2023

Producing long-distance dependencies in English and Japanese

Mari Kugemoto

Shota Momma

Mari Kugemoto and Shota Momma

Chapter 8
Producing long-distance dependencies in English and Japanese

1 Introduction

In sentence production, it is widely assumed that speakers can start speaking sentences without extensive *look-ahead*; later-coming words and structures in a sentence are not necessarily planned before its articulation onset (e.g., Griffin 2001; Levelt 1989; De Smedt 1990; Allum and Wheeldon 2007, 2009; Schriefers, Teruel, and Meinshausen 1998, Brown-schmidt et al. 2006; Brown-Schmidt and Konopka 2008; among others). For instance, Griffin (2001) suggested that when uttering sentences like *The A and the B are above the C*, speakers began to speak "The A. . ." before planning "B" and "C". However, previous studies mostly examined relatively simple sentences where sentence-initial constituents do not depend on later coming words, and those studies tend to focus on whether the later-coming *words* are planned before the initiation of an utterance and thus little is known about how the later-coming *structures* are planned before the initiation of an utterance (cf. Wheeldon et al. 2013). Consequently, how speakers plan structural representations of complex sentences is largely unknown. For example, it remains unclear how speakers plan sentences that contain *filler-gap dependencies,* as in *what do you think the dog ate?* In this sentence, the sentence-initial constituent *(who)* is the "*filler*" that fills the "*gap*," the missing object position, after the verb *ate.* Filler-gap dependencies are intensively studied in analytical linguistics (Chomsky 1957, 1965, 1995; Frank 2004; Kroch and Joshi 1985; Pollard and Sag 1994; Ross 1967; among many other) and in sentence comprehension research (Aoshima, Phillips, and Weinberg 2004; Frazier and Clifton 1989; Fodor 1978; Frazier, Clifton, and Randall 1983; Frazier and d'Arcais 1989; Garnsey, Tanenhaus, and Chapman 1989; Omaki et al. 2015; Wanner and Maratsos 1978; among many other). In comparison, limited attention has been paid to filler-gap dependency production. Studying filler-gap dependency production is important in constructing theories of production that are not limited in scope and connecting sentence production research with analytical linguistics. It is also likely to be useful in understanding the relationship between production and working memory mechanisms. Against this background, the current chapter aims to study the nature of the production mechanisms involved in planning sentences involving filler-gap dependencies, specifically focusing on the production of wh-dependencies, a type of filler-gap dependencies found in wh-questions.

https://doi.org/10.1515/9783110778946-008

1.1 Two strategies for producing wh-dependencies

The current study investigates the time-course of wh-dependency formation in English and Japanese sentence production. We compare two possible hypotheses about how speakers form wh-dependencies in speaking: the *late commitment hypothesis* and the *early commitment hypothesis* (Momma 2021). The late commitment hypothesis claims that the grammatical status of the gap is not specified when speaking the filler. For example, in sentences like *What do you think the dog ate?*, the grammatical status of *what* is not specified when *what* is uttered; in the extreme case, it may not be determined up until the materials immediately preceding the gap need to be uttered. The late commitment hypothesis allows sentence production to proceed flexibly because speakers can keep various options open throughout their production. This flexibility may be beneficial because speakers can avoid having to say a word that they are not ready to say (Ferreira 1996). For instance, when an agent noun is difficult to retrieve, speakers may want to use the passive voice to postpone it (e.g., when speakers have difficulty retrieving the word *professor*, speakers may want to say *Who was introduced by the professor?* instead of saying *Who was the professor introducing?*). If speakers commit to the object status of the filler before saying *what*, this strategy would be unavailable. At the same time, the late commitment strategy may be disadvantageous because speakers could "talk themselves into the corner." For example, when the gap happens to correspond to the participant of the event described by a relative clause, wh-dependencies would fail to be established due to the constraint that the gap cannot be posited inside a relative clause (i.e., due to the relative clause island; Ross 1967 among others). If speakers do not decide the structural position of the gap until late in the utterance, they might start speaking the filler and later realize that the filler cannot be associated with the appropriate grammatical position due to various constraints on filler-gap dependencies (Ross 1967).

In contrast to the late commitment hypothesis, the early commitment hypothesis claims that the grammatical status of the filler is already determined before the filler is uttered. For example, in *What do you think the dog ate?*, before *what* is spoken, speakers already represent *what* as the object of the verb *ate*. This strategy is beneficial because speakers can avoid positing an illicit gap. But one disadvantage is that speakers lose flexibility in their production. For example, the passivization strategy discussed above would not be available if speakers have already decided the grammatical position of the filler before starting to speak it. Considering both the late and early commitment hypotheses have advantages and disadvantages, either of those hypotheses is plausible. The present study aims to test those hypotheses in both English (Experiment 1) and Japanese (Experiment 2).

Of course, speakers of different languages may use different strategies for filler-gap dependency production, depending on the properties of the languages they

speak. For instance, English and Japanese differ in the usual position of wh-phrases. In English, wh-phrases are moved to the left edge of a clause in most cases, while in Japanese they often stay in their canonical positions. When wh-phrases are moved in Japanese, the movement may be driven by a different cause than in English. Because wh-phrases in English and Japanese show different distributional properties and their movement may be driven by distinct causes, English and Japanese speakers may plan filler-gap dependencies differently. For instance, in English, speakers may develop the strategy to plan the grammatical status of the filler early to avoid violating the various constraints on long-distance dependencies, according to the early commitment hypothesis. In contrast, Japanese speakers may not adopt the early commitment strategy because wh-phrases often do not appear at the sentence-initial position, and perhaps because the constraints on long-distance dependencies may be generally more relaxed in Japanese (Kuno 1973; Omaki et al. 2020). Thus, English and Japanese speakers may reasonably differ in how they establish filler-gap dependencies in production. But they may also use fundamentally similar mechanisms for filler-gap dependency production. The present study thus aims to compare the time-course of wh-dependency formation in English and Japanese, to explore how typological differences may affect wh-dependency planning mechanisms.

