The Effect of 'Weight' of the Relative Clause Construction in L2 Japanese Production

Kazue Kanno
University of Hawaii at Manoa

要 旨

第二言語習得の発話データにおいて関係節の使用頻度の低さが報告されいるが、本稿では頻度の低さの要因として、関係節の構造的複雑さがもたらす「重さ」が関与しているのではないかと考え、検討した。調査は英語を母語とする上級日本語学習者を対象として、2 つの発話データにおいて使用された関係節を様々な面から分析、調査した。その結果、学習者の関係節の特徴として「短い」ということ、そして文頭の部分に関係節が使われる頻度が高いことが分かった。これは関係節の構造的複雑さ、又日本語の関係節が前置型であることが即時的になされなければならない発話にもたらす影響であり、関係節の頻度の低さに貢献しているのではないかと考える。

1. Introduction

Relative clauses (RCs) are known to be one of the more difficult patterns for second language (L2) learners, and previous research shows that these patterns are used very infrequently in spontaneous speech by non-native speakers (e.g., Tamaru et al., 1993; Saito, 2002). Why is this the case?

Previous studies (e.g., Kanno 2001, 2007) report that even beginning-level learners can comprehend simple subject relatives quite accurately, suggesting that second language learners do not lack knowledge of the construction per se.

Schachter (1974) puts forward avoidance as an explanation for the infrequency of relative clauses in the spontaneous written production data that she studied. In other words, learners avoid using RCs because they are difficult in some sense. Can avoidance also explain the infrequency of RCs in spontaneous

oral production, or are there other factors limiting learners' use of RCs? This study investigates the latter possibility by exploring how the production factor of 'weight' might contribute to the infrequent use of RCs in the spontaneous oral production of L2 speakers of Japanese.

2. Weight

Weight can be defined in various ways, but a constituent's string length and structural complexity are generally accepted factors in the calculation of 'heaviness' and are known to affect the way speakers organize sentences. (See Arnold et al., 2000, for more discussion of this topic). In English, for instance, the direct object of a directional transitive verb generally precedes the goal phrase, as shown in (1a) below. However, when the direct object noun phrase (NP) contains a 'heavy' constituent, such as a sentence modifier, it may move to the right of the goal phrase, as in (1b).

(1) Heavy NP Shift

a. The boy brought [$_{NP}$ **the dish [$_{IP}$ I had ordered]]** [to the table].
direct object *goal phrase*

b. The boy brought [to the table] **[$_{NP}$ the dish [$_{S}$ I had ordered]].**
goal phrase *direct object*

In contrast, in a head-final language such as Japanese in which a verb has to come at the end of the sentence, a heavy constituent tends to be preposed, as illustrated by the leftward scrambling of the sentential complement in the following sentence.

(2) Preposing of a sentential complement:

a. Subject-Indirect Object-Clause-V:
[John-ga] [Mary-ni] [$_{S'}$ [$_{S}$ **Susan-ga kaisha-o yameta**] **to**] itta.
John-Nm Mary-to Susan-Nm company-Ac quit that said
'John said to Mary that Susan quit the job.'

b. Clause- Subject-Indirect Object -V:
[$_{S'}$ [$_{S}$ **Susan-ga kaisha-o yameta**] **to**] [John-ga] [Mary-ni] *e* itta.

A similar effect is observed with relative clauses: Yamashita (2002) reports that an NP containing a relative clause (RC) tends to be 'scrambled' leftward over a 'lighter' constituent.

