Fragments in Child English and Spanish

Koji Sugisaki and William Snyder

MIE UNIVERSITY AND UNIVERSITY OF CONNECTICUT

1. Introduction: Grammatical Conservatism

There are a growing number of acquisition studies which indicate that children are “grammatically conservative”: At least in their natural, spontaneous speech, children do not begin using a new syntactic structure until they have both determined that the structure is permitted in the adult language, and identified the adults' grammatical basis for it (e.g. Sugisaki and Snyder 2003, Snyder 2007, 2008, Rodríguez-Mondoñedo 2008, Sugisaki 2010). This claim of Grammatical Conservatism (GC) is based on the observation that, when we systematically examine the transcripts of a child's spontaneous speech, it is strikingly rare to find an actual error of “co-mission,” where the child puts words or morphemes together in an ungrammatical way. Instead, the vast majority of errors are errors of omission, where required words or morphemes are simply omitted from the child's utterance. Thus, in their spontaneous speech, children appear to reserve judgment on points of grammatical variation, and refrain from actually putting elements together in ways that might turn out to be disallowed in the target language.

The goal of this study is to evaluate the claim of children's GC in a new syntactic domain: children's short, fragmentary answers to prepositional *wh*-questions in English and Spanish. The results not only provide a new piece of evidence for children's GC, but also suggest that there is a case in which GC overrides the general tendency to make omission errors.
2. Grammatical Conservatism with English Particles


(1)  
   a. Mary stood up.  
   b. Mary lifted the box up.  
   c. Mary lifted it up.  
   d. Mary lifted up the box.

The child must rely on the input, at least in part, to determine that the examples in (1) are permitted in English, because languages like French, Russian, and Swahili lack any such possibility for a directional particle (like *up*) to be semantically connected to a verb (like *lift*) yet surface as an independent word.

A few of the logically possible error-types that a child could make when acquiring this syntactic property are provided in (2).

(2)  
   a. * Mary lifted up it.  
   b. * Mary lifted up the box out.  
   c. * Mary lift up+ed the box.  
   d. * Mary will up+lift the box.

For example, the error in (2a) is quite likely to occur if a child is reasoning by analogy from (1b-d). The error in (2b) could also result from simple distributional reasoning, if the child surmises from (1b) and (1d) that English provides two independent positions for directional particles. The error in (2c) could result if the child were swayed by the close semantic connection between *lift* and *up*, as well as their frequent occurrence side by side, and incorrectly concluded that they constitute a single word.

The error in (2d) might have a different source: It involves a prefixed
particle, which is ungrammatical in English but fully grammatical in other West Germanic languages. In other words, the child could be led to (2d) because it corresponds to an option of Universal Grammar (UG), although it happens to be the wrong option for English. Thus, whether the child is reasoning by analogy (as in certain domain-general approaches to language acquisition), or trying out the options compatible with UG (as expected in many Chomskyan approaches to acquisition), the opportunities for co-mission errors are ample.

In order to determine whether English-learning children actually produce co-mission errors of the kinds in (2) and others, Snyder (2007) conducted a near-exhaustive search for errors with the English verb-particle construction in the longitudinal corpus for Sarah (Brown 1973) that is available in the CHILDES database (MacWhinney 2000). This corpus includes over 37,000 child utterances, covering an age span from 2(years);03(months) to 5;01. The results were striking: Sarah made almost no co-mission errors. From the beginning of her corpus through the age of 2;10, Sarah produced 102 examples of verb-particle constructions, of which 70 were correct from the standpoint of adult English. Of the remaining 32, at least 29 of them (90.6%) were errors of omission. Of the other three, only one was unambiguously a grammatical error, which is provided in (3). Thus, the findings from the acquisition of English verb-particle construction provide us with clear evidence for GC in children's spontaneous speech.

(3) I […] go down+ed. [Transcript 34, line 569, age 2;10:20]

In light of this background, we now turn to the question of whether children's GC holds for a different area of syntax: short, fragmentary answers to prepositional questions.
3. Cross-linguistic Variation in P-questions and Fragment Answers

It is widely known that languages differ with respect to the movement possibilities for prepositional complements (e.g. van Riemsdijk 1978, Hornstein and Weinberg 1981, Kayne 1981, Abels 2003). For example, in everyday spoken English, *wh*-movement of a prepositional complement strands the preposition, while in Romance languages like Spanish, the preposition has to be “pied-piped” along with the *wh*-word.