1.2 A method for investigating the time-course of wh-dependency production

To investigate wh-dependency planning processes in English and Japanese, we used a close relative of the method that Momma (2021) used to investigate the time-course of filler-gap dependency planning. Before we elaborate on the current method, the basic logic of the method used in Momma (2021) should be explained. The method relied on two previously well-established phenomena: the structural priming effect (Bock 1986; see Pickering and Ferreira 2008 and Mahowald et al. 2016 for a recent overview) and the *that*-trace constraint (Perlmutter 1971; see Pesetsky 2017 for a recent overview).

The structural priming effect refers to a phenomenon that speakers tend to re-use the same structures they recently encountered (Bock 1986). For instance, after encountering a prepositional dative sentence like *I showed my drawing to her*, speakers are more likely to produce the prepositional dative structure *The boy gave the ball to the dog* than its double object counterpart, *The boy gave the dog the ball*. Structural priming can occur without any overlap in words between prime and target sentences (Bock 1986). Usually, the structural priming effect is measured as the increase in the production rate of a particular structure, but structural priming

has also been shown to speed up the production of the primed structure (Wheeldon and Smith 2003; Seagert, Wheeldon, and Hagoort 2016). Most relevantly for current purposes, the complementizer *that* can be structurally primed. Ferreira (2003) reported that sentences with the complementizer *that* increased the likelihood of speakers using *that* in the subsequent production. For example, after encountering prime sentences like *The director announced that Hollywood's hottest actor would be playing the part*, speakers were more likely to produce *that* in target sentences like *The jury believed that the young witness told the truth* than after encountering minimally different prime sentences like *The director announced Hollywood's hottest actor would be playing the part*. This complementizer priming is not reducible to the priming of the phonological form of *that*. This is because Ferreira (2003) showed that the demonstrative *that* as in *that dog* did not prime the complementizer *that*, and because the null complementizer also primed the null complementizer. Thus, the complementizer priming is best characterized as priming at the structural level, not the phonological level.

Momma (2021) also used the constraint known as the *that*-trace effect (Perlmutter 1971; see Ritchart et al. 2016 for laboratory-based experimental evidence for this effect). The *that*-trace constraint bans the structures where the complementizer *that* is followed by the gaps in the following:

(1) *Which girl do you think that ate the cake?

This effect is not observed in sentences where the gap corresponds to the embedded object position, as in the following:

(2) Which cake do you think that the girl ate?

Importantly, the *that*-priming effect and *that*-trace constraint conflict with each other. The *that*-priming effect encourages speakers to say *that* while the *that*-trace constraint prohibits them to say *that.* Momma (2021) showed that this conflict between the *that*-priming effect and the *that*-trace constraint slowed down the planning process. That is, speakers are slower to speak sentences like *Who do you think met the girl?* given prime sentences with *that* like *The boy thinks that the dog liked them* than given minimally different prime sentences without *that*, presumably due to the conflict between the *that*-priming effect and the *that*-trace constraint. Critically, in a series of picture description experiments, it was observed that this slow-down effect appeared before the sentence onset of utterances, that is, before starting to say the filler. This suggests that speakers already plan the grammatical function of the filler, as well as the complementizer structure of the gap-containing clause, in accordance with the early commitment hypothesis.

1.3 Current experiments

Having explained the logic used in Momma (2021), we are now ready to describe the current experiments. There are two experiments in the current study. Previous studies on sentence planning often used picture description tasks (Allum and Wheeldon 2007, 2009; Schriefers, Teruel, and Meinshausen 1998; Smith and Wheeldon 1999; Konopka and Meyer 2014; among others). However, because it is difficult to elicit complicated target sentences of interest in English and Japanese using a picture description task, the current study alternatively used a variant of sentence recall task. The working assumption is that sentence recall involves the regeneration of memorized sentences from their conceptual representations (Potter and Lombardi 1998). In both Experiments 1 and 2, participants memorized one target sentence and one prime sentence in this order and recited the target sentence. In this task, because the prime sentence is the last sentence they encounter before uttering the target sentence, the structure of the prime sentence would be primed in the target production.

In Experiment 1, we examined if English speakers plan the grammatical status of the gap before starting to speak the filler, as in Momma (2021), but using the sentence recall task. We aim to evaluate if the results from Momma (2021) can be conceptually replicated and if they can generalize to different task contexts. In Experiment 1, prime and target sentences were like the following:

(3) Prime sentences
  a. Do you think that the student solved the question? (*that* prime)
  b. Do you think the student solved the question? (null prime)

(4) Target sentences
  a. Which trainer do you think loved the lion? (subject extraction)
  b. Which trainer do you think the lion loved? (object extraction)

Given prime sentences with the complementizer *that* like (3a), speakers should be more inclined to say *that* in target sentences than given prime sentences like (3b). However, when the target sentence is an embedded subject wh-question like (4a), the complementizer *that* cannot be used because the *that*-trace constraint prohibits the complementizer *that* followed by the subject gap. Thus, the *that* priming and the *that*-trace constraint creates a conflict in production of sentences like (4a) given a prime sentence with the complementizer *that* like (3a).

Experiment 2 aimed to test whether Japanese speakers plan wh-dependencies before speaking the wh-phrase scrambled to the sentence-initial position. However, because, as far as we know, Japanese does not have a structure that can potentially

violate the *that*-trace constraint, we used a different effect to make inferences about the timing of filler-gap dependency planning. Namely, we used simple structural priming on two types of wh-questions with two different scope relations. In Japanese, wh-phrases are associated with the question particle, *-ka*. When a sentence is bi-clausal and a wh-phrase is extracted from the embedded clause, the position of the Q-particle determines the scope of the wh-phrase.