(3) Preposing of a heavy NP containing a relative clause

a. Canonical: Subject [[RC] $N]_{obj}$ – V
Kare-wa [[$_{RC}$ watashi-ga oboeteiru] hanashi]-o sukkari
he-Tp I-Nm remember story-Ac completely
wasurete-simatteiru kamosirenai
forget-have maybe
'He may have forgotten completely the story that I remember.'

b. Scrambled : [[RC] $N]_{obj}$ -Subject- V (Yamashita, 2002: p. 604)
[[$_{RC}$ *Watashi-ga oboeteiru*] hanashi]-o kare-wa *e* sukkari
wasuretesimatteiru kamosirenai

Based on a corpus study of various types of scripts ranging from casual conversation to formal writing, Yamashita reports the following two findings:

(i) A scrambled phrase tends to be relatively 'long'. (it typically contains 3 to 13 content words, with an average of 5.6).
(ii) A scrambled phrase tends to be syntactically complex (i.e., it contains either a sentential complement or a relative clause).

The syntactic complexity of RCs is increased still further by the fact that they contain a gap. In a language such as Japanese, in which RCs precede the head noun, a gap has to be planned in advance (i.e, prior to the occurrence of its filler—the reverse of what happens in English), placing further burden on the production mechanisms.

(4) [John-ga _ katta] hon
John-Nm gap bought book
'the book that John bought'

Overall, then, there is good reason to believe that the addition of a relative clause to a head noun makes a phrase 'heavier' with respect to both string length and structural complexity. This in turn is consistent with the key proposal of this paper, which is that the weight of relative clauses contributes to their relative infrequency in the speech of Japanese L2 learners. More specifically, I propose that the weight of RCs interacts with the incremental planning of the sentence in the course of spontaneous production. As I will show next, evidence for this hypothesis comes from the properties of the relative clauses produced by L2 learners.

3. Current study

3.1 Data and participants

Two sets of spoken data were collected from 7 English-speaking (non-heritage) L2 learners of Japanese who were at the ACTFL OPI Advanced Level(1): (i) a guided narrative(2) and (ii) "free" conversation (i.e. the Oral Proficiency Interview). Narrative data were also collected from three native speakers of Japanese and are included here for the purposes of comparison. (The data used in this study are part of the data collected for Kanno, Hasegawa, Ikeda, Ito and Long (2007).)

The reason for selecting advanced-level learners is a practical one: Ozeki's (2005) data suggest that intermediate-level learners produce a mean of just 2.6 relative clauses in their Oral Proficiency Interviews, compared to 11.7 for advanced-level learners.

3.2 Relative clauses extracted from the narrative data

As the following examples illustrate, it is often not easy to distinguish RCs from phrasal modifiers in Japanese. This is not only because all types of modifiers precede the head noun in Japanese (a head-final language), but also because Japanese lacks the relative pronouns and other RC markers found in many Indo-European language.

(5) a. [ooki-i] hito
big-Prs person
'a big person'

b. [me-ga ooki-i] hito
eyes-Nm big- Prs person
'a person with big eyes'

(6) a. [yude-ta] tamago
boil-Pst eggs
'boiled eggs'

b. [kinoo yude-ta tamago]
yesterday boil-Pst eggs
'the eggs that (someone) boiled yesterday'

In this study, only clausal modifiers that include both a gap and a verbal predicate were extracted for analysis; modifiers whose predicate was an adjective or an adjectival nominal were deliberately excluded. This was because English adjectival modifiers [e.g., *the very tired man*] are phrases rather than clauses and precede the head noun, as in Japanese. Given the possibility of transfer, this made it difficult to know what status the English-speaking L2 learners assign to comparable phrases in Japanese.

4. Results

4.1 Degree of embedding/subordination

Table 1 presents the overall mean dependent clause/clause (DC/C) ratio in the narrative data for the native speakers and for the learners of Japanese. (This was calculated by dividing the number of dependent clauses (DC) by the total number of clauses (C).) These ratios were then used to examine the degree of embedding/subordination with a view to determining whether the learners are indeed at a stage where they can handle complex sentences.

Table 1: Dependency clause ratios

	Dependency clause (DC)/Clause (C)	DC/C Ratio
Learners	83/269	0.308
Native speakers	184/293	0.628

The DC/C ratio for the learner group (0.308) clearly suggests that they are able to produce complex sentences involving embedded clauses.

