A comparison of JEFLLs' and NESs' referential expressions in written English narratives in terms of distribution patterns and referential distance

INOI, Shin'ichi

茨城大学教育学部紀要（教育科学） 59: 97-108

2010-03-31

http://hdl.handle.net/10109/1335

このリポジトリに収録されているコンテンツの著作権は、それぞれの著作権者に帰属します。引用、転載、複製等される場合は、著作権法を遵守してください。
A Comparison of JEFLLs’ and NESs’ Referential Expressions in Written English Narratives in terms of Distribution Patterns and Referential Distance*

Shin’ichi IINO

(Received November 30, 2009)

Abstract

This study compares 39 Japanese EFL learners’ (JEFLLs’) and 9 native English speakers’ (NESs’) referential expressions for story characters in written English narratives in terms of distribution patterns and referential distance. It shows that JEFLLs produced pronouns and bare nominals frequently during L2 English narrative production, whereas it was pronouns and definite nominals that NESs produced frequently during L1 English narrative production. It also shows that referential distance is a key factor in NESs’ use of referential expressions in narratives, whereas it was not a determining factor in JEFLLs’ use of referential expressions in narratives.

Introduction

As a language teacher and researcher, I frequently encounter Japanese EFL (English as a foreign language) learners’ peculiar narrative texts. The following English narrative text, which was taken from one of my previous studies (Inoi, 1991), is a typical example.

[1]

① Once upon a time there lived an old man and an old woman in a small village.
② One day an old man went gathering firewood in the mountains, and an old woman went washing clothes in the stream….
③ When an old woman was washing, a peach came floating topsy-turvy down the stream….

(Inoi, 1991, p.102)

The narrative was actually produced by a Japanese EFL learner in translating the beginning part of Momotaro (The Peach Boy), a traditional Japanese folk tale. In [1], the referential expression “an old man” in the first sentence was repeated exactly in the second sentence. In this case the definite article “the” should have been used instead, as in “the old man went gathering firewood in the mountains” since the same person was

* This paper was orally presented at the 48th JACET Annual Convention at Hokkai Gakuen, Sapporo, Hokkaido, in September, 2009.
* * The Department of English, College of Education, Ibaraki University, Mito 310-8512 Japan
mentioned again. The phrase “an old woman” in the first sentence was also repeated in the second and third sentences. In other words, the noun phrases “an old man” and “an old woman” were employed as fixed referential expressions for the old couple in the narrative as if they were proper nouns; the phrases were neither pronominalized nor specified by the definite article “the.” As a result, each of the three individual sentences is grammatically correct at the sentence level, that is, when isolated from the rest of the narrative. However, when analyzed across sentence boundaries, they turn out to be grammatically incorrect. The sample above clearly shows the importance and necessity of discourse-level analysis of learner language in the field of second/foreign language teaching and learning.

Larsen-Freeman (1980) also takes the same view, arguing for the need of analysis across sentence boundaries or discourse analysis in language teaching and learning as follows:

Indeed, in focusing only upon structures at the sentence level, researchers may have perpetrated another misleading simplification of the language acquiring process. They have overlooked the need for the learners to acquire knowledge of the rule-bound system that exists at the discourse level, both in the oral mode, e.g., conversation (Sacks, Schegloff, and Jefferson, 1974), and in written discourse (e.g., Halliday and Hasan, 1976). (Larsen-Freeman, 1980, vi)

The present study tries to address the issue of reference, a particular type of cohesion, in written English narrative texts produced by Japanese EFL learners (JEFLs). The study analyzes referential expressions across sentence boundaries (i.e., at the level of discourse) as well as within sentences (i.e., at the level of syntax). In this paper I take a quantitative approach to data analysis. For qualitative analysis, see Inoi (2006a, 2007, 2008). The study focuses mainly on a variety of referential expressions they employed to refer to story characters during narrative production. Various types of English referential expressions produced by JEFLs are compared with those produced by native English speakers (NESs) mainly in terms of distribution patterns and referential distance.