(5) a. どの ライオンが 逃げた と 言いました か? (matrix)
Which lion-NOM ran-away that said-POLITE Q
'Which lion did you say ran away?'
b. どの ライオンが 逃げた か 言いました か? (embedded)
Which lion-NOM ran-away Q said-POLITE Q
'Did you say which lion ran away?'

In both sentences, the wh-phrase *which lion* occurs in the initial position of the sentence. But in (5a), it is associated with the sentence-final Q-particle and has the matrix scope. In contrast, when the wh-phrase is associated with the Q-particle in the embedded clause as in (5b), it is usually interpreted to have the embedded scope, although it could have the matrix scope when prosodically licensed. Based on the finding by Wheeldon and Smith (2003) and Segaert et al. (2017) that speakers are faster to speak the primed structures, we predicted that speakers should be faster to plan target sentences when prime sentences have the same scope relation as target sentences. If this potential speed-up effect is observed in the onset latency of target utterances where wh-fillers are fronted, it can be inferred that speakers plan (a) whether the wh-filler is associated with the embedded or matrix complementizer and (b) the type of complementizer used for the embedded and matrix clause, before starting to speak sentence-initial wh-fillers. If this prediction is met, it can be argued that Japanese speakers plan the structural representations of wh-dependencies early, before starting to speak the scrambled wh-filler, just like English speakers. More specifically, it can be argued that both English and Japanese speakers minimally plan the complementizer structure of the clause that the relevant wh-phrase is taking scope over, before starting to speak the sentence-initial wh-fillers.

2 Experiment 1

Like in Momma (2021), Experiment 1 examined the timing of wh-dependency formation in English using the conflict between the *that*-priming effect and the

that-trace constraint in subject extracted wh-questions. The early commitment hypothesis predicts that this conflict would cause a slow-down effect at the onset of subject-extracted wh-questions, but not object-extracted wh-questions.

2.1 Method

2.1.1 Participants

Forty-eight monolingual English speakers were recruited via Prolific Academic. Informed consent was obtained from each participant. Each participant was paid five US dollars as compensation for the 20–30 minutes experiment. We replaced eleven participants who did not follow instructions or whose recordings were not intelligible and two additional participants who had less than half error-free trials.

2.1.2 Materials

For the target sentences, forty-eight pairs of subject-extracted wh-questions and object-extracted wh-questions like (4a) and (4b) were constructed (see Table 1). All sentences began with *Which NP do you think. . .* The prime sentences like (3a) and (3b) were forty-eight yes-no questions either with or without the complementizer *that*. They began with either *Do you. . .* or *Do they. . .* The prime sentences were paired with the target sentences so that they did not share the content words aside from the embedding verb *think*. They also did not have any obvious semantic relationship.

Table 1: The four conditions in Experiment 1.

condition	target sentence	prime sentence
subject-extraction / *that* prime	Which trainer do you think loved the lion?	Do you think that the student solved the question?
subject-extraction / null prime		Do you think the student solved the question?
object-extraction / *that* prime	Which trainer do you think the lion loved?	Do you think that the student solved the question?
object-extraction / null prime		Do you think the student solved the question?

2.1.3 Procedure

The experiment was conducted online using PCIbex (Zehr and Schwarz 2018). At the beginning of the experiment, there were three practice trials, which had the same task structure as the experimental trials. The experimental trials were structured as follows. First, a target sentence was presented for 5000 ms. Participants were instructed to read it aloud and memorize it. Subsequently, a prime sentence was presented for 5000 ms, which participants also read aloud and memorized. After a blank screen presented for 2000 ms, either '1' or '2' in the red font was presented as the prompt for recall. Participants were instructed to recite the first sentence when '1' was presented. They were instructed to recite the second sentence when '2' was presented. In critical trials, '1' was always presented, as the target sentences were always presented as the first sentence. In filler trials, which were indistinguishable from critical trials from the participants' perspectives, participants were presented with '2.' Thus, speakers could not reliably predict which sentence they needed to recall. There were 48 critical trials and 24 filler trials.

2.1.4 Scoring and analysis

All audio files were first transcribed and coded for errors. Errors were defined as any deviations from target sentences. Incomplete utterances, trials where participants were still uttering the previous sentence after the recall prompt, and trials where participants uttered overt hesitation (uh, am, etc.) before finishing the sentence, were also coded as erroneous. The erroneous trials were excluded from the subsequent analysis. Trials where participants said the complementizer *that* in the object extracted wh-question (e.g., *Which trainer do you think <u>that</u> the lion loved?*) and trials where participants replaced *you* with *they* (e.g., *Which trainer do they think the lion loved?* for the target *Which trainer do you think the lion loved?*) were included in the analysis. The onset latency of the error-free trials was manually measured using Praat, by the authors and a research assistant who were all blind to the prime type condition.

Using R (R Core Team 2020) and *lmer* package (Bates et al. 2015), a linear mixed-effects model was fit for the onset latency of target sentences. The model was initially maximal in the sense of Barr et al. (2013), but due to the convergence issue, the random slopes were removed from the model. When simplifying the model, the random slope that accounted for the least amount of variance was removed successively, until the model converged. The final model had PrimeType (that vs. null), ExtractionType (subject vs. object) and their interaction as fixed effects, and by-subject and by-item random intercepts.

2.2 Result

In Experiment 1, 30.8 % of the trials (606 out of 1968 trials) were excluded from the subsequent analyses as erroneous trials. The error rates in each condition are shown in Table 2. The trials where the onset latency is longer than 2500 ms (16 out of 1968 trials; 0.8%) were excluded as well.

