A Discourse-Pragmatic Analysis of Referential Choice in English and Japanese Monolingual Children: A Longitudinal Study

MISHINA-MORI Satomi

Key words: referential choice, discourse-pragmatics, Japanese, English, longitudinal study

Abstract

The current study examines if the discourse-pragmatic principles for referential choice can account for young children’s use of verbal arguments at different stages of development in children acquiring both null and overt languages. Longitudinal data of two English and two Japanese monolingual children approximately two to three years of age were divided into three stages, and the relationship between the children’s referential choice and the discourse-pragmatic feature of the referent was analyzed. The results showed that from an early stage of development, English-speaking children chose arguments based on the information status—mostly lexical arguments for new information and non-lexical arguments for given information, although the two children showed different patterns in the development of the use of pronouns. Japanese children also showed some sensitivity to the information status of the argument but not in obvious ways: the majority of subjects were null for both new and given information, although there was a slight increase in the proportion of lexical arguments for new information compared with given information. The results suggest that discourse-pragmatic principles are at work from the earliest stages of development in children acquiring languages of distinct syntactic properties, but the developmental patterns seem to be affected by language-specific features.
1. Introduction

The frequent omission of verbal arguments is a prominent characteristic of children’s speech in the earliest stages of development. This is observed not only in null argument languages, such as Chinese (Wang et al., 1992), Inuktitut (Allen, 2000), Japanese (Hirakawa, 1993; Nakayama, 1994; Guerriero et al., 2006) and Korean (Clancy, 1993, 1997; Kim, 2000) but also in languages such as English, in which overt arguments are required (Bloom, 1990, 1993; Valian, 1991; Guerriero et al., 2006). Observe the examples below:

\[
\begin{align*}
(1) & \quad a. \; \emptyset \; \text{kore} \; \text{tabeta} \quad \text{(I ate this)} \\
& \quad \text{this} \; \text{eat-PAST} \\
& \quad b. \; \emptyset \; \emptyset \; \text{tabeta.} \quad \text{(I ate this)} \\
& \quad \text{eat-PAST} \\
(2) & \quad a. \; \emptyset \; \text{want} \; \text{this.} \\
& \quad b. \; \text{I} \; \text{want} \; \emptyset
\end{align*}
\]

Some studies that have documented the development of argument choice patterns report that children acquiring an overt argument language gradually come to realize arguments in accordance with development by approximately age three (Valian, 1991), whereas those who are learning null languages do not exhibit such change but continue to drop arguments at a certain rate (Nakayama, 1994).

Thus far, three different approaches have been proposed to account for this phenomenon in child language: the competence-based approach (Hymes & Wexler, 1993; Radford, 1990), the performance-based approach (Bloom, 1990, 1993; Kim, 2000; Valian, 1991), and the discourse-pragmatic approach (Allen, 2000; Clancy, 1993, 1997; Guerriero et al., 2006). The competence-based approach assumes that adults and young children have different underlying grammar. Adult grammar requires overt arguments, whereas children’s grammar does not, and later in development, children’s grammar approaches that of adult’s. It has been noted, however, that this explanation fails to capture the gradual decline in the use of null arguments in the children observed, as the approach would predict a sudden change from one stage to the next (Allen, 2000; Guerriero et al., 2006). More importantly, the competence-based approach was developed based on a number of Indo-European languages in which subjects can be null but objects cannot. Therefore, the explanation is targeted towards explaining why subjects can be dropped and does not necessarily apply to the object drop observed in languages such as Japanese and Korean.

The performance-based approach (Bloom, 1990, 1993; Kim, 2000; Valian, 1991), in
contrast, proposes that children do have the same underlying grammar as adults but fail to realize arguments due to the limitations in their processing capacities. Although this approach better explains the gradual change in the children’s referential choice patterns, it does not account for the argument omission in children learning null argument languages, as the rate of argument omission remains approximately the same throughout the developmental stages (Hirakawa, 1993; Nakayama, 1994).

