

# Supplemental materials for “Shared syntax between comprehension and production: Multi-paradigm evidence that resumptive pronouns hinder, not facilitate, comprehension”

Adam M. Morgan, Titus von der Malsburg, Victor S. Ferreira, Eva Wittenberg

May 2020

## 1 Fillers in Experiments 2 and 3

We included 60 fillers that were specifically designed to deter subjects from developing heuristic parsing or response strategies (e.g., ‘the first animal in the sentence is always the one that pronoun refers to’). These consisted of five sets of twelve items. Examples of each of these types and the types appear in Table 1, and the types of parsing/response strategies they were meant to prevent are summarized in Table 2. We designed the fillers to match the critical items in several ways so that the critical items would not stand out to participants. Each filler type included two embedded clauses which were comprised of roughly equal numbers of non-islands, weak islands, and strong islands. All fillers began a wh-dependency with a clefted animal character and the dependency ended with a gap in the middle clause. Note that because there were no gaps in the lowest clause (i.e., inside the island), there were no island violations in the fillers.

We controlled for many of the same features here as we did for the critical items. Similarly, the pattern of genders across the characters within a given filler type remained the same. For instance, every filler of Type B had three characters: feminine, masculine, feminine (in that order) or masculine, feminine, masculine (in that order). Within each filler type, half of the twelve items began with a feminine character and half with a masculine character. Also as with critical items, within each type of filler root clause tense was half present, half past; subordinators were “that” for half of the stimuli and “who” for the other half; and all clause boundaries were marked with an overt subordinator. All logical possible combinations of these features appeared a roughly equal number of times across each filler type.

Table 1: Experiment 2 fillers.

Filler label	Sample stimulus
Type A	<p>It is Mr. Dino who Mr. Bear asked _ if Miss Cat immobilized Miss Rabbit with a rope.</p> <p>Who did what to whom?</p> <p>a) Miss Cat immobilized Miss Rabbit with a rope. (✓)</p> <p>b) Mr. Bear immobilized Miss Rabbit with a rope.</p> <p>c) Mr. Bear immobilized Miss Duckie with a rope.</p> <p>d) Miss Cat immobilized Miss Duckie with a rope.</p>
Type B	<p>It was Miss Rabbit that Mr. Bear told _ that he jabbed Miss Piggy with a fork.</p> <p>Who did what to whom?</p> <p>a) Mr. Bear jabbed Miss Piggy with a fork. (✓)</p> <p>b) Mr. Bear jabbed Miss Cat with a fork.</p> <p>c) Miss Duckie jabbed Miss Piggy with a fork.</p> <p>d) Miss Duckie jabbed Miss Cat with a fork.</p>
Type C	<p>It is Miss Rabbit who _ informed Mr. Dino that she whacked Mr. Dog with a bottle.</p> <p>Who did what to whom?</p> <p>a) Miss Rabbit whacked Mr. Dog with a bottle. (✓)</p> <p>b) Miss Rabbit whacked Mr. Froggy with a bottle.</p> <p>c) Mr. Dino whacked Mr. Dog with a bottle.</p> <p>d) Mr. Dino whacked Mr. Froggy with a bottle.</p>
Type D	<p>It was Mister Bear who _ saw Miss Duckie when he hit Mr. Froggy with a rock.</p> <p>Who did what to whom?</p> <p>a) Mr. Dog hit Mr. Froggy with a rock. (✓)</p> <p>b) Mr. Dog hit Miss Piggy with a rock.</p> <p>c) Miss Duckie hit Mr. Froggy with a rock.</p> <p>d) Miss Duckie hit Miss Piggy with a rock.</p>
Type E	<p>It was Miss Cat that _ understood why Mr. Dino swore that Mr. Froggy cleaned him with a loofa.</p> <p>Who did what to whom?</p> <p>a) Mr. Froggy cleaned Mr. Dino with a loofa. (✓)</p> <p>b) Mr. Froggy cleaned Ms. Cat with a loofa.</p> <p>c) Miss Cat cleaned Mr. Dino with a loofa.</p> <p>d) Miss Cat cleaned Mr. Dog with a loofa.</p>









*Note.* Gaps (underscores) were not shown to participants.