Table 2: Error rates in each condition in Experiment 1.

condition	error rate
subject-extraction / *that* prime	28.7 %
subject-extraction / null prime	32.5 %
object-extraction / *that* prime	30.5 %
object-extraction / null prime	28.2 %

As shown in Figure 1, in the subject extraction condition, speakers were 46 ms slower in the *that* prime condition than in the null prime condition, but in the object extraction condition, they were 8 ms slower in the *that* prime condition than in the null prime condition. Supporting this pattern, the statistical model showed that the interaction between ExtractionType and PrimeType was significant ($\beta = 0.05$, $SE = 0.02$, $|t| = 2.13$, $p = 0.03$). In addition, the planned comparison based on the nested models showed that the simple effect of PrimeType was significant in the subject extraction condition ($\beta = 0.05$, $SE = 0.02$, $|t| = 2.73$, $p = 0.006$), but not in the object extraction condition ($p = 0.78$). The main effect of Prime Type was marginally significant ($p = 0.08$), but this is not interpretable given the interaction involving this term. The main effect of Extraction Type was not significant ($p = 0.77$).

2.3 Discussion

The results of Experiment 1 showed that there was a slow-down effect in onset latency selectively in the subject extraction condition, but not in the object extraction condition. This pattern replicates Momma (2021) but in a different task environment. This suggests that speakers know that, as early as at the sentence onset, the subject-extracted question is not compatible with the *that* complementizer. That is, speakers plan the structural properties of wh-dependency, specifically the grammatical function of the extracted wh-phrase and the complementizer type of the gap-containing clause before uttering it. This supports the early com-

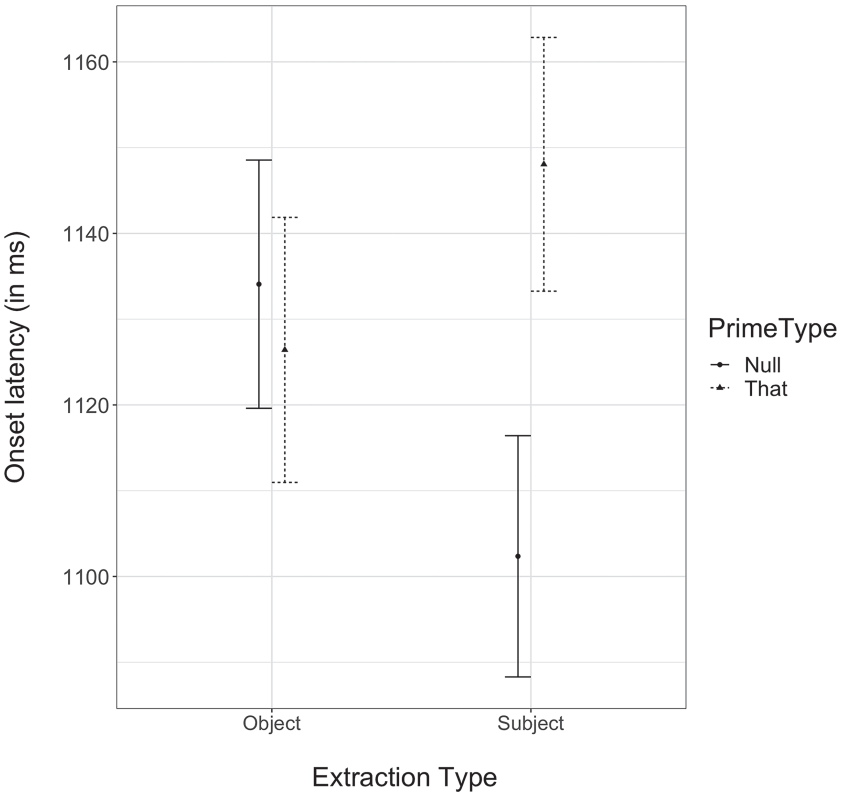

Figure 1: By-subject mean onset latency across four conditions in Experiment 1. Error bars represent the standard error of the means.

mitment hypothesis, which claims that speakers plan the grammatical details of wh-dependencies before uttering the sentence-initial filler.

3 Experiment 2

Experiment 1 showed that speakers plan the grammatical details of wh-dependency in English. However, this early commitment strategy may be language-specific. For example, because some constraints on filler-gap dependencies may be relaxed (or even absent) in Japanese (Kuno 1973; Omaki et al. 2020), Japanese speakers may have weaker motivations for planning the grammatical status of the filler/gap before the filler production. Experiment 2 investigated if Japanese speakers nevertheless use the early commitment strategy for planning wh-dependency despite

relevant typological differences. As discussed in the introduction, we used the potential speed-up effect in onset latency due to structural priming. Specifically, we hypothesized that the scope of wh-phrases can be primed, and this priming effect would lead to faster onset latency when target sentences share the same wh-scope with prime sentences. If this potential speed-up effect is observed before the onset of sentence-initial wh-phrases, it can be inferred that speakers plan at least the scope relation of wh-phrases and by extension the complementizer type of the embedded clause, before starting to speak the wh-filler.

3.1 Method

3.1.1 Participants

Thirty-five native Japanese speakers participated in Experiment 2 online. For those who live outside of Japan, it was confirmed that they acquired Japanese in their infancy and use Japanese daily via a questionnaire. No demographic information was collected other than language backgrounds. Each participant was paid ten US dollars or 1000 yen per an hour as compensation for the 30–45 minute experiment. We replaced two participants who did not follow instructions or whose recordings were not intelligible and nine additional participants who had less than half error-free trials.