Most recent studies on children’s referential choice have taken the discourse-pragmatic approach, which has proved to be successful in explaining adult argument drop phenomenon (Chafe, 1994; Du Bois, 1987; Givon, 1983) and has thus been applied to the analysis of child data (Allen, 2000; Clancy, 1993, 1997; Guerriero et al., 2006). This approach integrates syntax and discourse-pragmatics and proposes that the information status of the referent determines its form: new information, or a referent introduced to the discourse for the first time thus having a high informative value, tends to be expressed in lexical form, whereas given information, or a referent that has already been recognized by the participants, is more likely to be realized in non-lexical (either null or pronominal) form. The underlying assumption here is that speakers select referential forms based on their inferences about how the listeners perceive the information status of the referent in discourse. Studies on adult speakers (Chafe, 1994; De Bois, 1987; Givon, 1983), mostly based on the analysis of narrative data, have revealed that adults are highly sensitive to the flow of information in discourse: they lexicalize new information, but express given information in non-lexical forms, either a null or pronominal form, depending on the syntactic feature of the speaker’s language. The strong relationship between the information status of a referent and its form is a robust finding that has been observed in a variety of typologically distinct languages, including overt argument languages, such as English or French; languages that drop subjects, such as Spanish; and null argument languages, such as Japanese and Mandarin Chinese (Chafe, 1994; Du Bois, 1987; Givon, 1983).

Studies taking a discourse-pragmatic approach to the analysis of referential choice in child language development have shown that the alternation of referents in children as young as one to two years old can also be explained by the information status of the referent (Clancy, 1993, 1997; Allen, 2000; Guerriero et al., 2001). Clancy (1997) is one of the earliest studies to show that the discourse-pragmatic approach can account for young children’s alternation of verbal arguments. She analyzed the relationship between a number of discourse factors and the choice of referents in spontaneous speech samples of two children acquiring Korean, a null argument language. The study revealed that the distribution of the three referent forms (ellipsis, pronouns, and lexical noun phrases) varied depending on the discourse-pragmatic contexts (i.e., query, contrast, absence and
prior mention including given, accessible, and new information). For example, given information, or the referent being activated in discourse and therefore available to the participants, was likely to be expressed in elliptical forms, whereas in the context of query, contrast or absence, when there is a necessity to activate the referent for the first time or disambiguate between two or more referents, the referent tended to be in lexical form.

Allen (2000) presented similar findings in children acquiring Inuktitut, also a null argument language, as their first language. Four children aged approximately two were tracked for nine months, and the relationship between informative features and argument forms was examined. It was found that the more informative features the referent had, the more likely it was to be realized. This finding indicates that discourse features can adequately account for the alternation of referential forms in children as young as two years of age.

More recently, Guerriero et al. (2006) investigated children’s sensitivity to informative features from a developmental perspective by contrasting Japanese, a null argument language, and English, an overt argument language. Speech samples of six Japanese- and six English-speaking children of approximately two to three years of age were collected to analyze the following: 1) if both English- and Japanese-speaking children showed sensitivity to language-universal discourse-pragmatic features by using lexical forms for new information and non-lexical forms for given information; and 2) if the children showed sensitivity to language-specific discourse-pragmatic features by using pronominal forms in English and null forms in Japanese for given information. The results revealed that English-speaking children used referential forms in accordance with both the language-universal and language-specific discourse-pragmatic features as early as two years of age, whereas Japanese-speaking children did not show such sensitivity, even when they turned three. More specifically, Japanese children continued to use null forms for new information and thus did not appear to differentiate between new and given information. The authors further examined whether null forms in reference to new information accompanied non-linguistic cues, such as joint attention, to make the referent more accessible to the listener, but that was not necessarily the case. They also determined that parental input was a factor to a certain extent. From these findings, Guerriero et al. (2006) concluded that discourse-pragmatic principles cannot fully explain the children’s alternation of verbal arguments, at least in null argument languages such as Japanese, and that parental input also plays a role.