(4) English: Preposition-stranding (P-stranding) possible

a. Who was Peter talking with \( t \) ?

b. ?? With whom was Peter talking \( t \) ? [Odd, in spoken English]

(5) Spanish: P-stranding impossible / Pied-piping obligatory

a. *Quién hablaba Pedro con \( t \) ?
   
   *who was-talking Peter with*

b. Con quién hablaba Pedro \( t \) ?
   
   *with who(m) was-talking Peter*

A study by Merchant (2004) revealed that the same point of cross-linguistic variation has an effect on fragment answers to prepositional *wh*-questions (“P-questions”). Fragment answers are answers to *wh*-questions which consist of a non-sentential XP like the (a) examples in (6) and (7), which nevertheless convey the same propositional content as a fully sentential answer like the (b) examples.

(6) Who did she see?


b. She saw John.
(7) When did he leave?
   a. After the movie ended.
   b. He left after the movie ended.

According to Merchant (2004), when the wh-phrase in a question is a complement of a preposition, the corresponding fragment answer can be either a ‘bare’ DP or (at least marginally) a PP in English, as shown in (8). In sharp contrast, only PP answers are permitted in Spanish, as illustrated in (9).

(8) English: Who was Peter talking with?
   a. Mary.
   b. ?? With Mary.

(9) Spanish: Con quién hablaba Pedro?
   a. *María.
   b. Con María.

This contrast between English and Spanish is part of a larger cross-linguistic generalization that Merchant (2004) has found: Bare-DP answers to P-questions are found only in those languages that permit P-stranding. The results of Merchant's cross-linguistic survey are summarized in (10).

In order to account for this cross-linguistic correlation between P-stranding on the one hand and bare-DP answers to P-questions on the other, Merchant (2004) argues that fragment answers have fully sentential syntactic structures: These answers are derived by A'-movement of the fragment to the clause-peripheral position in the syntactic component, followed by an operation of ellipsis at PF, as illustrated in (11). Since fragment answers involve leftward A'-movement under this analysis, the grammatical constraints governing P-stranding will be operative in these structures as well. Thus, the ban on
bare-DP answers to P-questions in Spanish is immediately accounted for.


<table>
<thead>
<tr>
<th>Language</th>
<th>P-stranding?</th>
<th>Bare-DP Answer?</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Swedish</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Norwegian</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Danish</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Icelandic</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Greek</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>German</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Yiddish</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Czech</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Russian</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bulgarian</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Hebrew</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

(11) a. Derivation of (8a):

a. In syntax: \[ \text{Mary} \ \text{TP Peter was talking with} \ t \]
b. In PF: \[ \text{Mary} \ \{\text{TP Peter was talking with} \ t \} \]

b. Derivation of (8b):

a. In syntax: \[ \text{with Mary} \ \text{TP Peter was talking} \ t \]
b. In PF: \[ \text{with Mary} \ \{\text{TP Peter was talking} \ t \} \]

To summarize this section, Merchant (2004) observes that fragment answers to P-questions are subject to cross-linguistic variation: While English permits bare-DP answers (and PP answers, to a certain extent), Spanish disallows bare-DP answers. In order to account for this observation, Merchant proposed an analysis in which fragment answers are derived from fully sentential structures through the combination of movement and ellipsis.
4. Grammatical Conservatism with English and Spanish P-questions

In Sugisaki and Snyder (2003), we examined children's acquisition of P-questions in English and Spanish (as in (4) and (5), repeated here as (12) and (13)), and found that children's wh-questions conform to the pattern in the target language as soon as children begin producing them.

(12) English:
   a. Who was Peter talking with t ?
   b. ?? With whom was Peter talking t ? [Odd, in spoken English]

(13) Spanish:
   a. *Quién hablaba Pedro con t ?
      who was-talking Peter with
   b. Con quién hablaba Pedro t ?
      with who(m) was-talking Peter