3.1.2 Materials

The stimuli were questions like (5a) and (5b). Table 3 shows the four conditions of the prime and target sentence combinations. All sentences had the same matrix verb and ending 言いましたか (‘said-POLITE-Q’) to make the sentences easier to memorize. In addition, to make the sentences as simple as possible, wh-phrases were always the subject and all verbs were intransitive verbs or verbs whose objects can be omitted naturally without contextual support. Because Japanese is a pro-drop language that allows pronouns to be omitted, the sentences like (5a) and (5b) are in principle ambiguous between the parse where the matrix subject is dropped and the parse where the embedded subject is dropped. However, to force participants to interpret the subject as extracted from the embedded clause, all subject noun phrases were headed by non-human nouns except for ‘the baby’. This would prevent the parse where the embedded subject is dropped because the parse where non-human noun phrases function the subject of the matrix verb *say* yields implausible interpretation (e.g., *Which lion said you ran away?*).

Table 3: The four conditions in Experiment 2. The sentences are translated from Japanese.

condition	target sentence	prime sentence
matrix scope / matching scope prime	Which lion did you say ran away?	Which train did you say stopped?
matrix scope / mismatching scope prime		Did you say which train stopped?
embedded scope / matching scope prime	Did you say which lion ran away?	Which train did you say stopped?
embedded scope / mismatching scope prime		Did you say which train stopped?

3.1.3 Procedure

The same procedure as in Experiment 1 was used.

3.1.4 Scoring and analysis

All audio files were transcribed and coded for errors using the same criteria as in Experiment 1. Onset latencies were measured with the same procedure as in Experiment 1. The onset latency of target sentences was analyzed using linear mixed-effects modeling. The model was initially maximal but was simplified in the same way as in Experiment 1 due to the convergence issue. The final model had PrimeType (match vs. mismatch), Scope (matrix vs. embedded) of the target sentence, and their interaction as fixed effects, and by-subject and by-item random intercepts.

3.2 Result

In Experiment 2, 31 % of the trials (521 out of 1680 trials) were excluded from the subsequent analyses as erroneous trials. The error rates in each condition are shown in Table 4. Onset latencies longer than 2500 ms (0.5 %, 8 out of 1680 trials) were also excluded.

As can be seen in Figure 2, in the embedded scope condition, speakers were 61.1 ms slower in the mismatch condition than in the match condition, but in the matrix scope condition, they were 30.6 ms faster in the mismatch condition than in the match condition. Supporting this pattern, the statistical model showed that the interaction between Scope and PrimeType was significant (β = -0.08, SE = 0.02, $|t|$ = 3.16, p = 0.002).

Table 4: Error rates in each condition in Experiment 1.

condition	error rate
matrix scope / matching scope prime	21.7 %
matrix scope / mismatching scope prime	46.4 %
embedded scope / matching scope prime	15.5 %
embedded scope / mismatching scope prime	40.5 %

In addition, the planned comparison based on the nested models showed that the simple effect of PrimeType was significant in the embedded scope condition (β = 0.04, *SE* = 0.02, |t| = 2.6, *p* = 0.01), and it was marginally significant in the matrix scope condition (β = -0.03, *SE* = 0.02, |t| = 1.87, *p* = 0.06). Neither the main effect of Prime Type (*p* = 0.65) nor the main effect of Scope (*p* = 0.44) was significant.

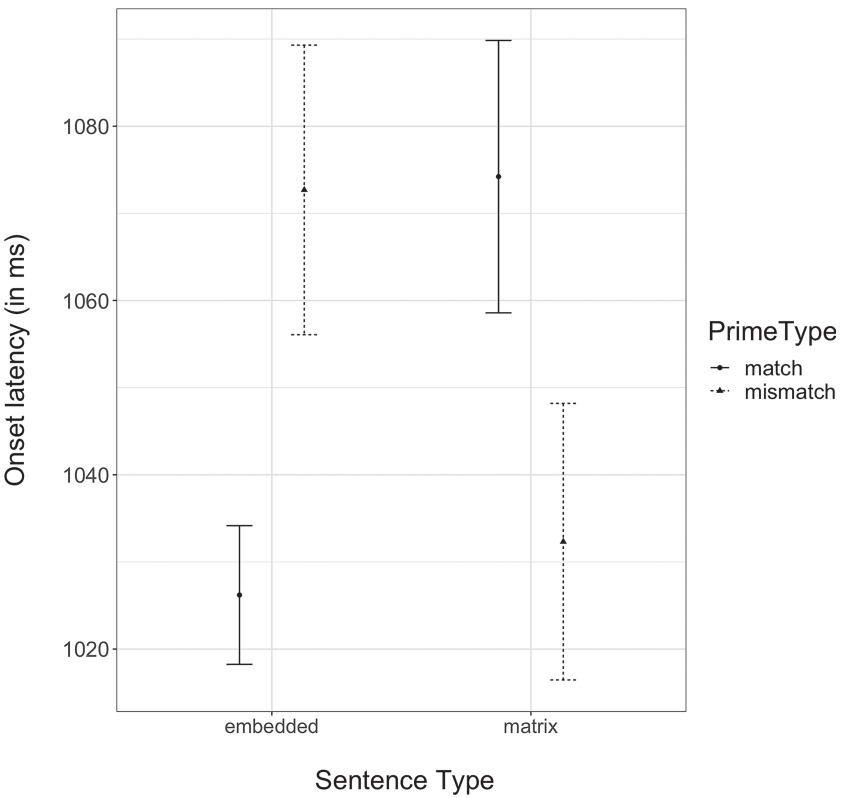

Figure 2: By-subject mean onset latency across four conditions in Experiment 2. Error bars represent the standard error of the means.