To summarize, the recent attempts to explain children’s referential choice based on the discourse-pragmatic approach have thus far been partially successful: although discourse-pragmatic features appear to predict the referential choice patterns in children acquiring a number of different languages, it did not fully account for the argument
choice in Japanese-speaking children. The deviation from the predicted pattern observed in the Japanese children in Guerriero et al.'s (2006) study was the frequent use of elliptical forms with reference to new information, one of the most obvious informative features that require lexical form. Guerriero et al. (2006) argued that this deviation can be explained by parental input; however, this argument is weak because it is unclear what motivates the adults to choose referents in that particular manner.

The purpose of the current study is to partially replicate Guerriero et al. (2006) and to determine if the results are consistent with theirs. That is, we examine whether English children develop sensitivity to discourse-pragmatic features in the early stages, whereas Japanese children do not, and both English- and Japanese-speaking children are able to use non-lexical forms for given contexts in language-specific ways. Furthermore, if the major deviation from the language-universal principle observed in Guerriero et al. (2006), that is, the predominant use of null forms with reference to new information in the Japanese data, is repeatedly found in the current study, we will attempt to give an explanation for the phenomenon and further discuss the validity of the discourse-pragmatic approach as an account for children's referential choice patterns. Below we summarize the basic characteristics of the structure of English and Japanese in terms of argument omission and realization.

2. Argument representation in English and Japanese

   English is an overt argument language, in which the realization of both subjects and objects is required, and null forms are a rare option. Observe the following examples, where both b. and c. are ungrammatical:

   (3)   a. Ken saw the movie.
        *b. Ø saw the movie.
        *c. Ken saw Ø.

   Children at the earliest stages of development are reported to omit arguments, but it has been reported that such omission gradually disappears by their third year (Bloom, 1990, 1993; Valian, 1991).

   Japanese, in contrast, allows subject as well as object omission as long as the referent is recoverable from the context. All three sentences below mean "Ken saw a movie", although the subject (Ken) and the object (movie) are missing in example b. and example c., respectively.
Nakayama (1994) conducted a developmental study on argument omission in a two-year-old and reported that the omission rate is more or less constant across different stages. That is, null argument is a prevalent phenomenon in both adult and child language. Another important characteristic of Japanese is that there are no grammatical cues that help recover the missing referent: it can only be recovered from contextual information (Huang, 1984). This is because Japanese does not mark agreement on the verb, and case markers can be dropped with the argument. Such feature is in stark contrast to other languages that allow null option, such as Spanish, Italian or Inuktitut, in which verb forms or agreement markers hold information about the characteristic of the referent.

3. Research question

The research question to be addressed in the current study is as follows: Can the discourse-pragmatic principles for referential choice account for the use of verbal arguments at different stages of development in monolingual children acquiring English and Japanese? More specifically, this study asks: 1) Do both English and Japanese children lexicalize new information and non-lexicalize given information? 2) Do English children mainly use pronouns, whereas Japanese children use null forms for given context? 3) Are there any developmental trends observed in these two aspects of argument choice?

The current study will focus on the use of subjects and exclude objects from analysis because subjects and objects show distinct patterns in referential choice. That is, within the flow of discourse, subjects tend to provide given information, whereas objects typically introduce new information (Clancy, 1993).

It is predicted that the results will be similar to those reported by Guerriero et al. (2006): English children will show sensitivity to both language-universal and language-specific discourse-pragmatic principles, whereas Japanese children will not show clear sensitivity to the language-universal principle, although their use of referential form may be consistent with the language-specific principle. More specifically, English-speaking children will alternate between different referential forms according to discourse-pragmatic features from early stage: the majority of new information will be expressed in
lexical forms from early on, whereas they will mainly use non-lexical forms for given information (language-universal principle). Furthermore, given information will be expressed predominantly in pronominal forms (language-specific principle) compared with other non-lexical forms, such as ellipsis. However, this development may occur later, as children produce more null forms than pronominal forms in the early stages of development. In contrast, Japanese-speaking children may not necessarily differentiate new and given information when selecting a referential form: they drop arguments in both contexts (language-universal principle), though they will follow the language-specific principle to use null forms rather than pronominal forms for given information.

