Not all filler-gap dependencies are perceived alike: Evidence from Tagalog

Jed Sam Pizarro-Guevara & Matt Wagers (University of California, Santa Cruz)
jpquevar@ucsc.edu

Overview. Comprehension of filler-gap dependencies (FGDs) proceeds predictively: a filler is linked to its gap before direct evidence of the gap’s position. In Tagalog, an Austronesian language of the Philippines, verbal agreement can signal the location of the gap site. In this study, we asked whether speakers could use that agreement to predict the identity of the gap, and whether the validity of the agreement cue affects how aggressively they do so. We tested three types of FGD and we found that agreement was indeed used to predict the gap. But it did so in a way that varied across construction, and that was not directly tied to cue validity.

Tagalog verbs can agree with either its subject or object (S- and O-AGR, respectively) [1]. In FGDs, the only argument that can be extracted is the one the verb agrees with [2], the subject under S-AGR, or the object under O-AGR. Upon processing agreement, perceivers should be able to predict the matching gap site. However, this cue’s validity is probably variable: O-AGR permits subject extraction for some speakers and for some constructions [3]. This variability led us to ask if agreement cues were exploited identically across three types of FGDs: wh-questions (WHQ), relative clauses (RC) and ay-topicalization (AY); or if they were used according to their FGD-specific validity. We used an acceptability judgment task (AJT) as a proxy for FGD-specific validity; and a stops-making-sense task (SMS) [4] to track the time-course of FGD-resolution.

Experiment 1: AJT (n = 64). To assess the extent to which subject extraction was permitted under O-AGR, we conducted an AJT that crossed agreement (S- and O-AGR) by extraction match (MATCH, MISMATCH) in 3 types of FGDs (WHQ, RC, AY). We always found a strong mismatch penalty for S-AGR and a weaker, variable one for O-AGR. Subject extractions under O-AGR were best tolerated in RCs, moderately so in AY-topicalizations, and least tolerated in WHQs (Table 1). We thus expected S-AGR to be a robust trigger for FGD formation in all environments. Meanwhile, O-AGR should trigger less predictive parsing as it becomes less diagnostic of an object gap: (most predictive) WHQ >> AY >> RC (least; Table 1 shaded cells).

Experiments 2–4: SMS (n = 85). Design and Analysis. We used SMS to compare the time-course of FGD formation under S- and O-AGR. We crossed AGREEMENT (S- and O-AGR) with the PLAUSIBILITY of the filler phrase (+/-PLAUS) in 3 experiments, each realizing a different FGD (WHQ, RC, AY; 24 item sets each). See Table 2. The experiments were combined together in one session and randomized with 24 fillers (declaratives with no FGD). In SMS, participants have the option to reject a trial at any region, which lets us define a d’ series by scaling cumulative correct rejections in -PLAUS conditions against misses in +PLAUS. See Figure 1.

Results. For S-AGRs, we found that d’ was positive at the verb and immediately before the co-arguments, in all FGDs. Participants were thus linking the filler to the correct gap site before full disambiguation. When O-AGR was a highly valid cue (WHQs), we found early discrimination; and as expected, when it was minimally valid (AY), discrimination did not emerge until after the verb. RCs confounded our predictions. We found early discrimination in O-AGR conditions, comparable to S-AGR, despite the low validity of the O-AGR cue in that FGD.

Discussion. Our results provide clear evidence of predictive FGD parsing in Tagalog and mixed evidence on how agreement is used in FGD-specific ways. Acceptability may simply be a poor proxy for O-AGR cue validity. Alternatively, structural differences between ay-topicalization and WhQs/RCs may be relevant. While WhQs are formed from null-headed RCs [5], the filler in ay-topicalization is directly related to the gap in a single dependency [6]. Ay-topicalization may thus be a simpler FGD to recognize and interpret.
### Experiment 1

**Table 1.** Mean acceptability ratings by AGREEMENT and MATCH (with S.E.). Significant interaction in each construction: 
RC: $\chi^2(1) = 9.4, p = .002$; AY: $\chi^2(1) = 23.9, p < .001$; WHQ: $\chi^2(1) = 3.9, p = .04$.

<table>
<thead>
<tr>
<th></th>
<th>AY</th>
<th>RC</th>
<th>WHQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S-Agr</td>
<td>O-Agr</td>
<td>S-Agr</td>
</tr>
<tr>
<td>Match</td>
<td>5.86 (.14)</td>
<td>5.27 (.18)</td>
<td>4.96 (.18)</td>
</tr>
<tr>
<td>Mismatch</td>
<td>2.03 (.15)</td>
<td>4.53 (.18)</td>
<td>2.24 (.14)</td>
</tr>
</tbody>
</table>

### Experiments 2–4

**Table 2.** A schematized item set. The filler is introduced in region $j$, the agreeing verb in $j+2$, an adverb in $j+3$, and the fully disambiguating co-argument in $j+4$. The $j+1$-region, 'X', corresponds to the nominative marker $ang$ in exp 2, the complementizer $na$ in exp 3, and the topic marker $ay$ in exp 4.

<table>
<thead>
<tr>
<th>REGION</th>
<th>$j$</th>
<th>$j+1$</th>
<th>$j+2$</th>
<th>$j+3$</th>
<th>$j+4$</th>
<th>$j+5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>...</td>
<td>Filler</td>
<td>X</td>
<td>Verb</td>
<td>YP</td>
<td>Co-argument</td>
</tr>
</tbody>
</table>

**Figure 1.** Mean $d'$ by voice at each region (dots) with 95% confidence intervals (bands) derived by a bootstrap over participants ($n = 10,000$).

**LEGEND**
- S-AGR plotted in Blue;
- O-AGR plotted in Yellow.

**Critical regions** (Verb + Adverb) boxed.

### REFERENCES.