The Acquisition of Parametric Variation in Count Noun Modification Using Numerals: Comparing Japanese and English

Tetsuya Sano

There is cross-linguistic parametric variation in count noun modification using numerals (Chierchia 1998, Krifka 1995). In some languages (e.g., English), a numeral directly modifies a count noun (Direct Modification, DM). Other languages (e.g., Japanese) use classifiers: A numeral cannot directly modify a count noun and instead the count noun is modified with an accompanying classifier (Indirect Modification, IM). However, the question then arises of how Japanese-speaking children come to know that DM is impossible in their target language. To answer this question, I investigated whether DM overgeneration is seen in Japanese-speaking children’s spontaneous production data, and my search found no DM overgeneration. This observation can be explained if we assume that innateness plays a role in avoiding such overgeneration. To be specific, the lack of DM overgeneration in Child Japanese is consistent with Chierchia’s (1998) proposal that the Japanese setting is the default for the parameter in question.

The outline of this paper is the following. In Section 1, I introduce the issue of this paper. In Section 2, I set a concrete research question. In Section 3, I give an answer to the research question by showing the result of my data search. In Section 4, I discuss my research results and point out a further issue. In Section 5, I give a conclusion.

1. Introduction

As pointed out in Chierchia (1998) and Krifka (1995), there is cross-linguistic variation in count noun modification using numerals. Specifically, there are languages that do not use classifiers, for example English, and there are other languages that do, such as Japanese and Chinese.

In some languages (e.g., English), a numeral directly modifies a count noun. In English example (1), the numeral two modifies the count noun cups without the use of a classifier. This is called DM in this paper.

Two cups. (DM)

In other languages (e.g., Japanese), DM is impossible. In a Japanese example (2), a numeral ni ‘two’ modifies a count noun koppu ‘cup(s)’ without a classifier.

Ni (no) koppu. (DM)

Two (Gen) cup (Gen: Genitive)
ungrammatical as ‘two cups’ with or without the Genitive Particle no

The phrase in (2) is ungrammatical in Japanese as ‘two cups’, with or without the Genitive Particle no. It is grammatical as ‘cup(s) with number two’ with the Genitive Particle no. Since this latter usage is different from DM, it will not be considered in this paper.

In a Japanese-like language that disallows DM, a classifier is used when a numeral modifies a noun. In the Japanese examples in (3), the classifiers -ko and -tsu are used when a numeral modifies a noun.

Ni-ko no koppu. / Futa-tsu no koppu. (IM)

Two-Cl Gen cup / two-Cl Gen cup (Cl: Classifier)
‘two cups’

This is called Indirect Modification with a numeral (IM) in this paper.

With the cross-linguistic variation in nominal modification with a numeral as a background, I would like to address the following issue in child language acquisition: How do Japanese-speaking children come to know that DM is impossible in the target language?

2. Setting up a research question

In this section, I set up a concrete research question to address the issue introduced in Section 1. Specifically, I compare two scenarios of how Japanese-speaking children come to know that DM is impossible in the target language in order to better determine the research question of this paper.

One idea is that DM is impossible because it is the innately determined default value of the parameter in question and becomes possible only when it appears in child-directed input data. An idea along this line is suggested in Chierchia (1998). We call this “the Default scenario” in this paper. Chierchia (1998) took consideration of the Subset Principle and did not literally use the word “Default,” but I would like to acknowledge that he proposed the Japanese-type value as the unmarked setting of the parameter in question.

Another idea is that the impossibility of DM is learned by Indirect Negative Evidence, which is suggested in Chomsky (1981). We call this “the Indirect Negative Evidence scenario” in this paper. Under this scenario, DM is allowed at
first and children come to know that it is impossible by observing from the child-directed input data that the form with a classifier is always used in modification with a numeral.

In comparing the two scenarios, overgeneration of DM plays an important role. First, the Default scenario is inconsistent with overgeneration of DM in L1 Japanese acquisition. Under the Default scenario, DM is impossible at the initial stage of acquisition due to innateness and becomes possible only when it appears in the child-directed input data. Since Japanese child-directed input data involves no DM, no DM overgeneration is expected under the Default scenario.