4.2 Number of RCs extracted from the narratives and the OPI data

A total of 59 RCs were extracted from the learners' narratives and OPI data, and 45 from the native speakers' narratives.

Table 2: RCs extracted from the narratives and the OPI data

	No. of RCs extracted from the narrative data	No. of RCs extracted from the OPI data	Total
Learners	30 (out of 269 clauses)	29	59
Native speakers	45 (out of 293 clauses)	No data	45

These RCs were then analyzed in terms of various factors relevant to their 'weight'. I will begin with a report on the analyses of the narrative data.

5. Analyses

5.1 Analyses of RCs in the narratives

5.1.1 Weight of RCs

In order to assess the weight of RCs with respect to their string length, we counted the number of words per clause, including postpositions.(3) Table 3 presents the mean length of RCs for the two groups of subjects.

Table 3: Length of RCs produced by the two groups

	No. of content words + postpositions	No. of RCs	Mean length of RCs
Learners	**70 (=59+20)**	30	2.33
Native speakers	186 (=127+59)	45	**4.13**

The average mean length of RCs for the learner group is 2.33, compared to 4.13 for the native speakers.(4) This raises the question of whether 'shortness' is restricted just to this one pattern, or whether it is a property of their utterances in general, regardless of their type. In order to answer this question, I calculated the length of two structure types that are comparable to the RC construction—the NP-

complement pattern illustrated in (7a) and the nominalized clause pattern exemplified in (7b), both of which include a head noun preceded by a clause.

(7) a. NP complement
[$_S$ kodomo-ga otona-o tasuketa] hanashi-o kiita
child-Nm adult-Ac rescued story-Ac heard
'I heard the story in which a child rescued an adult.'

b. Nominalized clause
[$_S$ kodomo-ga otona-o tasuketa] koto-o kiita
child-Nm adult-Ac rescued nominalizer-Ac heard
'I heard that a child rescued an adult.'

Table 4 presents the mean average length of these patterns for each group. The mean length of the two structure types is 5.14 for the learner group, compared to 6.52 for the native speakers. Crucially, this is more than twice the mean length of the RCs produced by the learner group (a mere 2.33, as noted in table 3).[(5)] This clearly suggests that 'shortness' is indeed a characteristic of learners' RCs, even though the learners' utterances are in general shorter than those of the native speakers.

Table 4: Mean length of Nominalized clauses + NP complements produced

	No. of content words + postpositions	No. of Nml clauses+ NP-complements	Mean length
Learners	**72 (=55+17)**	14	5.14
Native speakers	202 (=143+59)	31	**6.52**

5.1.2 Effect of weight on production

Now let us consider other ways in which the weight of RCs is manifested in production. We begin by considering the precise content of NPs containing an RC—in particular, whether they include other constituents modifying the head noun. To do this, we divided NPs into three types: those containing a single RC, those containing an RC and another clausal modifier, and those containing an RC and a phrasal modifier. Examples for the last two types are given below.

(8) Examples of Double Clausal Type and RC + Phrase Type

a. Double Clausal Type:
[[$_{RC}$ [$_{RC}$ Ooku-no hito-ga atsumaru] tokoro]-de okita] jiken]
many-Pos person-Nm gather place-Loc occurred incident
'the incident that occurred at the place where many people gather'

b. RC + Phrase Type:
[[$_{RC}$ soko-de hataraiteita] [$_{AP}$ mukankeina] hito]
there-Lc was.working unrelated people
'innocent people who were working there'

Table 5 presents the distribution of these three types of phrases.

Table 5: Three types of NPs containing an RC

	Single RC Type: [RC] N	Double Clausal Type: [RC + clausal modifier] N	RC+ Phrase Type: [RC + phrasal Modifier] N
Learners	30	0	0
Native speakers	35	7	3

None of NPs produced by the L2 learners contained more than a single RC. In contrast, the native speaker group used the two-clause NP type 7 times and the RC + phrasal modifier type 3 times. This asymmetry once again appears to reflect the heaviness of RCs; for L2 learners, the presence of a single RC precludes the occurrence of another modifier in the same NP.