Distribution patterns refer to the frequency ratios of various types of referential expressions in narratives, while referential distance refers to the distance between a referential expression and its antecedent, or between two references to the same referent. As in Clancy (1980) and Givón (1983a), the referential distance is measured by the number of clauses or clause boundaries which separate a referential expression and its antecedent. For example, in the sentences “Wash and core six cooking apples. Put them into a fireproof dish” (Halliday & Hasan, 1976, p. 2), the pronoun “them” in the second sentence is a referential expression and its antecedent is “six cooking apples in the first sentence; the pronoun has a referential distance score of one clause (boundary) because its antecedent is explicitly mentioned in the immediately preceding clause. As will be exemplified later, the referential distance was calculated only in the case of explicit or direct anaphora.

The notion “referential distance” is in fact based on Givón’s (1983a) topic continuity scale:
more continuous/accessible topics

- zero anaphora
- unstressed/bound pronouns ('agreement')
- stressed/independent pronouns
- full NP's

more discontinuous/inaccessible topics

Figure 1. Topic continuity scale (Givón, 1983a, p. 18)

According to this scale, “unstressed/bound pronouns” have their antecedents in the immediately preceding or following text. Anaphoric pronouns belong to this category. On the other hand, “stressed/independent pronouns” are used either contrastively or as topic switchers. Givón argues that a correlation exists between topic continuity and ease of topic identification: “The more disruptive, surprising, discontinuous or hard to process a topic is, the more coding material must be assigned to it” (1983a, p. 18).

Givón’s claim is based upon the assumption that as referential distance increases, so does the amount of coding material used to maintain reference. His research shows that referential distance has quite a significant influence on the choice of referential expressions in native English speakers’ oral narratives. The present study examines whether referential distance is a key factor in JEFLLs’ choice of referential expressions in their L2 English narratives.

Literature Review

There have been a limited number of studies which address referential expressions in a quantitative approach (e.g., Clancy, 1980; Chaudron & Parker, 1990; Givón, 1983a, 1983b, 1984; Hinds, 1983; Inoi, 2007; Tanimura, 2001, 2003; Williams, 1989). This section focuses mainly on Clancy (1980) and Williams (1989) because these studies analyzed referential expressions in terms of referential distance.

Clancy’s (1980) study examined and compared English and Japanese referential expressions in terms of time (i.e., the number of clauses and sentences separating two references to the same referent), and interference (i.e., the number of other referents intervening between the two references). These measurements are quite similar to Givón’s (1983a, 1983b, 1984) referential distance and topic ambiguity, respectively. Clancy reported that in both English and Japanese narratives, as the interval from the last reference to a referent or the number of other intervening referents increased, the explicit forms, such as full nominal phrases, were more likely to be used. This finding supports Givón’s argument for the topic continuity scale (Figure 1). However, Clancy also indicated that even when there were none of these cognitive constraints, Japanese speakers used a greater number of nominal forms than English speakers did.

Clancy noted another difference between Japanese and English speakers: the Japanese speakers’ stronger reliance on ellipsis or use of zero anaphor. She attributed these differences to language-specific factors. She speculated that Japanese speakers’ frequent use of nominal forms might be due to their tendency
to need more time to establish a new character as old information after his or her introduction in the narrative. She also speculated that Japanese speakers’ frequent use of ellipsis might be because of the number of ambiguity-clarifying devices available to the Japanese language, such as repeated and postponed subjects. She suggested further research on the issue of repetition of nominal phrases in Japanese discourse. Clancy’s (1980) findings and explanations are all based on the analyses of spoken L1 Japanese and English narratives. The present research, however, compares written L1 and L2 English narratives, with the former produced by NESs and the latter by JEFLLs.

Williams (1989) investigated the referential devices produced by three speaker groups: native speakers of American English, Singaporean English speakers, and ESL learners. L1 backgrounds of Singaporean English speakers were Malay, Tamil, and Chinese, whereas those of ESL learners included Chinese, Malay, Korean, and Vietnamese. The task given each subject was to have a 45-minute, unguided, free conversation in pairs. A total of six pairs were formed. None of the pairs shared an L1 or dialect, except the native English speakers. Williams focused on the third-person referential expressions which acted as subjects of main verbs. The forms analyzed were zero anaphors, pronouns, definite nouns, and pronoun copies. A pronoun copy is a co-referential subject pronoun following a noun. For example, in “His wife she died at 39” (Williams, 1989, p. 155), the pronoun “she” is referred to as a pronoun copy since it is co-referential with the immediately preceding noun phrase “his wife.”