Second, the Indirect Negative Evidence scenario is consistent with overgeneration of DM in L1 Japanese acquisition. Under this scenario, DM is allowed at first. Hence, it is expected that DMs are overgenerated until they gradually taper off.

In this way, overgeneration of DM plays an important role in comparing the two scenarios for learning that DM is impossible in the target language. Given this consideration, the research question of this paper is as follows: Is there DM overgeneration in L1 Japanese acquisition?

3. Data

This section presents the results of the data search I performed to see if there is DM overgeneration in Child Japanese. For this purpose, several corpora in the CHILDES database (MacWhinney 2000) were used.

First, I examined the emergence of DM in Child English for comparison. Table 1 presents the results of my search of the CHILDES database for three English-speaking children at age 1–2. The names of the three children are: Eve (Brown 1973), Peter (Bloom, Hood, & Lightbown 1974, Bloom, Lightbown, & Hood 1975), and Adam (Brown 1973).

<table>
<thead>
<tr>
<th>Child name</th>
<th>Eve</th>
<th>Peter</th>
<th>Adam</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1;6–2;3</td>
<td>1;9–2;10</td>
<td>2;3–2;11</td>
<td>1;6–2;11</td>
</tr>
<tr>
<td>Total utterances</td>
<td>11868</td>
<td>27184</td>
<td>16209</td>
<td>55261</td>
</tr>
<tr>
<td>Utterances including numerals</td>
<td>112</td>
<td>337</td>
<td>224</td>
<td>673</td>
</tr>
<tr>
<td>DM</td>
<td>51</td>
<td>134</td>
<td>98</td>
<td>283</td>
</tr>
</tbody>
</table>

In the English data, there are around 300 DMs out of around 700 utterances with the numerals 1, 2, and 3, showing that at age 1–2 English-speaking children are old enough to produce DM.
In the English data, about 40% of the utterances with numerals are DMs. In contrast to the English data, in the Japanese data in Table 2, there is no DM in around 170 utterances with numerals 1, 2, and 3, without classifiers. For Table 2, I conducted searches with three Japanese-speaking children’s CHILDES database at age 1–2. The names of the three children are: Tai (Miyata 2000), Aki (Miyata 1995), and Nanami (Miyata 2012).

**Table 2: Numerals in Child Japanese (for numbers 1, 2, and 3) at age 1–2**

<table>
<thead>
<tr>
<th>Child name</th>
<th>Tai</th>
<th>Aki</th>
<th>Nanami</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1;6–2;11</td>
<td>2;0–2;11</td>
<td>1;10–2;11</td>
<td>1;6–2;11</td>
</tr>
<tr>
<td>Total utterances</td>
<td>32218</td>
<td>20857</td>
<td>15016</td>
<td>68091</td>
</tr>
<tr>
<td>Utterances including numerals w/ classifiers</td>
<td>145</td>
<td>55</td>
<td>33</td>
<td>233</td>
</tr>
<tr>
<td>Utterances including numerals w/o classifiers</td>
<td>27</td>
<td>114</td>
<td>32</td>
<td>173</td>
</tr>
<tr>
<td>IM</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>DM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(Classifiers examined in this search are: -tsu (Inanimate Generic), -ko (3D Object), -ri/-nin (Human), -hiki (Animate).)

A chi-square test found the contrast between English and Japanese to be statistically significant at $p < .01$. The data in Table 2 show that there is no DM overgeneration in Child Japanese at age 1–2. By the way, I would like to point out that even IM is not abundant in Table 2. There are only 4 IMs in total in Table 2. It would be desirable to examine the existence of DM in larger corpora with more IMs.