3.3 Discussion

The results showed that speakers were faster to start speaking sentences with the embedded scope given prime sentences with the embedded scope. In contrast, speakers were marginally slower to start speaking sentences with the matrix scope given prime sentences with the matrix scope. There was an interaction between Prime Type and Scope. We suggest that this interaction can be explained by assuming a facilitatory effect of scope priming (cf. Wheeldon and Smith 2003) and an inhibitory effect of similarity-based interference (Lewis 1996), which to some extent cancel each other out. First, the similarity-based interference slows production planning when the prime and the target sentences are similar in scope, perhaps because two sentences are less discriminable from each other when they have the same scope properties (in the match condition) than when they have distinct scope properties (in the mismatch condition). This effect of similarity-based interference is masked by the facilitatory effect of scope-related structural priming in the embedded scope condition. However, the similarity-based interference effect in the matrix scope condition remains observable because the scope-related structural priming effect is less strong in the matrix scope condition. The reason that the structural priming effect is less strong in the matrix scope condition may be due to the effect known as the *inverse preference* effect (Jaeger and Snider 2008; Reitter, Keller, and Moore 2011; Bernolet and Hartsuiker 2010; Ferreira 2003; among others). In the structural priming literature, it is widely observed that less frequent structures are more easily primed than more frequent structures. It is reasonable to assume that the matrix scope is less primable than the embedded scope because the matrix scope wh-questions occur even in sentences without any embedded clauses (i.e., in mono-clausal sentences). If the matrix wh-scope in bi-(or multi-) clausal sentences and mono-clausal sentences are treated as the same type of dependency configuration, the matrix wh-scope would be more frequent than the embedded scope interpretation. Given the inverse preference effect, it may be harder to prime the matrix scope structures than to prime the embedded scope structures. If this is the case, the similarity-based interference effect should mask the small structural priming effect in the matrix scope condition, but the relatively large structural priming effect should mask the similarity-based interference effect in the embedded scope condition. Thus, the combination of similarity-based interference and structural priming may explain the pattern we observed in the current data, although this explanation remains speculative and the assumptions we made here should be independently verified with further studies.

4 General discussion

Both Experiment 1 and 2 show that the structural properties of wh-dependency are planned before the wh-phrase is spoken, as the early commitment hypothesis predicts. In Experiment 1 in English, the slow-down effect caused by the conflict between the *that*-priming and the *that*-trace constraint was observed in onset latency, replicating Momma (2021). This suggests that speakers already plan the grammatical function of the filler and the complementizer of the gap-containing clause before starting to speak the wh-filler, across different task environments. In Experiment 2, we use the structural priming of wh-scopes in Japanese to make inferences about the time-course of wh-dependency planning. The results show a complicated pattern, but under our interpretation, they minimally suggest that speakers plan the scope of wh-phrases early and, assuming that the complementizer structures are critically involved in determining the scope relation, the complementizer of the clause that wh-phrase is taking scope over. For the target sentences with the embedded wh-scope, speakers were faster to start speaking when the prime sentences also had the embedded wh-scope. In contrast, for the target sentences with the matrix wh-scope, speakers were marginally slower when the prime sentences also had the matrix wh-scope. Although this pattern was not entirely predicted and deserves further investigation, we speculate that this pattern was caused by the interplay between the facilitatory effect of scope priming and the inhibitory effect of similarity-based interference. Taken together, Experiment 1 and Experiment 2 both suggest that speakers plan the complementizer structure of the clause containing the gap before starting to speak the filler. This in turn suggests some abstract similarity between how English and Japanese speakers plan wh-dependencies, despite surface differences in how such dependencies are realized.

We argue that the inhibitory effect found in the matrix scope condition was due to the similarity-based interference. Previous research suggested that the similarity-based interference arises in the process of retrieving words from memory during comprehension (see Van Dyke and McElree 2006 for an overview). The similarity-based interference in word retrieval also occurs in production. For instance, in Smith and Wheeldon (2004), the latency of sentences containing two semantically related nouns such as *the saw and the axe move down* is longer than when the two nouns are not related as in *the saw and the cat move down*, suggesting that the later-coming nouns interfered with the retrieval of the initial noun (at least when they are planned together). Thus, the similarity-based interference arises both in comprehension and production when a word similar to the retrieval target is co-present in memory. Given that the current study uses a memory-based task, it is conceivable that the retrieval of a sentence with the embedded or matrix scope can be more difficult in the presence of another sentence with the same

scope property in memory. The relevant notion of similarity here can be about the complementizer type (question particle vs. declarative complementizer), the scope relation (embedded vs. matrix), or the sentence type (wh- vs. yes-no question). Experiment 2 does not provide evidence to determine which of those properties are relevant to the similarity-based interference effect we postulated here. Nevertheless, the slow-down effect we found in the matrix scope condition may reflect the interference based on the similarity of the properties related to wh-scope.

We also speculate that the similarity-based interference arises in both the embedded and matrix scope conditions, but it is canceled out by the facilitatory effect of scope-related structural priming in the embedded scope condition. We attribute the lack of the facilitatory priming effect in the matrix scope condition to the inverse preference effect, based on the assumption that the mono-clausal wh-scope is also counted as the matrix wh-scope. That is, the matrix scope is difficult to prime because it is frequent. Although that assumption about frequency counting needs to be tested independently, the matrix wh-dependency structure in multi-clausal sentences is the same as that in the mono-clausal questions in the sense that they both involve the dependency between the wh-phrase and the question particle in the matrix clause.