4. Method

4.1. Subjects

Subjects for the current study were two English-speaking children, Sarah and Eve, studied by Brown (1973), and two Japanese-speaking children, Aki and Ryo, from Miyata (1995). The transcribed data of these children interacting with their caretakers were drawn from the CHILDES corpus (MacWhinney, 1995). To observe the developmental trend in their referential choice patterns, three stages were defined based on the children’s MLUs: stage 1 MLU 1.0-1.5, stage 2 MLU 1.6-2.0, and stage 3 MLU 2.1-2.5. Two to three files that matched the MLU range were chosen for each child for analysis. The biological ages of the children at stage 1 ranged from 1;6 to 2;3, at stage 2 from 1;7 to 2;9, and at stage 3 from 1;9 to 3;2. The ages and MLU range for each subject at different stages are presented in Table 1.

| Table 1. Ages and MLUs of subjects at each stage |
|---------------------------------|-----------------|-----------------|-----------------|
|                                 | Stage 1 MLU 1.0-1.5 | Stage 2 MLU 1.6-2.0 | Stage 3 MLU 2.1-2.5 |
|                                 | age | MLU | age | MLU | age | MLU |
| Japanese                       |     |     |     |     |     |     |
| Aki                            | 2;0–2;1 | 1.3 | 2;3 | 1.6–1.8 | 2;5–2;7 | 2.2–2.5 |
| Ryo                            | 1;10–1;11 | 1.1–1.3 | 2;2 | 1.8–2.0 | 2;5 | 2.3–2.5 |
| English                        |     |     |     |     |     |     |
| Sarah                          | 2;3 | 1.5 | 2;9 | 1.7–1.9 | 3;0–3;2 | 2.2–2.4 |
| Eve                            | 1;6 | 1.5–1.6 | 1;7–1;8 | 1.7–2.0 | 1;9 | 2.3–2.6 |

The stages were defined not by biological age but by MLUs because it has been argued elsewhere (e.g., Mishina, 1997) that there is a huge amount of individual difference in language development, which makes age-based comparison less reliable.
4. 2. Coding and analysis

The children’s utterances that accompanied verbs were coded according to the system adapted from Matsuoka et al. (2008), which is based on CHAT conventions (Mac Whinney, 1995). 1) The grammatical status of the argument in focus was coded as either subject or object, although the objects were excluded from the analysis. 2) The argument form of each subject was categorized into one of the three following forms: null, pronominal, and lexical. As for the null forms, omissions in English, such as of imperative forms, were excluded from the analysis because the omission is the norm, regardless of the discourse context. Pronominal forms include those that identify the referents that have been mentioned in the prior discourse. 3) The information status of the argument was coded as either new or given information. Whether the argument referred to new or given information was determined by Guerriero et al.’s (2006) criteria: new information refers to those that appear in the discourse for the first time or 20 or more turns after the previous appearance within the same discourse, and given information are those that have been introduced to the discourse. First and second person pronouns (I, you, we) were always coded as pronominal-given, also following Guerriero et al. (2006).

Quantitative analyses were conducted by CLAN programs (Mac Whinney, 1995), by which the frequencies of null, pronominal, and lexical forms with reference to new information and given information for each child were obtained.