Our analysis of ten English corpora and four Spanish corpora revealed that (i) no child acquiring English ever used pied-piping of prepositions when asking a P-question, and (ii) no child acquiring Spanish ever used P-stranding. Furthermore, each of four children acquiring English (Abe, Eve, Naomi, and Shem) had a statistically significant gap between the point by which she was clearly using both PP complements and direct-object wh-questions, on the one hand, and the point at which she began to use P-questions, as shown in (14). (For all of the children, PP complements appeared earlier than direct-object questions.) Most probably, these four children were actively avoiding P-questions in their spontaneous speech, for a period of up to nine months.
<table>
<thead>
<tr>
<th>child</th>
<th>direct-object question</th>
<th>P-stranding</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
<td>2:05:00</td>
<td>2:07:07</td>
<td>14(.583)11=.037</td>
</tr>
<tr>
<td>Eve</td>
<td>1:08:00</td>
<td>2:02:00</td>
<td>14(.818)48=.009</td>
</tr>
<tr>
<td>Naomi</td>
<td>1:11:30</td>
<td>2:08:30</td>
<td>14(.833)42=.007</td>
</tr>
<tr>
<td>Shem</td>
<td>2:02:16</td>
<td>2:06:06</td>
<td>14(.714)18=.033</td>
</tr>
</tbody>
</table>

5. **Grammatical Conservatism with English and Spanish Fragments**

A very strong test of children's GC is provided by fragment answers to P-questions. In contrast to prepositional questions, where the P-stranding and pied-piping versions contain exactly the same number of words, the bare-DP answer requires one fewer word than the P+DP answer. If the Spanish-learning child nonetheless favors P+DP answers, then she is actually performing extra labor (articulating a longer phonetic string) in order to achieve GC.

To see whether this strong form of GC is indeed operative, we examined spontaneous-speech data from English-learning and Spanish-learning children. The details are reported below.

5.1. **Subjects and Method**

We analyzed spontaneous-speech data from CHILDES (MacWhinney 2000) for five children acquiring English, and five children acquiring Spanish. The corpora we examined are summarized in (15) and (16).

In all cases we began at the point when the child's mean length of utterance (MLU) first reached 2.50 words, to ensure that the child had left the telegraphic stage behind, and could in principle produce both DP (which consists of two words in many cases, e.g. my mother) and P+DP (which consists of three words in many cases, e.g. with my mother) utterances. The CLAN program Combo, together with lists of prepositions and wh-words in English and Spanish, was used to locate every non-child utterance that could possibly have contained a
prepositional question, together with the two utterances that immediately followed. The output was searched by hand to locate all of the child's fragment answers to prepositional questions. Results were checked against the original transcripts to exclude imitations, repetitions, and formulaic routines.

(15) English Corpora Analyzed:

<table>
<thead>
<tr>
<th>Child</th>
<th>Collected by</th>
<th>Ages</th>
<th># of utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe</td>
<td>Kuczaj (1976)</td>
<td>2;04:24 - 2;08:18</td>
<td>3,110</td>
</tr>
<tr>
<td>Adam</td>
<td>Brown (1973)</td>
<td>2;03:04 - 2;07:00</td>
<td>9,254</td>
</tr>
<tr>
<td>Eve</td>
<td>Brown (1973)</td>
<td>1;06:00 - 2;03:00</td>
<td>12,473</td>
</tr>
<tr>
<td>Nina</td>
<td>Suppes (1973)</td>
<td>1;11:16 - 3;00:03</td>
<td>23,586</td>
</tr>
<tr>
<td>Sarah</td>
<td>Brown (1973)</td>
<td>2;03:05 - 3;05:13</td>
<td>17,881</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td><strong>66,304</strong></td>
</tr>
</tbody>
</table>

(16) Spanish Corpora Analyzed:

<table>
<thead>
<tr>
<th>Child</th>
<th>Collected by</th>
<th>Ages</th>
<th># of utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emilio</td>
<td>Serrat &amp; Vila</td>
<td>0;11 - 4;08</td>
<td>7,126</td>
</tr>
<tr>
<td>Irene</td>
<td>Ojea and Llinas-Grau</td>
<td>0;11 - 3;02</td>
<td>12,055</td>
</tr>
<tr>
<td>Koki</td>
<td>Rosa Montes</td>
<td>1;07 - 2;11</td>
<td>4,548</td>
</tr>
<tr>
<td>María</td>
<td>Susana López-Ornat</td>
<td>1;07 - 3;11</td>
<td>8,433</td>
</tr>
<tr>
<td>Magín</td>
<td>Carmen Aguirre</td>
<td>1;07 - 2;10</td>
<td>10,916</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td><strong>43,168</strong></td>
</tr>
</tbody>
</table>

