning and could be additive to the SI effect from the related distractors.

References

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experi-*

mental Psychology: Learning, Memory, and Cognition, 33, 791–810. http://dx.doi.org/10.1037/0278-7393.33.4.791

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

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

Boersma, P., & Weenink, D. (2014). *Praat: Doing phonetics by computer* (Version 5.3.63) [Computer program]. Retrieved from http://www.praat.org/

Bresnan, J. (2001). *Lexical-functional syntax*. Oxford, UK: Blackwell.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition, 109,* 274–280. http://dx.doi.org/10.1016/j.cognition.2008.07.011

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language, 54,* 592–609. http://dx.doi.org/10.1016/j.jml.2005.12.008

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammar. In L. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). Cambridge, MA: MIT Press.

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82,* B1–B14. http://dx.doi.org/10.1016/S0010-0277(01)00138-X

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. http://dx.doi.org/10.3758/BF03196521

Haegeman, L. (1991). *Introduction to government and binding theory*. Oxford, UK: Blackwell.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109,* 340–347. http://dx.doi.org/10.1037/0033-2909.109.2.340

Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1363–1376. http://dx.doi.org/10.1037/a0036797

Iwasaki, N. (2011). Incremental sentence production: Observations from elicited speech errors in Japanese. In H. Yamashita, Y. Hirose, & J. Packard (Eds.), *Processing and producing head-final structures* (pp. 131–151). Dordrecht, Heidelberg, London and New York: Springer.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11,* 201–258. http://dx.doi.org/10.1207/s15516709cog1102_5

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14,* 185–209. http://dx.doi.org/10.1016/0010-0277(83)90029-X

Kratzer, A. (1996). Severing the external argument from the verb. In J. Rooryck & L. Zaring (Eds.), *Phrase Structure and the Lexicon* (pp. 109–137). Dordrecht, The Netherlands: Kluwer. http://dx.doi.org/10.1007/978-94-015-8617-7_5

Kratzer, A. (2002). *The event argument and the semantics of verbs*. Amherst: University of Massachusetts.

Lasnik, H. (2003). On the extended projection principle. *Studies in Modern Grammar, 31,* 1–23.

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

Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38,* 283–312.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language, 35,* 477–496. http://dx.doi.org/10.1006/jmla.1996.0026

Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66,* B25–B33. http://dx.doi.org/10.1016/S0010-0277(98)00009-2

Momma, S., Slevc, R., & Phillips, C. (2015, March 19–21). *The timing of verb planning in active and passive sentence production*. Poster presented at the 28th annual CUNY Conference on Human Sentence Processing, Los Angeles, CA.

Pollard, C., & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Chicago, IL: University of Chicago Press.

Scaltritti, M., Navarrete, E., & Peressotti, F. (2015). Distributional analyses in the picture-word interference paradigm: Exploring the semantic interference and the distractor frequency effects. *The Quarterly Journal of Experimental Psychology, 68,* 1348–1369.

Schnur, T. T. (2011). Phonological planning during sentence production: Beyond the verb. *Frontiers in Psychology, 2,* 319.

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

Schnur, T. T., Costa, A., & Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of Psycholinguistic Research, 35,* 189–213. http://dx.doi.org/10.1007/s10936-005-9011-6

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

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. http://dx.doi.org/10.1006/jmla.1998.2578

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

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition, 114,* 447–454. http://dx.doi.org/10.1016/j.cognition.2009.11.003

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., . . . Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51,* 247–250. http://dx.doi.org/10.1016/j.jml.2004.03.002

Tabossi, P., & Collina, S. (2002). The picture–word interference paradigm: Conceptual effects in the production of verbs. *Rivista di Linguistica, 14,* 27–41.

Vincent, S. B. (1912). The function of the vibrissae in the behavior of the white rat. *Behavioral Monographs, 1*(5).

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language, 37,* 356–381. http://dx.doi.org/10.1006/jmla.1997.2517

(*Appendices follow*)

Appendix A

R Code Used in Experiment 1, 2 and 3

 # R Code Used for Experiment 1

$$Exp1m1 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 + Transitivity * Relatedness \mid Subject) \; + \; (1 \; + Relatedness \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

$$Exp1m2 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 + Transitivity + Relatedness \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

$$Exp1m3 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 + Relatedness \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

$$Exp1m4 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 + Transitivity \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

$$Exp1m5 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

$$Exp1m6 < -lmer(RT \sim Transitivity * Relatedness \; + \; (1 \mid Subject) \; + \; (1 \mid Item), \; data \; = \; Exp1\_data, \; RMEL \; = \; FALSE)$$

 # R Code Used for Experiment 2

$$Exp2m1 < -lmer(RT \sim NounType * Relatedness \; + \; (1 + NounType * Relatedness \mid Subject) \; + \; (1 \; + Relatedness \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

$$Exp2m2 < -lmer(RT \sim NounType * Relatedness \; + \; (1 + NounType + Relatedness \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

$$Exp2m3 < -lmer(RT \sim NounType * Relatedness \; + \; (1 + Relatedness \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

$$Exp2m4 < -lmer(RT \sim NounType * Relatedness \; + \; (1 + NounType \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

$$Exp2m5 < -lmer(RT \sim NounType * Relatedness \; + \; (1 \mid Subject) \; + \; (1 + Relatedness \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

$$Exp2m6 < -lmer(RT \sim NounType * Relatedness \; + \; (1 \mid Subject) \; + \; (1 \mid Item), \; data \; = \; Exp2\_data, \; RMEL \; = \; FALSE)$$