Under our interpretation, the current results suggest that both English and Japanese speakers plan the grammatical details of wh-dependencies before starting to speak the wh-filler. This way of planning sentences involving wh-dependencies is generally congruent with the broad class of production theories that allow the generation of structural representations before selecting lexical items (e.g., Garrett 1975; see Bock and Ferreira 2014). In the current studies, we provide evidence that structural representations encoding wh-dependencies are at least to some extent planned, but we cannot tell from current results if words and structures intervening the filler and the gap are planned or not planned before the speech onset. However, Momma (2021) showed that the words intervening between the filler and the gap are likely not planned before the filler production. Momma (2021) argued that the formalism known as Tree-Adjoining Grammar (Joshi, Levy, and Takahashi 1975; Frank 2004) naturally captures this idea that the planning of sentences involving filler-gap dependencies starts with first building the non-contiguous parts of sentences (the filler and the gap). Under this view, words and structures intervening between the filler and the gap are planned later. In other words, filler-gap dependency production can still be incremental, in the sense that planning and articulation are still frequently interleaved in the production of a single sentence. Although the current experiments do not provide direct evidence for or against this view, given that speakers in current experiments took only slightly more than 1 second to start speaking, we deem it implausible that speakers in the current experiments planned all words and structures intervening the filler and the gap in details before starting to speak. Thus, the current results are naturally compatible with the view that the structural

representations of non-contiguous parts (the filler and the gap) are planned as a unit, and the words and structures intervening the filler and the gap are inserted later. This hypothesis about filler-gap dependency production can be subsumed under the view that speakers can build structural representations prior to lexical selection (e.g., Garrett 1975; see Bock and Ferreira 2014 for a recent overview). Given that the current results show some high-level parallelism between English and Japanese, this view might be applied to both English and Japanese sentence production.

Lastly, we acknowledge that the current study has the limitation that sentence recall tasks may differ from naturalistic language production processes in relevant respects. Sentence recall tasks are not widely used as a method to investigate the time-course of language production, and speakers in recall experiments may deploy planning procedures that are fundamentally different from those in naturalistic production. However, it is worth noting that the accessibility effect on word order, the effect that is usually assumed to arise from the temporal dynamics of sentence planning, can be observed in recall-based experiments (e.g., Bock and Irwin 1980; McDonald, Bock, and Kelly 1993; Tanaka et al. 2011), as in naturalistic production (e.g., Kempen and Harbusch 2011). Also, in our lab, several lines of study show that the time-course of verb planning is similar between recall-based experiments and picture-description experiments (e.g., Momma and Yoshida 2021). Thus, we assume that the time-course of sentence-recall mirrors the time-course of naturalistic sentence production as a reasonable starting point, although of course this assumption should be evaluated further. Finally, we also acknowledge that the results of Experiment 2 have an alternative interpretation. For example, it may be that speakers were simply slower to start speaking after reading and memorizing a matrix scope prime sentence (that is, after reading a match prime sentence in the matrix scope condition and after reading a mismatch prime sentence in the embedded scope condition), perhaps because matrix scope sentences are more complex than embedded scope sentences. This possibility cannot be ruled out in the current study, but future studies should examine the relationship between the complexity of prime sentences and the production latency of the target sentence production in the current task. If this interpretation is correct, more complex prime sentences should increase the onset latency of subsequent target production.

5 Conclusion

The current study shows the grammatical details of wh-dependencies are predominantly planned before the utterance of the sentence-initial wh-phrases both in English and Japanese, in accordance with the early commitment hypothesis. Of

course, the current study does not show that the planning processes involved in wh-dependency formation in production are identical between English and Japanese. However, English and Japanese sentence production may plausibly involve similar planning mechanisms for formulating wh-dependencies, despite the surface differences in how wh-dependencies are realized in the two languages. Specifically, wh-dependency formation in English and Japanese may both involve planning the complementizer structure before producing the filler.

References

Aoshima, Sachiko, Colin Phillips, & Amy Weinberg. 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51. 23–54.

Allum, Paul H. & Linda R. Wheeldon. 2007. Planning scope in spoken sentence production: the role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(4). 791.

Allum, Paul H. & Linda Wheeldon. 2009. Scope of lexical access in spoken sentence production: Implications for the conceptual–syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(5). 1240.

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1). 1–48.

Barr, Dale J., Roger Levy, Christoph Scheepers, & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.

Bernolet, Sarah & Robert Hartsuiker. 2010. Does verb bias modulate syntactic priming? *Cognition* 114(3). 455–461.

Bock, Kathryn J. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18(3). 355–387.

Bock, Kathryn & Victor S. Ferreira. 2014. Syntactically speaking. In Matthew Goldrick, Victor Ferreira, & Michele Miozzo (eds.), *The Oxford Handbook of Language Production*, 21–46. Oxford: Oxford University Press.

Bock, Kathryn J. & David E. Irwin. 1980. Syntactic effects of information availability in sentence production. *Journal of Verbal Learning and Verbal Behavior* 19(4). 467–484.

Brown-Schmidt, Sarah & Agnieszka E. Konopka. 2008. Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition* 109(2). 274–280.

Brown-Schmidt, Sarah & Michael K. Tanenhaus. 2006. Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language* 54(4). 592–609.

Chomsky, Noam. 1957. *Syntactic Structure*. The Hague: Morton.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.

De Smedt, Koenraad. 1990. IPF: An incremental parallel formulator. In Robert Dale, Chris Mellish, & Michael Zock (eds.), *Current Research in Natural Language Generation,* 167–192. Academic Press.