5. Results

5. 1. English

The proportions of null, pronominal, and lexical form with reference to new information were calculated and are presented in Figures 1 and 2 for Sarah and Eve, respectively. The proportions of the three forms with reference to given information are presented in Figures 3 and 4 for Sarah and Eve, respectively. The data confirm our prediction regarding language-universal discourse-pragmatic principle—that English-speaking children would show sensitivity to the discourse-pragmatic context in their use of referential form. Children lexicalized the majority of the arguments when referring to new information, whereas they tended to use non-lexical forms (either null or pronominal forms) when referring to given information. Furthermore, this pattern is observed from stage 1 onwards and indicates that the children knew the distinction from the earliest stages of development.
Language-specific discourse-pragmatic principle predicts that children will predominantly use pronouns compared with other forms when referring to given information. However, this prediction was not exactly substantiated. At stage 1, both children mainly dropped the arguments when referring to given information, and pronoun use was infrequent. From stage 2 and beyond, Sarah’s referential choice showed a dramatic change, namely that pronouns dominated the arguments for given context. The change indicated that she had acquired the language-specific rule by the time her MLU approached 2.0. In contrast, Eve showed a developmental pattern totally different from that of Sarah’s. Throughout the 3 stages, null forms dominated the arguments, though the proportion of pronouns increased over the course of development. Thus, the data seem to suggest that Sarah was still on her way to acquiring the language-specific rule to pronominalize given arguments.

5. 2. Japanese

Figures 5 and 6 present the distributions of different referential forms at each stage when referring to new information for Aki and Ryo, respectively, and Figures 7 and 8
show the same distribution when referring to given information for Aki and Ryo, respectively. It is clear from these figures that the children did not, at any stage, show sensitivity to the informative values of the arguments in obvious ways: their subjects were predominantly null regardless of the discourse-pragmatic context. It should be noted, however, that some developmental changes are observed when referring to new information: the use of lexical subjects emerges at stage 2 and continues to occupy approximately 30% of the total number of referents. At later stages, the children also used more lexical forms for given context, but the ratio was much smaller than the ratio observed in new context. This difference may be a slight indication of differentiating new and given context. Thus, although there appeared to be an emerging sensitivity to the informative value of arguments, there was no clear-cut evidence for the claim that children choose referents according to the language-universal discourse-pragmatic principles.

Language-specific principle, on the other hand, is observed among the two children: given information was expressed predominantly in null forms throughout the observation period, and the use of pronouns was rare.
To summarize, English-speaking children exhibited sensitivity to the informative value of the argument from the earliest stages of development by lexicalizing new information and using non-lexical forms to express given information. However, the referential choice patterns for given context reveal that language-specific principle was a slower development. Sarah started to pronominalize the majority of arguments from stage 2, but Eve used null forms throughout the three stages, although the use of pronouns did increase over the course of development. Japanese-speaking children, in contrast, showed a different pattern: they did not appear to differentiate between given and new information in their referential choice pattern and dropped subjects most of the time, regardless of the information status of the argument, throughout the three stages. Language-specific principle was observed from the start, with the vast majority of given information expressed in null forms rather than pronominal forms.

6. Discussion

The current results are in most cases consistent with the predictions for both English and Japanese children. With regard to language-universal discourse-pragmatic principle, English-speaking children were found to be highly sensitive to the discourse-pragmatic context: they lexicalized new information and non-lexicalized given arguments most of the time. The results differ from Guerriero et al. (2006), who reported that English-speaking children did not differentiate new vs. given context at the earliest stage of observation, which was when the children were 1;9 and that such a distinction was observed when the children were 3;0. Because the current subjects, Sarah and Eve, were 2;3 and 1;6, respectively, at stage 1, it could be speculated that the emergence of such a distinction occurs around the time children turn 2;0, with some individual variation. However, this speculation cannot be confirmed, as Guerriero et al. (2006) observed the children when they were 1;9 and 3;0 only.

Language-specific principles in English, that is, prominent pronominalization of given information, was not consistently observed among the two children. At stage 1, Sarah used more null forms than pronominal forms but appeared to have acquired the language-specific rule by stage 2. This acquisition pattern is consistent with the prediction. However, Eve did not show such development: although the ratio of pronouns was on the rise, the majority of the referents she used was null forms throughout the observation period. We conjecture that this may be because Eve was significantly younger than Sarah in the study, only 1;9 at stage 3, whereas Sarah was 3;0. This gap is caused by the fact that Eve was advanced in her language development when assessed by MLU; thus, her biological age ended up being younger than Sarah’s. If we
turn to the findings by Guerriero et al. (2006), the use of pronouns was also a later
development in their English subjects, who used more null than pronominal forms at 1;9
but developed the ability to use pronouns by 3;0. Thus, the distinct pattern observed in
Eve’s speech may suggest that developmental stages need to be defined by age rather
than MLU when it comes to the development of pronouns.