Hence, I extended my search by including data for three Japanese-speaking children in the CHILDES database at age 3–4. Table 3 includes the data from additional searches of the data for these children. The names of the three children are: Nanami, ArikaM, and Asato (Miyata 2012). For Tai and Aki, no corpora at age 3–4 exist in the CHILDES database.
Table 3: Numerals in Child Japanese (for numbers 1, 2, and 3) at age 3–4

<table>
<thead>
<tr>
<th>Child name</th>
<th>Nanami</th>
<th>ArikaM</th>
<th>Asato</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>3;0–4;11</td>
<td>3;0–4;11</td>
<td>3;0–4;11</td>
<td>3;0–4;11</td>
</tr>
<tr>
<td>Total utterances</td>
<td>12097</td>
<td>45918</td>
<td>11581</td>
<td>69596</td>
</tr>
<tr>
<td>Utterances including numerals w/ classifiers</td>
<td>93</td>
<td>278</td>
<td>74</td>
<td>445</td>
</tr>
<tr>
<td>Utterances including numerals w/o classifiers</td>
<td>13</td>
<td>249</td>
<td>45</td>
<td>307</td>
</tr>
<tr>
<td>IM</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>DM</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(Classifiers examined in this search are: -tsu (Inanimate Generic), -ko (3D Object), -ri/-nin (Human), -hiki (Animate).)

Considering Table 2 and Table 3 together, there are again no examples of DM among the total 480 utterances that include the numerals 1, 2, and 3 without classifiers. Thus, again, there is no DM overgeneration. In Tables 2 and 3, there are 14 IMs in total. This is not a large number, but statistically speaking, if IM and DM are equally probable, the probability of getting 14 IMs in succession is extremely low: \((0.5)^{14} = 0.000061035\).

To summarize, there was no DM overgeneration in my search, which suggests that something innate blocks overgeneration.

4. Discussion

The searches of Child Japanese corpora found no DM overgeneration. Japanese-speaking children do not overgenerate DM, even though it is produced abundantly in Child English at the equivalent age. This goes against the Indirect Negative Evidence scenario. If Japanese-speaking children learn the lack of DM by Indirect Negative Evidence, DMs would be expected to be overgenerated until they gradually taper off.

Also, since Japanese-speaking children do not overgenerate DM, the Default scenario is not falsified. If the lack of DMs in Japanese is due to that being the default value of a parameter, overgeneration of DM would not be expected. In fact, as there is no DM overgeneration in the Child Japanese data presented here, the search results are consistent with the Default scenario.
Let us discuss a further issue here. As we have seen, there is variation between Japanese and English regarding DM. Bale and Coon (2014) point out that this variation is in some cases found even within one language, for example in Mi’gmaq (Algonquian) and Chol (Mayan). In Mi’gmaq, for instance, only DM is possible with the numbers 1–5, while only IM is possible with the numbers 6 and higher.

Thus, given languages like Mi’gmaq and Chol, we have a further prediction for the Default scenario: Overgeneration of IM is possible in these languages. According to the Default scenario, DM is impossible by default and becomes possible only in the presence of positive evidence. Hence, no overgeneration of DM is expected at the initial stage under the Default scenario. However, the Default scenario under discussion does not exclude the possibility of IM overgeneration. Hence, in Mi’gmaq acquisition, for instance, IM overgeneration is predicted to be possible with the numbers 1–5 in accordance with the Default scenario under discussion.

If there is no IM overgeneration in Mi’gmaq and Chol, this would mean that the Default scenario for the respective parametric acquisition is inadequate. If so, children are very conservative in DM/IM acquisition. My speculation for how this would work in such a case is as follows: No parametric option is innately given by default, but the child does not adopt any parametric option unless there is positive evidence for it. This is an application of “grammatical conservatism,” as posited by Snyder (2007), to the issue of the parametric acquisition discussed in this paper. I leave this issue for future research.

5. Conclusion

In this paper, I examined how Japanese-speaking children come to know that DM is impossible in their target language. Searches of child language corpora found no DM overgeneration in Japanese-speaking children’s spontaneous production data. This observation can be explained if we assume that DM is banned by the innately determined default value of the parameter and argues against the idea that the ban on DM is learned from Indirect Negative Evidence.

References