5.2. Results

The results are summarized in (17) and (18). Remarkably, the children acquiring Spanish overwhelmingly used adult-like P+DP answers from the outset. Beginning at the child's first fragment answer to a prepositional question, the next five transcripts contained an average of 93.6% P+DP answers to the prepositional questions that they answered with a fragment (range: 83%-100%). In contrast, in the corresponding first five transcripts for each of the English-learning children, both of the options available to adults were employed: On average, children produced a P+DP answer, rather than a bare-DP
answer, to only 40.2% of the prepositional questions that they answered with a fragment (Range: 20%-67%). The Spanish-English contrast appears to be statistically reliable ($t(8)=5.88$, two-tailed $p<.001$).

(17) Results from Child English:

<table>
<thead>
<tr>
<th>Child</th>
<th>Types of Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Abe</td>
<td>1 (50%)</td>
</tr>
<tr>
<td>Adam</td>
<td>3 (75%)</td>
</tr>
<tr>
<td>Eve</td>
<td>1 (33%)</td>
</tr>
<tr>
<td>Nina</td>
<td>11 (61%)</td>
</tr>
<tr>
<td>Sarah</td>
<td>4 (80%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>20</td>
</tr>
</tbody>
</table>

(18) Results from Child Spanish:

<table>
<thead>
<tr>
<th>Child</th>
<th>Types of Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DP</td>
</tr>
<tr>
<td>Emilio</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Irene</td>
<td>1 (7%)</td>
</tr>
<tr>
<td>Koki</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>María</td>
<td>3 (8%)</td>
</tr>
<tr>
<td>Magín</td>
<td>1 (17%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>5</td>
</tr>
</tbody>
</table>

These results from child English and Spanish demonstrate that children's GC is operative even in the domain of fragment answers. Furthermore, in this case GC overrides the general tendency to make omission errors: Spanish-learning children overwhelmingly favored P+DP answers over bare-DP answers, which in turn suggests that they are actually performing extra labor (articulating a longer phonetic string) in order to achieve GC.

One final issue needs to be addressed. As discussed above, when English-learning children begin to produce full, sentential P-questions, these invariably have P-stranding, not pied-piping. Yet, the children's earliest fragment
answers are often PPs, not bare DPs. This is surprising under Merchant's analysis, because for him PP fragments are derived via pied-piping. If (as seems likely) the young child's grammar does not yet permit any pied-piping at all, how can it permit PP fragments?

A possible answer is suggested by a recent, wh-in-situ analysis of "sluicing" (i.e., elliptical questions, as in John was talking to somebody, but I don't know who) proposed by Kimura (2010). Rather than moving the wh-phrase to the left edge of the clause and deleting an XP to its right, which is Merchant's approach, Kimura proposes that the wh-phrase remains in situ, and that all the [-wh] material in the clause is simply deleted. The English-Spanish contrast with respect to P-stranding / pied-piping in P-questions would be derived from feature percolation (Kimura 2010:56, fn.8): In languages like Spanish, the wh-feature obligatorily percolates up to PP, yielding obligatory pied-piping in sentential questions, and a full PP in sluices.

Kimura's analysis can accommodate fragment answers to P-questions if either the prepositional object or the whole PP in such answers bears a special feature, let's say [+Focus], that protects it from deletion. The feature-percolation requirement in languages like Spanish would then need to cover [+Focus] as well as [+wh]. Instead of undergoing A-bar movement, as in Merchant's account, the PP/DP could remain in situ.

If we adopt this set of proposals, the benefit is that English-learning children's early PP-fragment answers do not require actual movement with pied-piping. As long as the child's grammar permits optional feature-percolation up to PP, then whenever the [+Focus] feature percolates and the [-Focus] material gets deleted, a PP-fragment will remain.

6. Conclusion

In this study, we demonstrated that children's fragment answers to prepositional wh-questions conform to the pattern predicted by the GC
hypothesis: While English-learning children produced both bare-DP and PP answers, Spanish-learning children largely restricted their answers to PPs. These findings provide a new piece of evidence for GC, and thereby bolster the utility of spontaneous speech data: When a child abruptly goes from never using a surface construction to using it frequently and correctly, we are entitled to conclude that she has acquired the final grammatical (or lexical) pre-requisite for the adult construction. As a direct consequence, the longitudinal records of children's spontaneous speech become an extremely valuable testing ground for theories of cross-linguistic variation.