Our data indicate that Japanese children do not differentiate new and given
information and drop subjects, regardless of the information status. These observations
are in line with our prediction as well as previous findings (Guerriero et al., 2006). It
appears that Japanese children, in natural discourse, omit subjects, even ones that have
high pragmatic saliency, that is, newness. The first possible explanation for such deviation
is the use of non-linguistic information. Allen (2000), who also observed a similar
phenomenon in her analysis of children acquiring Inuktitut, a null-argument language,
argued that this is possible if some non-linguistic information, such as joint attention
among the participants, helps them share the new information in the physical context
without the information being introduced to the discourse. This was a tempting
explanation, as it is generally believed that because young children tend to talk about the
here and now, parent-child communication can be significantly more context-dependent
than the adult norms. However, Guerriero et al. (2006) examined the relationship
between the use of non-linguistic information, such as pointing or gazing, but they found
that such information did not help clarify the referent of the null form. Therefore, the use
of non-linguistic information is not a strong candidate to account for this deviation.

A second possible explanation, as proposed by Guerriero et al. (2006), is adult input.
Guerriero et al. reported that there was a relationship between referential choice patterns
of mothers and those of their children and concluded that input may also play a role.
However, this conclusion still does not explain why adults omit arguments which are,
according to the discourse-pragmatic principle, expected to be expressed in overt forms.

The third and most likely explanation, as also proposed by Guerriero et al. (2006), may
be that the mother and child were engaged in a typical Japanese interactional style
observed in adults who are familiar with each other: that is, they freely omit arguments,
regardless of information status, and expect the listener to understand it correctly based
on the discourse context as well as shared knowledge. One way to confirm this analysis
is to observe children interacting with unfamiliar adults and examine if the referential
choice pattern is different from the current results. It is also necessary to study the
referential choice in adult Japanese interaction more in detail, in order to examine if the
proposed universal discourse-pragmatic principle applies to all languages.

The language-specific principle was observed in the Japanese children throughout
the three developmental stages: they used mostly null arguments for given information,
and pronominal forms were rarely observed. This observation is consistent with the prediction as well as the previous findings (Guerriero, et al., 2006).

7. Conclusion

The current findings seem to support the view that the discourse-pragmatic approach can explain referential choice in young children, although the frequent use of null forms for new information is yet to be accounted for. Further research includes more developmental studies on Japanese children’s referential choice for a longer period of time to confirm when the children start to select referents according to the proposed principle in Japanese. More cross-linguistic studies on children acquiring other languages, both null and overt argument languages, are also needed to see if the persistent use of null forms, even when the use of lexical form is more desirable, is observed only in null argument languages, as observed in the current study as well as Guerriero et al. (2006). Studies on children interacting with unfamiliar adults would also clarify if the frequent use of null arguments for new information is a result of communication style established among familiar individuals. Analysis on the adult norm based on interactional data may also be necessary because many of the discourse-pragmatic analyses of adult speech are based on narrative data, where the speaker and the listener do not share the physical context of most of the topics being discussed. There is a possibility that different patterns of referential choice may be found, and this variation would change the principle itself. Accumulation of substantial longitudinal studies on children acquiring typologically diverse languages, together with further studies on the referential choice patterns in adults, would shed light on the mechanism of referential choice in young children.

Notes

1) The abbreviations used in the gloss in the current paper are as follows: Ø = omitted argument; ACC = accusative case marker; TOP = topic case marker; PAST = past tense marker.

2) The MLU values for each interval had been calculated for each child based on Brown (1973) for English and Miyata (1995) for Japanese and are indicated in the CHILDES file for all children except for Ryo. Ryo’s MLUs were obtained from Suzanne Miyata through personal communication.

References


A Discourse-Pragmatic Analysis of Referential Choice in English and Japanese Monolingual Children