Williams analyzed the referential devices in terms of three quantitative measures: referential distance, potential ambiguity, and persistence or decay. Referential distance was measured by counting the number of clauses between two references to the same referents, as in Clancy’s (1980) study and Givón’s (1983a, 1983b) studies. A clause was defined as a grammatical unit that contains a finite verb. Potential ambiguity refers to “the presence of other referents in the immediately preceding discourse which are semantically compatible with the referent being measured” (Williams, 1989, p. 156). Persistence or decay refers to “how long the topic survives, that is, how far to the right the reference to that topic persists in subsequence discourse” (1989, p. 156).

As for referential distance, Williams reported that all three speaker groups showed a similar pattern in the production of the referential devices: in general, referential distance was the longest in the case of pronoun copies, and it decreased in the cases of definite nouns, pronouns, and zero anaphors in this order. However, ESL speakers did not show any significant difference in referential distance between pronouns and zero anaphors, which means they used these two devices interchangeably in the same contexts. As for potential ambiguity, the three groups “tended to use pronoun copies and definite nouns in potentially ambiguous contexts, and conversely, they generally used pronouns where there are no referents which might cause confusion, and the repetition of the full noun could be considered redundant” (Williams, 1989, p. 174). This tendency was even observed among ESL speakers, though they did not differentiate between zero anaphors and pronouns. As for topic persistence or decay, no salient tendency was observed in the use of referential devices across groups.

In addition, Williams reported that the variable use of devices for topic marking was found within rather than between groups, particularly in the non-native English speaker groups. She concluded that “regardless of
what variety of English is spoken—a native, institutionalized, or learner variety—the function of these linguistic devices remains relatively constant” (Williams, 1989, p.185).

However, Williams noted two differences across the three speaker groups. One is the non-English native speaker groups’ frequent use in a wide variety of contexts of zero anaphora (or the omission of pronouns) in the subject position of a sentence. The other difference is the ESL speaker group’s extensive use of pronoun copies, compared to native English speakers and Singaporean English speakers. She attributed the former difference in part to L1 transfer, or the transfer of the speakers’ first languages, in which subject pronouns were more often optional. She added another reason in the case of ESL speakers. It is “the combinations of the limited linguistic resources and the demands of immediate production” (Williams, 1989, p. 183) which she claims may also have led to ESL speakers’ frequent omission of references to concrete nouns. As for the second difference (i.e., the ESL speaker group’s extensive use of pronoun copies), she did not give any explicit explanation.

The present study differs from Williams’ (1989) study in that it analyzes non-spontaneous written narrative production, whereas her data was spontaneous oral interactive production between two persons. She admits that analysis of data in a different type may produce different findings, particularly concerning the referential devices for topic persistence:

Finally, there is some indication that these devices may operate differently within extended narratives [italics added]. In these cases, once a theme for the narrative is established, it may be developed with more continuous referential devices. (Williams, 1989, p. 184)

The present study will also include Williams’ (1989) approach to analysis of referential expressions, particularly the measurement of referential distance. It would also be interesting to find whether pronoun copies and zero anaphors occur in Japanese EFL learners’ written English production. Kanzaki (1994), in fact, reports that zero anaphors frequently occur in the Japanese language. If JEFLLs transfer this feature into their English narratives, zero anaphors will likely take place.

The present study has the following research questions:
1) What types of referential expressions are likely to occur in JEFLLs’ and NESs’ written English narratives?
2) Is the average referential distance score of each type of referential expressions different between JEFLLs’ narratives and NESs’ narratives?