Finally, to the extent that children exhibit GC, this confronts theoretical linguists with a stronger version of the Logical Problem of Language Acquisition: Traditionally, linguists have needed to indicate how a learner could possibly identify the correct grammar, from among the available options, given the types of evidence that are in fact adequate for human children. Now a linguistic theory needs to explain how a grammatically conservative learner could do so. Even though we do not have a definite answer at this moment, some of the possibilities are discussed in Snyder (2007, 2008).

Notes
* We are grateful to Stephen Crain, Ken Hiraiwa, Tetsuya Sano, and the audience at TCP 2010 for valuable comments and discussion. The usual disclaimers apply. This research was supported in part by a Grant-in-Aid for Young Scientists (B) from the Japan Society for the Promotion of Science (#21720174) (Sugisaki).

1 For an experimental study concerning children's knowledge of Condition C in fragment answers (which is illustrated in (i)), see Conroy and Thornton (2005).

(i) Speaker A: Where did he₁ send the letter?
Speaker B: * To Chuckies's₁ house.
The $p$-value is calculated as follows. Abe (for example) produced 11 direct-object $wh$-questions prior to his first P-question. In transcripts slightly later than that first P-question, when Abe asked either a direct-object question or a P-question, 58.3% of the time it was a direct-object question. The probability of producing 11 or more direct-object questions before the first P-question just by chance, under the null hypothesis that P-questions were available to Abe as early as direct-object questions, and had the same likelihood of being used as in slightly later transcripts, is $p=(.583)^{11}=.00264$. Given that a total of 14 children's corpora were examined, a Bonferroni correction is appropriate: Corrected $p=14(.583)^{11}=.037$.

On our version of Kimura's approach, we may also need to assume that optional percolation of a $[+\text{wh}]$ feature to PP is strongly disfavored (for the adult as well as the child) when the result will be overt movement of a larger (PP) constituent. This would account for the general lack of pied-piping in children's (and for the most part, adults') English P-questions, while still allowing for the ready availability of PP-fragment questions and PP-fragment answers – neither of which involve movement.

A potential problem for our extension of Kimura's analysis of sluicing is the following contrast between sluicing and fragment answers with respect to their island (in)sensitivity:

(i)  
   a. They want to hire someone who speaks a Balkan language, but I don't remember which.
   b. * They want to hire someone who speaks a Balkan language, but I don't remember which (Balkan language) they want to hire someone who speaks.

(ii) Does Abby speak the same Balkan language that Ben speaks?
   a. *No, Charlie.
   b. No, she speaks the same Balkan language that Charlie speaks.
According to Merchant (2004), while island violations can be repaired in sluicing as illustrated in (ia), these effects persist even after ellipsis in the case of fragment answers, as exemplified by (iia). While the \textit{wh}-in-situ analysis of sluicing by Kimura (2010) immediately accounts for the absence of island effects in this construction, our \textit{in-situ} analysis of fragments faces the question of why fragment answers to P-questions exhibit island effects, despite the absence of movement.

A possible answer comes from a careful examination of Merchant's example (iia). Note that the ellipsis “straddles” the focused word \textit{Charlie}, in the sense that there is unpronounced material both before and after. If we eliminate this property, we see that ellipsis in a fragment answer actually can repair an island violation:

(iii) [Context: John and Bill are friends attending a wedding reception. Earlier, when Bill was not present, John saw their mutual friend Abbie speaking to a woman whom he did not recognize. Now Bill is trying to help John identify her.]

Bill: Did Abbie speak to the woman who's sitting next to \textit{Ben}?

The native English-speakers we polled accepted both versions of John's fragment answer, although some preferred the full-PP version. All speakers rejected the corresponding \textit{wh}-questions, however:

(iv) a. ** Who did Abbie speak to the woman who's sitting next to __?

b. ** Next to whom did Abbie speak to the woman who's sitting __?

We take this evidence to show that Merchant's claims in this specific area merit re-examination, and that we might in fact be correct in extending Kimura's version of his approach to ‘fragment’ answers.
References
Snyder, W. (2008) Children's grammatical conservatism: Implications for linguistic theory. In T. Sano et al. (Eds.), *An Enterprise in the Cognitive*


(sugisaki@human.mie-u.ac.jp)

(william.snyder@uconn.edu)