Filling the gap in gap-filling: Long-distance dependency processing in sentence production

Shota Momma & Victor S. Ferriera (UCSD)
shotam2@gmail.com

In long-distance dependency processing, it is widely known that comprehenders fill the gap before the gap becomes apparent in the bottom-up input (Crain & Fodor, 1985; Stowe, 1986) and even before encountering gap-hosting verbs (Omaki et al., 2016). In contrast to the gap-filling process in sentence comprehension, the gap-filling process in sentence production is poorly understood. Here we show that speakers plan the gap before the production onset of wh-questions, and that this gap-planning precedes the planning of gap-hosting verbs. We suggest that long-distance dependency processing in comprehension and production share the same active, top-down structure building process.

Design: This experiment took advantage of the *that-trace* constraint (Perlmutter, 1971). This constraint prohibits subject extraction from an embedded clause with an overt complementizer *that*. e.g., *Who does John think (*that) is kissing Mary?* is ungrammatical only when *that* is present. In contrast, this constraint does not apply to object extractions: *Who does John think (that) Mary is kissing?* The that-trace constraint is cross-linguistically observed (Pesetsky, 2016) and is robust across individuals (Ritchart et al., 2016). The main purpose of this study is to exploit the that-trace constraint to probe the timing of gap planning. On each trial, separated by filler trials, participants (n = 48) were first syntactically primed for the presence or absence of *that* (Ferreira, 2003), and then described picture stimuli, e.g. Fig. 1, Fig. 2. The participants were pre-trained to produce sentences like the following as a response to the picture stimuli:

(1) *Who does the artist think (*that) is chasing the ballerina.* [Fig. 1: subject extraction]

(2) *Who does the artist think (that) the chef is chasing.* [Fig. 2: object extraction]

When speakers were primed with *that*, they were expected to experience processing difficulty in the subject extraction sentences. This is because the that-priming conflicts with the that-trace constraint. Critically, the timing at which this "adverse that-priming effect" should correspond to the latest time point by which speakers plan the gap (and *that*). This is because, without planning the gap, speakers should not be able to know whether the that-trace constraint is relevant to the sentence they are producing. Independently, distractor verbs that are either semantically related to the verbs (e.g., *stalk*), or unrelated to the verbs (e.g., *kick*, relatedness was normed according to Latent Semantic Analysis, Landauer & Dumais, 1997) were superimposed on the pictures (Fig. 3). Comparing the related with the unrelated distractor conditions, we expected to observe the semantic interference effect (Schriefers et al. 1990). Importantly, our main interest was *when* the semantic interference effect is observed. We localized the adverse *that*-priming and semantic interference effects to a specific word in sentence production using a text-to-speech alignment algorithm (Yuan & Lieberman, 2008; Momma et al., 2017). We reasoned that, by comparing the relative timing of the adverse that-priming effect and the semantic interference effect, we can estimate the relative timing at which gaps vs. gap-hosting verbs are planned. **Results:** All the analyses we report here were pre-specified according to the the minimally different experiment that preceded this experiment (n = 30), the results of which were consistent with the current results. On filler trials, speakers produced *that* more when primed with sentences with *that* than without (p = .01). Speakers also almost never violated the that-trace constraint (~3% of trials, produced by two participants). Critically, in the *that*-primed condition, speakers were slower to start speaking the subject extraction sentences, but not object extraction sentences (Fig. 4, interaction p < .05). In contrast, in related verb distractor conditions, speakers elongated the production of *is* in both subject and object extraction sentences (*ps* < *.05*). This suggests that speakers planned the gap before the sentence onset, and planned the verbs as they were producing the pre-verbal words (at least when they produce subject questions). Average production time in each region is summarized in Table 1. **Conclusion:** Speakers plan the gap early on in sentence production, before the sentence onset and before planning the gap-hosting verbs, at least in subject extraction sentences. This time-course of planning is consistent with the view that speakers as well as comprehenders plan the gap prior to verbs in an active, top-down fashion.