Method

The participants were 39 Japanese EFL learners (JEFLs) and 9 native English speakers (NESs). The JEFLs (7 males and 32 females) were aged between 18 and 20, majoring in food and nutrition, social welfare, and industry and design, at a junior college in Fukushima Prefecture, Japan. None of them were English majors, nor had any ever been abroad at the time when the data was collected. They were all learning
English as a foreign language.

The NESs (5 males and 4 females) were from either the United States or Canada. Seven of them were English teachers at colleges or universities in Japan, their length of residence in Japan raging from several years to over 20 years; the rest (i.e., two native English speakers) were college students visiting Japan temporarily at the time the data was collected. They all willingly took part in the project and produced written L1 English narratives.

Both JEFLLs and NESs were instructed to look at a series of pictures about what a family did on a weekend and to write in English what the family did in such a way that people who had not seen the pictures would be able to understand it. There were three characters depicted in the pictures: a man, his wife, and their son. The subjects were also instructed to build upon the lead sentence “A family was talking about what to do next weekend.” The JEFLLs were allowed to use English dictionaries if necessary. (See Inoi, 2006a, 2009 for the pictures and detailed data collection procedure.) JEFLLs originally produced both L1 Japanese and L2 English narratives on the basis of pictures, whereas NESs only produced L1 English narratives. In this paper, however, only JEFLLs’ L2 English data was compared with NESs’ L1 English data in a quantitative manner.

Referential expressions analyzed include such forms as zero anaphors (zero pronouns), definite nominals with “the,” third-person pronouns, full nominal forms, and bare nominals (nominals without any articles). Excluded from analysis were the expressions referring to non-human objects such as “the river” and “a restaurant” because the present study focuses specifically on the referential expressions employed for the three people depicted in the pictures.

Analysis

Frequency of referential expressions

Table 1 describes a summary of the frequencies of a variety of referential expressions in JEFLLs’ L2 and NESs’ L1 English narratives. In this study, a clause was defined as a grammatical unit consisting of a subject and a finite verb, and quotes did not count as clauses. However, there were two expressions in JEFLLs’ English narratives which did not contain a finite verb. One of them was “Father and child getting

Table 1. Descriptions of the Numbers of Referential Expressions and Clauses in JEFLLs’ L2 and NESs’ L1 English Narratives.

<table>
<thead>
<tr>
<th></th>
<th>Pronouns</th>
<th>Definite nominals</th>
<th>Zero anaphors</th>
<th>Bare nominals</th>
<th>Indefinite nominals</th>
<th>Total</th>
<th>Clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>JEFLLs’ L2</td>
<td>111 (34.05%)</td>
<td>44 (13.50%)</td>
<td>8 (2.45%)</td>
<td>148 (45.40%)</td>
<td>15 (4.60%)</td>
<td>326</td>
<td>295</td>
</tr>
<tr>
<td>Average</td>
<td>2.85</td>
<td>1.13</td>
<td>0.21</td>
<td>3.79</td>
<td>0.38</td>
<td>8.36</td>
<td>7.56</td>
</tr>
<tr>
<td>NESs’ L1</td>
<td>42 (46.15%)</td>
<td>30 (32.97%)</td>
<td>14 (15.38%)</td>
<td>5 (5.49%)</td>
<td>-</td>
<td>91</td>
<td>103</td>
</tr>
<tr>
<td>Average</td>
<td>4.67</td>
<td>3.33</td>
<td>1.56</td>
<td>0.56</td>
<td>-</td>
<td>10.44</td>
<td>11.44</td>
</tr>
</tbody>
</table>
fishing,” a finite verb missing. Still such an expression counted as a clause. As shown in Table 1, the 39 JEFLLs produced a total of 326 referential expressions in 295 clauses in L2 English narratives, each JEFLL participant generating an average of 8.36 referential expressions and an average of 7.56 clauses. The 9 NESs, meanwhile, produced a total of 91 referential expressions in 103 clauses in L1 English narratives, each NES participant generating an average of 10.44 referential expressions and an average of 11.44 clauses.