Table 2: Summary of the possible comprehension heuristics that the fillers were designed to prevent. The specific filler types that were designed to prevent each strategy are listed in the right hand column.

Concern to be addressed	Solution
All pronouns in the critical stimuli are resumptive pronouns, which means they all refer to the first animal character in the sentence (i.e. the head noun). Participants may develop a strategy of interpreting all pronouns as referring to the first animal, which would allow participants to correctly interpret resumptive pronouns without parsing them.	We designed fillers that contain pronouns with referents in various positions so as to diversify the types of pronouns in the experiment. Some referred to the subject of middle clause (Types B & E), others to the high subject (Type C), and some even to extra-sentential referents (Type D).
Similarly, if all pronouns in accusative case (i.e., <i>her/him</i> , as opposed to <i>she/he</i> ) are resumptive pronouns, participants may come to rely on case cues to compensate for any difficulties with parsing resumptive pronouns.	Pronouns in Type E fillers have accusative case and are not resumptive pronouns.
If all pronouns in the study have referents within the sentence, participants may learn to disregard the dangle multiple choice option for critical items.	The answer choices for Type D fillers require participants to settle on an extra-sentential referent for a pronoun.
Participants may develop task-specific parsing strategies if they only ever encounter gaps in one position.	Fillers contain gaps in various positions: direct object of middle clause (Types A & B), subject of middle clause (Types C, D, & E), in addition to gaps in the critical items which are in the lowest object position.
Some fillers should have embedding verbs that take clausal complements so that the critical items do not stand out in this regard. (Filler Types A, B, & C take two complements each: a NP followed by a clause.)	Filler Type E contains embedding verbs which take single clausal complements.

## 2 Experiment 2 Visual Stimuli

Table 3: Animal characters for Experiment 3

Mr. Bear	Miss Cat	Mr. Dino	Mr. Dog	Miss Duckie	Mr. Froggy	Miss Piggy	Miss Rabbit
							

## 3 Experiment 2 Setup



Figure 1: Setup for Experiment 3

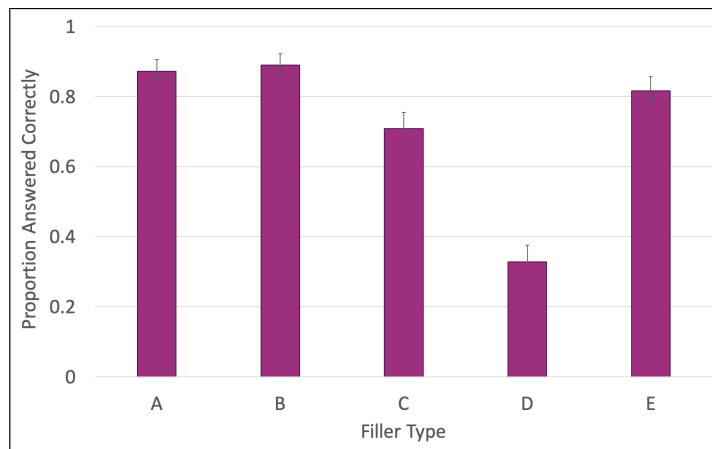


Figure 2: Experiment 2 results: Accuracy on filler trials, by filler type (see Table 1)

## 4 Remarks on Experiment 2 Filler results

Filler items all included one unambiguously correct interpretation among the four multiple choice options. Overall, accuracy on these trials was high (Figure 2), indicating that participants performed the task as intended. A notable exception is in Type D fillers, which were answered correctly only 32.8% of the time (a conservative estimate of chance for these trials would be 25%). Type D fillers were designed so as to require participants to establish a referent for the pronoun that was not present in the sentence – a *dangle* interpretation. This indicates that participants felt a strong pressure to choose a referent for the pronoun from among the characters mentioned in the sentence.

It is worth noting that the *dangle* response in our multiple choice options also represented a gender- and number-congruent referent for resumptive pronouns in critical trials, albeit one which is not mentioned in the sentence. One might then argue that, if resumptive pronouns truly led to chance performance, we should see increases in dangle interpretations as well as local interpretations relative to gap conditions. However, the fact that participants chose the correct interpretation to Type D fillers only a third of the time indicates that participants appear to feel pressure to find an intra-sentential referent for pronouns, at least in our tasks. This suggests that for critical trials, the dangle interpretation was probably not an a priori reasonable option.