Ferreira, Victor S. 1996. Is it better to give than to donate? syntactic flexibility in language production. *Journal of Memory and Language* 35(5). 724–755.
Ferreira, Victor S. 2003. The persistence of optional complementizer production: Why saying "that" is not saying "that" at all. *Journal of Memory and Language* 48. 379–398.
Fodor Janet D. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9. 427–473.
Frank, Robert. 2004. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.
Frazier, Lyn, Charles Clifton, & Janet Randall. 1983. Filling gaps: Decision principles and structure in sentence comprehension. *Cognition* 13(2). 187–222.
Frazier, Lyn & Charles Clifton. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4(2). 93–126.
Frazier, Lyn & Giovanni B. Flores d'Arcais. 1989. Filler driven parsing: a study of gap filling in Dutch. *Journal of Memory and Language* 28. 331–344.
Garrett, Merrill F. 1975. The analysis of sentence production. In G. H. Bower (ed.), *Psychology of Learning and Motivation*, Vol. 9, 133–177. New York: Academic Press.
Garnsey Susan M., Michael K. Tanenhaus, & Robert M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18. 51–60. 10.1007/BF01069046
Griffin, Zenzi M. 2001. Gaze durations during speech reflect word selection and phonological encoding. *Cognition* 82(1). B1–B14.
Jaeger, T. Florian & Neal E. Snider. 2008. Implicit learning and syntactic persistence: Surprisal and cumulativity. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, Vol. 827812. Austin: Cognitive Science Society.
Joshi, Aravind K., Leon S. Levy, & Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences* 10(1). 136–163.
Kempen, Gerard & Karin Harbusch. 2011. Word order scrambling as a consequence of incremental sentence production. In Holden Härtl & Heike Tappe (eds.), *Mediating Between Concepts and Grammar*, 141–164. Berlin: De Gruyter Mouton.
Konopka, Agnieszka E. & Antje S. Meyer. 2014. Priming sentence planning. *Cognitive Psychology* 73. 1–40.
Kroch, Anthony S. & Aravind K. Joshi. 1985. The linguistic relevance of tree adjoining grammar. *Technical Reports MS-CIS-85-18*.
Kuno, Susumu. 1973. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4(3). 363–385.
Lewis, Richard L. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research* 25. 93–115.
Levelt, Willem J. 1989. *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
Mahowald, Kyle, Ariel James, Richard Futrell, & Edward Gibson. 2016. A meta-analysis of syntactic priming in language production. *Journal of Memory and Language* 91. 5–27.
McDonald, Janet L., Kathryn Bock, & Michael H. Kelly. 1993. Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology* 25(2). 188–230.
Momma, Shota. 2021. Filling the gap is gap filling: Long-distance dependency formation in sentence production. *Cognitive Psychology* 129. 104411.
Momma, Shota & Masaya Yoshida. 2021. Syntax guides sentence planning: evidence from multiple dependency constructions. Short talk at the 34th annual CUNY conference on human sentence processing. Philadelphia, PA, United States, 5–7 March.

Omaki, A, Shin Fukuda, Chizuru Nakao, & Maria Polinsky. 2020. Subextraction in Japanese and subject-object symmetry. *Natural Language and Linguistic Theory* 38. 627–669.
Omaki, Akira, Ellen F. Lau, Imogen Davidson White, Myles L. Dakan, Aaron Apple, & Colin Phillips. 2015. Hyper-active gap filling. *Frontiers in Psychology* 6. 384.
Perlmutter, David M. 1971. *Deep and Surface Structure Constraints in Syntax.* New York: Holt, Reinhart, and Winston.
Pesetsky, David. 2017. Complementizer-trace effects. In M. Everaert & H. van Riemsdijk (eds.), *The Blackwell Companion to Syntax*. 2nd edn. Blackwell.
Pickering, Martin J. & Victor S. Ferreira. 2008. Structural priming: A critical review. *Psychological Bulletin* 134(3). 427.
Pollard, Carl & Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
Potter, Mary C. & Linda Lombardi. 1998. Syntactic priming in immediate recall of sentences. *Journal of Memory and Language* 38(3). 265–282.
R Core Team 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. Accessed 22 June 2020.
Reitter, David, Frank Keller, & Johanna D. Moore. 2011. A Computational Cognitive Model of Syntactic Priming. *Cognitive Science* 35(4). 587–637.
Ritchart, Amanda, Grant Goodall, & Marc Garellek. 2016. Prosody and the that-trace effect: An experimental study. In K. Kim, P. Umbal, T. Block, Q. Chan, T. Cheng, K. Finney, M. Katz, S. Nickel-Thompson, & L. Shorten (Eds.), *33rd West Coast Conference on Formal Linguistics*, 320–328. Somerville: Cascadilla Proceedings Project.
Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Cambridge: Massachusetts Institute of Technology dissertation.
Schriefers, Herbert, Encarna Teruel, & Raik-Michael Meinshausen. 1998. Producing simple sentences: Results from picture–word interference experiments. *Journal of Memory and Language* 39(4). 609.
Segaert, Katrien, Linda Wheeldon, & Peter Hagoort. 2016. Unifying structural priming effects on syntactic choices and timing of sentence generation. *Journal of Memory and Language* 91. 59–80.
Smith, Mark & Linda Wheeldon. 1999. High level processing scope in spoken sentence production. *Cognition* 73(3). 205–246.
Smith, Mark & Linda Wheeldon. 2004. Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30(3). 675–686.
Tanaka, Mikihiro N., Holly P. Branigan, Janet F. McLean, & Martin J. Pickering. 2011. Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language* 65(3). 318–330.
Van Dyke, Julie A. & Brian McElree. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language* 55(2). 157–166.
Wanner, Eric & Michael P. Maratsos. 1978. An ATN approach to comprehension. In Morris Hale, Joan Bresnan, & George A. Miller (eds.), *Linguistic Theory and Psychological Reality*. Cambridge: MIT Press.
Wheeldon, Linda & Mark Smith. 2003. Phrase structure priming: A short-lived effect. *Language and Cognitive Processes* 18(4). 431–442.
Wheeldon, Linda, Natalie Ohlson, Aimee Ashby, & Sophie Gator. 2013. Lexical availability and grammatical encoding scope during spoken sentence production. *Quarterly journal of Experimental Psychology* 66(8). 1653–1673.
Zehr, Jeremy & Florian Schwarz. 2018. PennController for Internet Based Experiments (IBEX).