The extracted RCs were also examined with respect to the grammatical role of the head noun in the matrix clause. This was done because some L1 acquisition studies (e.g., Limber, 1973, Diessel & Tomasello, 2000 & 2005) report that the frequency and/or difficulty of RCs varies depending on their location. For instance, Diessel & Tomasello (2000) found that RCs modifying the predicate nominal (PN) in a copular clause and those modifying a nominal in an isolated NP are the two most frequent types of relative clauses in the speech of English speaking children, whereas RCs modifying a subject noun make up the least frequent type. They suggest that the high frequency of RCs modifying a PN can be traced to the fact that sentences containing this type of RC (e.g., *This is the sugar that goes in there*.) are similar to simple sentences in English. However,

Ozeki & Shirai (2007) report that this pattern was not found in the Japanese-speaking children's speech that they examined. They attribute this to the structural properties of Japanese RCs and to the fact that the word order is different from English. Table 6 presents the results.

Table 6: RCs by the syntactic role of the head noun in a matrix clause

	Learners	Native speakers
Subject (SUBJ)	**16 (53 %)**	20 (44%)
Object (OBJ)	5 (17%)	11 (24%)
Oblique (OBL)	2 (07%)	2 (04%)
Predicate Nominal (PN)	0	8 (18%)
Bare NP	2 (07%)	1 (02%)
Others (e.g. adverb, etc.)	5 (17%)	3 (07%)
Total number of RCs	30	45

The learners used no RCs with the predicate nominal in a copular clause and very few RCs with an NP in isolation. This is similar to Ozeki & Shirai (2007)'s L1 finding.

Both for the learner group and for native speakers, the subject was the most frequent type of head for relative clauses. However, this preference was much stronger among the learners than the native speakers (53% versus 44 %).

Also of interest is the fact that in all 16 sentences where an RC modified a subject in the speech of the learner group, the matrix clause was built around a one-place predicate (e.g., an intransitive verb, an adjective, etc). This was the case in just 11 of the 20 patterns of subject-modifying RCs produced by the native speakers. Evidently, the learners have a tendency to use a 'short' predicate when the subject NP includes an RC. This may be another manifestation of the effect of RC weight in that the structural complexity it introduces appears to induce simplifications elsewhere in the sentence. (Interestingly, Diessel & Tomasello (2005) report a similar tendency for the emergence of relative clauses in children acquiring English and German as first languages.)

Although Japanese is an SOV language, the presence of null pronouns and of scrambling does not ensure that the initial constituent of a sentence is always the subject. It is necessary to tease apart the effect of subjecthood and initial

position on the placement of relative clauses. Table 7 presents our findings.

Table 7: Location of RCs

	Initial position	Non-initial position
Learners	**24 (80%)**	6 (20%)
Native speakers	**20 (44%)**	**25 (55%)**

As the figures in Table 7 indicate, the L2 learners—but not the native speakers—predominantly use RCs in NPs that occur at the beginning of a clause. This clearly suggests that the L2 learners prefer to use RCs at the initial position of a clause.

5.2 Analyses of RCs in the OPI data

A total of 29 RCs were extracted from the learners' OPI data and were analyzed in the same manner as those in the narrative data.

5.2.1 Weight of RCs with respect to the length

Table 8 presents the mean length of RCs for the learner group. For purposes of comparison, the figures for the narrative data are also included here.

Table 8: Length of RCs in Learners' OPI data

	No of Content words + particles	No. of RCs	Mean length
OPI data	**76=(62+14)**	**29**	2.62
Narrative data	**79=(59+20)**	**30**	2.33

As their mean clause length of 2.62 indicate, the RCs in the OPI data are short, just as they are in the case of the narratives.

5.2.2 Effect of weight on production

Now let us consider other ways in which the weight of RCs might manifest itself in production. We begin with an examination of the content of NPs containing an RC. As before, NPs are grouped into three types: those containing a single RC, those containing an RC and another clausal modifier, and those containing an RC

and a phrasal modifier. Table 9 summarizes our results.