(*Appendices continue*)

 # R Code Used for Experiment 3

$$Exp3m1 < -lmer(RT \sim SentenceType * Relatedness + (1 + SentenceType * Relatedness \mid Subject) + (1 + Relatedness \mid Item), data = Exp3\_data, RMEL = FALSE)$$

$$Exp3m2 < -lmer(RT \sim Transitivity * Relatedness + (1 + SentenceType + Relatedness \mid Subject) + (1 + Relatedness \mid Item), data = Exp3\_data, RMEL = FALSE)$$

$$Exp3m3 < -lmer(RT \sim SentenceType * Relatedness + (1 + Relatedness \mid Subject) + (1 + Relatedness \mid Item), data = Exp1\_data, RMEL = FALSE)$$

$$Exp3m4 < -lmer(RT \sim SentenceType * Relatedness + (1 + SentenceType \mid Subject) + (1 + Relatedness \mid Item), data = Exp3\_data, RMEL = FALSE)$$

$$Exp3m5 < -lmer(RT \sim SentenceType * Relatedness + (1 \mid Subject) + (1 + Relatedness \mid Item), data = Exp1\_data, RMEL = FALSE)$$

$$Exp3m6 < -lmer(RT \sim SentenceType * Relatedness + (1 \mid Subject) + (1 \mid Item), data = Exp3\_data, RMEL = FALSE)$$

(*Appendices continue*)

Appendix B

Items Used in Experiments 1, 2, and 3

Target noun (Exp. 2 & 3)	Target verb (Exp. 1 & 2)	Sentence type (Exp. 2)	Related distractor	Unrelated distractor
アイス (aisu: ice cream)	溶ける (tokeru: melt)	SV	崩れる (kuzureru: lose shape)	浮かぶ (ukabu: float)
カエル (kaeru: frog)	跳ねる (haneru: jump)	SV	はずむ (hazumu: bounce)	噛む (kamu: bite)
家 (ie: house)	燃える (moeru: burn)	SV	焼ける (yakeru: burn)	舞う (mau: hover)
犬 (inu: dog)	吠える (hoeru: bark)	SV	鳴く (naku: cry)	崩れる (crash/break)
独楽 (koma: spinning top)	回る (mawaru: spin)	SV	巡る (meguru: move around)	焼ける (yakeru: burn)
石 (ishi: stone)	沈む (shizumu: sink)	SV	浮かぶ (ukabu: float)	滴る (shitataru: drip)
蜂 (hachi: bee)	刺す (sasu: sting)	SV	噛む (kamu: bite)	はずむ (hazumu: bounce)
鐘 (kane: bell)	鳴る (naru: ring)	SV	とどろく (todoroku: reverberate)	転ぶ (korobu: fall over)
雨 (ame: rain)	降る (furu: fall from the sky)	SV	滴る (shitataru: drip)	鳴く (naku: cry)
風船 (fuusen: balloon)	割れる (wareru: pop)	SV	砕ける (kudakeru: crash)	巡る (meguru: move around)
飛行機 (hikouki: airplane)	落ちる (ochiru: fall)	SV	転ぶ (korobu: fall over)	とどろく (todoroku: reverberate)
鳥 (tori: bird)	飛ぶ (tobu: fly)	SV	舞う (mau: hover)	砕ける (kudakeru: crash)
ピアノ (piano: piano)	弾く (hiku: play)	OV	鳴らす (narasu: sound)	解く (toku: open/untie)
ボール (boru: ball)	捕る (toru: catch)	OV	つかむ (tsukamu: grab)	さする (sasuru: rub/stroke)
レ モン (remon: lemon)	絞る (shiboru: squeeze)	OV	ねじる (nejiru: twist)	見る (miru: look/watch)
手 (te: hand)	叩く (tataku: hit/clap)	OV	打つ (utsu: strike)	吐く (haku: exhale)
本 (hon: book)	読む (yomu: read)	OV	見る (miru: look/watch)	ねじる (nejiru: twist)
猫 (neko: cat)	なでる (naderu: pet)	OV	さする (sasuru: rub/stroke)	鳴らす (narasu: sound)
花 (hana: flower)	摘む (tusmu: pick)	OV	むしる (mushiru: pluck)	湿らす (shimerasu: moisten)
財布 (saifu: wallet)	盗む (nusumu: steal)	OV	くすねる (kusuneru: pilfer)	こする (kosuru: rub)
鍵 (kagi: key)	開ける (akeru: open)	OV	解く (toku: open/untie)	打つ (utsu: strike)
髪 (kami: hair)	乾かす (kawakasu: dry)	OV	湿らす (shimerasu: moisten)	むしる (mushiru: pluck)
窓 (mado: window)	拭く (fuku: wipe)	OV	こする (kosuru: rub)	くすねる (kusuneru: pilfer)
たばこ (tabako: cigarette)	吸う (suu: inhale/suck)	OV	吐く (haku: exhale)	つかむ (tsukamu: grab)

Note. The items have their Romaji transcription and approximate English translations in parentheses. In Experiment 2, target nouns are inflected with appropriate case-marker (-ga in SV and --o in OV sentences). SV = subject-verb; OV = object-verb.

Received October 15, 2014
Revision received July 20, 2015
Accepted August 5, 2015 ■