Figure 2 graphically shows the distribution pattern of JEFLLs’ L2 referential expressions in English narratives. Bare nominals occurred most frequently (i.e., 45.40%), followed by pronouns (34.05%) and definite nominals (13.50%). Bare nominals and pronouns account for nearly 80% of the total number of referential expressions in JEFLLs’ narratives. The occurrences of indefinite nominals and zero anaphors only account for less than 5% of the total frequency, respectively.

The distribution pattern of referential expressions is compared between JEFLLs’ L2 and NESs’ L1 English narratives (see Figure 3). One easily identifiable difference between them is that JEFLLs produced pronouns (34.05%) and bare nominals (45.40%) frequently in L2 narratives, whereas it was pronouns
(46.81%) and definite nominals (32.98%) that NESs produced frequently. Another noticeable difference is that JEFLLs’ production rates of pronouns (34.05%), definite nominals (13.5%), and zero anaphors (2.45%) in L2 narratives were much smaller than NESs’ production rates of these forms (i.e., pronouns, 46.81%; definite nominals, 32.98%; zero anaphors, 14.89%). Still another difference is that JEFLLs produced indefinite nominals with “a” (4.60%) as referential expressions while none of the NESs did. A further difference is that JEFLLs’ L2 production rate of bare nominals (45.40%) was far larger than that of NESs (5.32%).

However, a similarity exists in the distribution pattern of various types of third-person pronouns between JEFLLs’ L2 and NESs’ L1 English narratives, as shown in Figure 4. In both JEFLLs’ and NESs’ narratives, the great majority of pronoun occurrences were the third-person plural pronoun “they.” Out of the 111 pronouns produced by JEFLLs in narratives (see Table 1), 89 (80.18%) were occurrences of the plural pronoun “they.” There were two common linguistic contexts in narratives where both JEFLLs and NESs usually produced “they.” One was in a sentence which described the father’s suggestion of going camping or the family’s decision on camping, as in “The father suggested they go camping to get away from the city for a while” or “They decided to go camping.” The other was in a sentence which described the family ending up eating a meal in a restaurant, as in “They went to a family restaurant for lunch.” In the former context 15 (38.46%) out of the 39 JEFLLs and 6 (66.67%) out of the 9 NESs produced “they,” whereas in the latter context as many as 28 JEFLLs (71.79%) and 6 NESs (66.67%) produced “they.” Probably this brought about the frequent occurrences of the pronoun “they” in both JEFLLs’ L2 and NESs’ L1 English narratives.

![Figure 4](image_url)

**Note.** The pronouns “he” and “she” are combined into one category because both are third-person singular nominative pronouns, while the genitive pronouns (‘his” and “her”) and the objective pronoun “him” are grouped into another category because of their low frequencies.

**Figure 4.** Comparison of the Distribution Pattern of Third-Person Pronouns between JEFLLs’ L2 and NESs’ L1 English Narratives.

**Referential distance**

As was demonstrated in my previous studies (Inoi, 1991, 1997, 2006a, 2007), JEFLLs are likely to use the same referential expressions, bare nominals in particular, for story characters as sort of fixed expressions or proper names in L2 English narratives: they use particular referential expressions irrespective of the length
of referential distance. However, in this section referential distance is measured for each type of referential expression in JEFLLs’ narratives for two reasons: one is to compare the referential distance between JEFLLs’ L2 and NESs’ L1 English narratives; the other reason is to examine whether referential distance functions as a key factor in the choices of referential expressions in L2 narratives, as researchers argue (Clancy, 1980; Givón, 1983a, 1983b, 1984; Hinds, 1983; Williams, 1989). Referential distance refers to the interval which has elapsed between two successive references to the same character in a narrative, which was measured by the number of clauses between a referential expression and its direct antecedent.

In the following sample [2], five out of the eight referential expressions for story characters were measured for their referential distance scores. The first pronoun “they” in line 2 has a referential distance of one clause (boundary) because it directly refers back to the referent “family” in the lead sentence (i.e., direct anaphora). The next pronoun “they” in the same line also has a referential distance of one because it refers back to the first pronoun “they” in the immediately preceding clause. The bare nominal expressions in line 3 ("mother", “father,” and “child”) were eliminated from referential distance measurement because each of these referential expressions did not have its direct antecedent in the preceding text. The bare nominal “mother” in line 4, in turn, has a distance of two clauses because its referent was explicitly mentioned in line 3. The bare nominal “boy” in line 4 also has a distance of two because it refers to the bare nominal “child” in line 3.