Table 9: Three types of NPs containing an RC

	Single RC Type: [RC] N	Double Clausal Type: [RC + another clausal modifier] N	RC+ Phrase Type: [RC + phrasal Modifier] N
OPI data	29	0	0
Narrative data	30	0	0

As can be seen here, the results are very similar to those obtained for the narrative data; the presence of an RC precludes the use of other modifiers.

Finally, RCs were examined with respect to where they occur. Table 10 shows the distribution of RCs based on the syntactic role of the head noun in the matrix clause, revealing that subjects are the single most likely head.

Table 10: RCs by the syntactic role of head noun in matrix clause

Syntactic role of head noun in matrix clause	Learners' OPI data	Learners' Narrative data
SUBJ	**9 (31%)**	**16 (53 %)**
OBJ	2 (07%)	5 (17%)
OBL	5 (17%)	2 (07%)
PN	**5 (17%)**	0
Bare NP	2 (07%)	2 (07%)
Others	6 (21%)	5 (17%)
Total number of RCs	29	30

As was the case with the narratives, there was a tendency to use a one-place-predicate in the matrix clause when the subject is modified by an RC. This occurred 8 times out of 9 in the speech of the learners.

Table 11 presents the distribution of RCs with respect to the location of the NP that they modify. As the figures in Table 11 indicate, the results are similar to those from the narratives—the L2 learners produce RCs in NPs that occur at the beginning of a clause more frequently than in a non-initial position.

Table 11: Location of RCs

	Initial position	Non-initial position
OPI	**19 (66%)**	10 (34%)
Narrative	**24 (80%)**	6 (20%)

In sum, as with the narrative data, the learners use RCs primarily to modify the subject head, and they place RCs predominantly at the beginning of a clause, avoiding center-embedded RCs. A similar finding is reported for Japanese children reported in Ozeki & Shirai (2007).

6. Discussion and conclusion

As we have seen, roughly the same picture emerges from the two sets of data. First, the RCs produced by the learners have two characteristics: (i) they are "short" with respect to the number of component words, and (ii) they tend to occur at the beginning of a clause. In contrast, the RCs produced by the native speakers do not exhibit these characteristics. This is consistent with Yamashita's (2002) characterization of NP heaviness in Japanese.

A possible explanation for the 'shortness' of RCs in L2 speech stems from their syntactic complexity, which in turn can be traced to the fact that they contain a gap. (Recall that gaps require considerable pre-planning in Japanese since relative clauses precede the noun that they modify.) Independent evidence for the difficulty of gaps comes from the fact that the learners are capable of producing longer sentential modifiers that do not involve a gap, as noted in our earlier discussion of sentence complements and nominalized clauses.

From this view then, 'shortness' offers learners a way to adjust the structural weight of RCs. Given that the gap cannot be eliminated, reduction to the length of the RC is the only adjustment that speakers can make—other than avoiding relative clauses altogether.

A further important finding has to do with the L2 learners' preference for using RCs in clause-initial position. The results show that, unlike native speakers, the learners prefer the pattern in (9a), in which the 'heavy' constituent is placed at the beginning of a clause.

(9)	a. #[[RC] N]	Constituent	V
	b. #Constituent	[[RC]N]	V

Why should this be?

Given that Japanese RCs are prenominal, their occurrence in the sentence-initial NP allows learners to produce one clause at a time. This creates a lighter burden on working memory than when there are two overlapping clauses, in which case the elements of one clause must be kept in a buffer while the other is being processed (e.g., Hawkins, 1994; Gibson, 1998). Even though the RCs produced by the learners are short, they are nonetheless structurally complex. Placement in sentence-initial position arguably offers learners a way to mitigate this complexity by avoiding the additional burden associated with overlapping clauses.