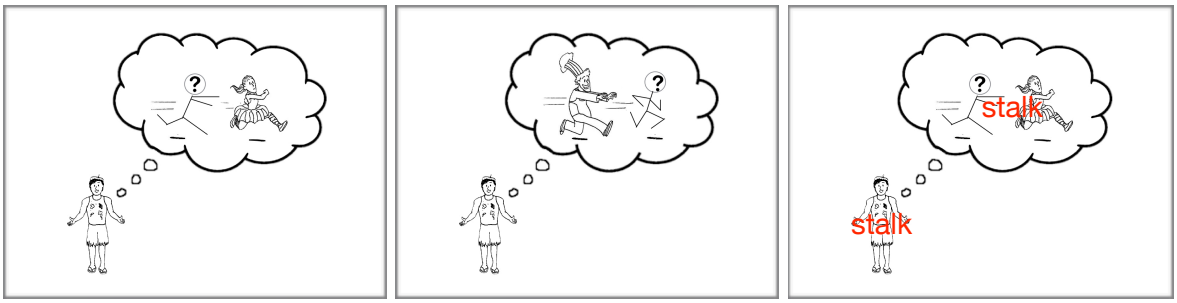

Fig. 1, Fig 2, Fig. 3. The sample picture stimuli used. With proper pre-training, the left picture elicited: *Who does the artist think is chasing the ballerina* (subject extraction). The middle picture elicited: *Who does the artist think (that) the chef is chasing* (object extraction). The right picture show the sample picture stimuli with the superimposed verb distractor (in the related distractor condition).

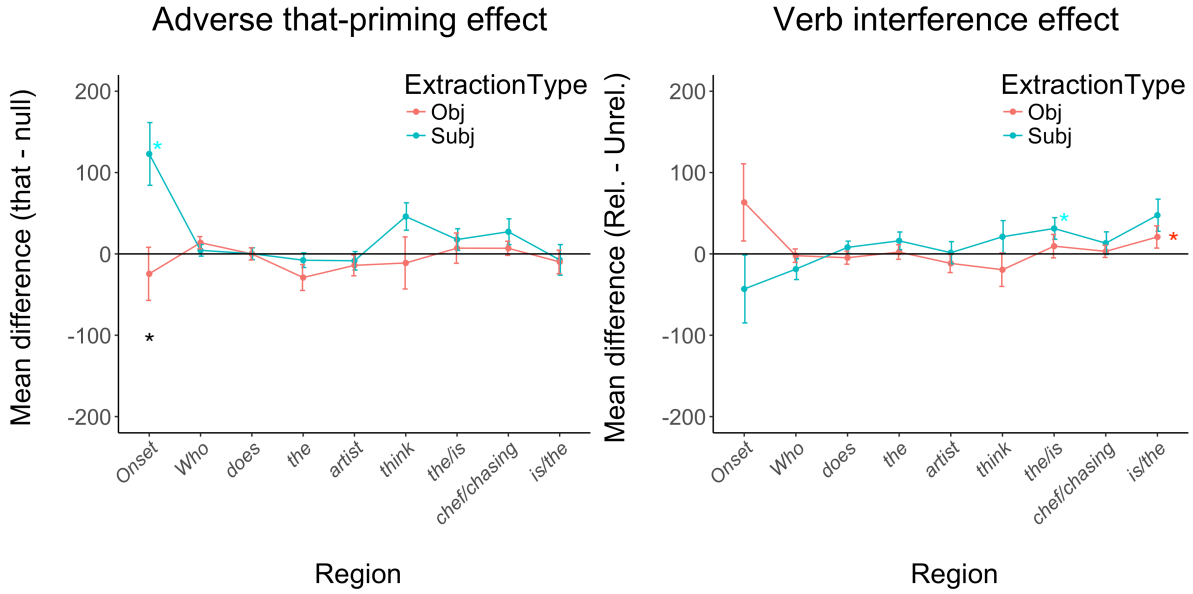

Fig. 4 (left): The adverse *that*-priming effect in the onset latency + word-by-word production duration (difference plot). Fig. 5 (right): The semantic interference effect in the onset latency + word-by-word production duration (difference plot). Black asterisks = significant interaction effect.

Condition	Onset	Who	does	the	artist	think	the/is	chef/ chasing	is/the
Subject extraction									
That. Related	1970 [41]	238 [8]	312 [7]	166 [12]	574 [13]	583 [22]	318 [20]	637 [15]	259 [21]
That. Unrelated	2009 [40]	252 [11]	295 [7]	139 [9]	562 [12]	615 [20]	285 [12]	651 [20]	182 [14]
Null. Related	1843 [42]	228 [10]	301 [8]	167 [13]	570 [12]	587 [20]	296 [17]	640 [20]	236 [25]
Null. Unrelated	1877 [42]	254 [13]	300 [9]	158 [12]	583 [13]	527 [24]	265 [15]	599 [15]	220 [23]
Object extraction									
That. Related	2024 [41]	247 [10]	308 [7]	164 [13]	580 [13]	619 [22]	247 [18]	532 [12]	266 [18]
That. Unrelated	1931 [45]	243 [8]	297 [6]	154 [10]	560 [13]	639 [37]	238 [14]	507 [10]	252 [17]
Null. Related	2013 [36]	227 [7]	294 [7]	190 [20]	570 [12]	646 [25]	239 [24]	503 [13]	292 [19]
Null. Unrelated	1979 [41]	237 [7]	310 [10]	187 [11]	606 [14]	649 [24]	230 [16]	526 [11]	260 [19]

Table 1: The region-by-region mean production time [SEM].