[2] [A family was talking about what to do next weekend.]

They decided camping./ And they was going to cook and fishing a fish./ Mother was cooking./ Father and child was arranging for fishing./ But mother burned./ And boy fished an empty can./ Ultimately they ate a meal in restaurant./

Note. The first sentence in parentheses is the lead sentence given; the slashes indicate clause boundaries; the referential expressions for story characters are underlined.

Figure 5 below shows the average referential distance scores of referential expressions in JEFLLs’ English narratives. The scores are more or less the same, or they fall into the small range between 2.45 and 2.67, except for zero anaphors. However, it does not support the findings by researchers (Clancy, 1980; Givón, 1983a, 1983b, 1984; Hinds, 1983; Williams, 1989). Givón’s topic continuity scale (see Figure 1) does not seem to apply to JEFLLs’ production of L2 English referential expressions. This seems to be due to JEFLLs’ tendency to generate particular referential expressions for story characters as fixed expressions irrespective of referential distance in L2 English narratives. To put this another way, JEFLLs tried to generate referential expressions in L2 English narratives in the same way they did in L1 Japanese narratives. This clearly shows that referential distance cannot be a determining factor in JEFLLs’ choices of various referential expressions in English narratives.
Figure 5. Average Referential Distance Scores of Referential Expressions in JEFLLs’ L2 English Narratives.

Figure 6 compares the average referential distance scores in JEFLLs’ L2 English narratives with those in NESs’ L1 English narratives. As shown in the figure, NESs’ choices of referential expressions varied in accordance with referential distance: definite nominals have the largest distance score (3.38), zero anaphors have the smallest distance (1.00), and pronouns have the score in-between (1.81). That is, referential distance was a determining factor in NESs’ choices of referential expressions during L1 narrative production. However, JEFLLs’ choices of L2 English referential expressions did not vary so much except for those of zero anaphors, which shows the non-significance of referential distance in JEFLLs’ choices of English referential expressions. Zero anaphors in JEFLLs’ L2 narratives, however, have the average distance score of one clause, which is exactly the same score as in NESs’ narratives. Like NESs, JEFLLs probably limited the use of a zero anaphor to the subject position of the second clause of a coordinate clause in L2 English narratives.

Figure 6. Comparison of Referential Distance between JEFLLs’ L2 Narratives and NESs’ L2 Narratives
Conclusion

This paper compared JEFLLs’ L2 and NESs’ L1 referential expressions for story characters in written English narratives in terms of distribution patterns and referential distance. The analysis showed that the distribution patterns of JEFLLs’ L2 referential expressions differed from those of NESs’ L1 referential expressions. JEFLLs produced pronouns and bare nominals frequently, whereas it was pronouns and definite nominals that NESs produced frequently. The study also showed that JEFLLs did not vary so much in choosing English referential expressions, while NESs produced referential expressions on the basis of referential distance. NESs were likely to produce full nominals as referential distance increased. These findings indicate that referential distance was not a determining factor in JEFLLs’ selection of referential expressions except for zero anaphors in L2 English narratives, while referential distance was a key factor in NESs’ use of referential expressions.

However, the findings are all based upon a limited number of JEFLLs’ and NESs’ written English narratives. Future research should examine orally produced data elicited from a large number of people. In addition, it would be interesting to collect data from English learners with a variety of L1 backgrounds.

This research clearly demonstrated the importance and necessity of discourse-level approach to language teaching and learning so that Japanese EFL learners can produce coherent and cohesive narratives. Narrative production seems to be a useful means toward teaching language learners, particularly JEFLLs, such linguistic forms as articles and pronouns because these linguistic forms are very sensitive to context. Narrative production may help language learners realize the differences in the way story characters are referred to in L1 and L2 narratives.

References


cross-language study, Vol. 3, 47-93.