In conclusion, the complexity of relative clauses affects their length and their distribution as well as the choice of matrix predicate. As mentioned earlier, there is a tendency to have a one-place predicate in the matrix clause. And, of course, complexity places a burden on the incremental planning that takes place during spontaneous production, contributing to the relative infrequency of relative clauses in the speech of Japanese L2 learners.

This is a preliminary report based on a small sample size. Further research has to be done with a larger sample in order to determine how exactly the weight of RCs affects production.

Notes

1. Table A: Participants

Participants	Descriptions	Group makeup	Gender Distribution
Learners of Japanese (n=7)*	All American native speakers of English resided in Hawaii at the time of data collection	2 graduate students 3 undergraduate students 2 business people	4 males & 3 females
Japanese native speakers (n=3)	Native speakers of Japanese from Japan	2 graduate students 1 business person	2 males and 1 female

*There were originally 7 participants, but 1 was excluded from this study because her narrative contained no relative clauses.

2. Narrative task: While looking at a set of four photos depicting the September 11, 2001 attacks on New York City , the participants were asked to describe the events and express their thoughts.
3. Case marking particles are also included.
4. No statistical analyses are done because there are only 3 native speakers. Because Japanese NP-complements are quite often connected to the head noun by the complementizer *toiu*, one might think that the longer length we obtained for the complex NP had to do with the use of this element, but none of the NP-complements produced by the learners included this item.

References

Arnold, J. E., Wasow, T., Losongco, A., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76, 28-55.

Diessel, H., & Tomasello, M. (2000). The development of relative clauses in spontaneous child speech. *Cognitive Linguisitcs* 11, 131-151.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language* 81, 882-906.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.

Hawkins, J. (1994). *A performance theory of order and constituency*. Cambridge : Cambridge University Press.

Kanno, K. (2001). Sentence processing by JSL learners. Paper presented at the Second Language Research Forum 2000, Madison, WI.

Kanno, K. (2007). Factors affecting the processing of Japanese relative clauses by L2 learners. *Studies in Second Language Acquisition*, 29, 197-218.

Kanno, K., Hasegawa, T., Ikeda, K., Ito, Y., & Long, M. (2007). Prior language-learning experience and variation in the linguistic profiles of advanced English-speaking learners of Japanese.' In D. M. Brinton & O. Kagan (Eds.), *Heritage language acquisition: A new field emerging* (pp. 165-180). Mahwah, NJ: Lawrence Erlbaum.

Limber, J. (1973). The genesis of complex sentences. In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 169-85). New York: Academic Press.

Ozeki, H., & Shirai, Y. (2007). The consequences of variation in the acquisition of relative clauses: An analysis of longitudinal production data from five Japanese children. In Y. Matsumoto, D. Oshima, O. Robinson & P. Sells (Eds.), *Diversity and universals in language: Perspectives and implications* (pp.243-270). Stanford, CA: CSLI Publications.

Ozeki, H. (2005). *Daini-gengo ni okeru nihongoshuushokusetu no sanshutsu wa huhenteki shuutoku-nanido-kaisoo ni shitagau ka* [Does the acquisition of noun-modifying constructions in L2 Japanese follow the noun phrase accessibility hierarchy?]. *Acquisition of Japanese as a Second Language, 8,* 64-82.

Saito, H. (2002). Second language acquisition research on Japanese noun modifying clauses. *Japanese Language Education: The state of the art in second language acquisition and instruction research—a guidepost to Japanese language education for the new century*, 45-69.

Schachter, J. (1974). An Error in Error Analysis. *Language Learning*, 24.2, 205-214.

Tamaru Y., Yoshioka, K., & Kimura, S. (1993). Gakushuusha no hatuswa-ni mirareru bunkoozoo no chookiteki kansatsu [a longitudinal development of sentence structures—a study of JSL adult learners]. *Nihongo Kyooiku*, 81, 43-54.

Yamashita, H. (2002). Scrambled sentences in Japanese: Linguistics properties and motivations for production. *Text*, 22, 597-633.