## 5 Frequentist analyses

Table 4: Experiment 1 results.

	$\beta$	$z$	$p$	
Intercept (GAP, NONISLAND)	1.62	5.58	< .001	***
RESUMPTION	-0.98	-1.69	.09	.
ISLANDHOOD:WEAK	-0.6	-1.6	.11	
ISLANDHOOD:STRONG	-1.2	-3.27	.001	**
RESUMPTION $\times$ ISLANDHOOD:WEAK	0.4	0.53	.6	
RESUMPTION $\times$ ISLANDHOOD:STRONG	-0.05	-0.07	.95	

Table 5: Experiment 2 results: Multiple choice interpretation responses.

	$\beta$	$z$	$p$	
Intercept	0.81	7.49	< .001	***
RESUMPTION	-0.19	-1.54	.12	
ISLANDHOOD:WEAK	-0.14	-1.49	.14	
ISLANDHOOD:STRONG	-0.58	-5.69	< .001	***
RESUMPTION $\times$ ISLANDHOOD:WEAK	0.08	0.45	.66	
RESUMPTION $\times$ ISLANDHOOD:STRONG	-0.1	-0.57	.57	

Table 6: Experiment 2 results: Reading speed.

	$\beta$	$t$	
Intercept	3.0	39.47	*
RESUMPTION	0.28	5.34	*
ISLANDHOOD:WEAK	-0.07	-2.3	*
ISLANDHOOD:STRONG	-0.08	-2.52	*
RESUMPTION $\times$ ISLANDHOOD:WEAK	-0.02	-0.32	
RESUMPTION $\times$ ISLANDHOOD:STRONG	0.01	0.2	

*Note.* Here and below, rather than estimating  $p$ -values, we assume sufficiently high degrees of freedom to safely interpret any  $t$ -value greater than 1.96 to be significant at  $\alpha = 0.05$ .

Table 7: Experiment 3 results: Multiple choice interpretation responses.

	$\beta$	$z$	$p$	
Intercept	0.49	5.03	< .001	***
RESUMPTION	-0.31	-2.64	.008	**
ISLANDHOOD:WEAK	0.06	0.66	.51	
ISLANDHOOD:STRONG	-0.48	-5.57	< .001	***
RESUMPTION $\times$ ISLANDHOOD:WEAK	0.18	1.14	.26	
RESUMPTION $\times$ ISLANDHOOD:STRONG	0.25	1.53	.13	

Table 8: Experiment 3 results: Gaze.

	$\beta$	$z$	$p$	
Intercept	0.28	2.34	.02	*
TIME	0.07	2.36	.02	*
RESUMPTION	-0.43	-2.35	.02	*
ISLANDHOOD:WEAK	0.24	1.34	.18	
ISLANDHOOD:STRONG	-0.54	-3.44	< .001	***
TIME $\times$ RESUMPTION	-0.05	-0.77	.44	
TIME $\times$ ISLANDHOOD:WEAK	0.08	1.74	.08	.
TIME $\times$ ISLANDHOOD:STRONG	-0.06	-1.24	.22	
RESUMPTION $\times$ ISLANDHOOD:WEAK	0.04	0.13	.9	
RESUMPTION $\times$ ISLANDHOOD:STRONG	0.25	0.99	.32	
TIME $\times$ RESUMPTION $\times$ ISLAND:WEAK	-0.05	-0.55	.59	
TIME $\times$ RESUMPTION $\times$ ISLAND:STRONG	0.0	-0.05	.96	

Table 9: Experiment 4 results: Multiple choice interpretation responses.

	$\beta$	$z$	$p$	
Intercept (ORDINARY PRONOUN)	1.74	4.25	< .001	***
GAP	0.02	0.04	.97	
RESUMPTIVE PRONOUN	-1.24	-2.43	.02	*

Table 10: Experiment 5 results.

	$\beta$	$z$	$p$	
Intercept (NON-ISLAND)	-2.0	-5.94	< .001	***
ISLANDHOOD:WEAK	1.06	2.49	.01	*
ISLANDHOOD:STRONG	2.26	5.34	< .001	***