UCLA Working Papers in Linguistics, no.13, December 2005
Papers in Psycholinguistics 2, Okabe and Nielsen (eds.)

PROCESSING RELATIVE CLAUSES IN JAPANESE[1]

TOMOKO ISHIZUKA
zukka@ucla.edu

This paper reports a self-paced reading study that compares the processing difficulity between object-gap and subject-gap relative clauses (RCs) in Japanese. The higher complexity of object-gap RCs compared with subject-gap RCs in SVO languages with postnominal relatives, such as English and French, is a well-established phenomenon in sentence processing. Two theories have been proposed to account for this phenomenon. One is in terms of resource, such as temporary storage or integration costs. The other is in terms of depth of embedding of the extracted element. Both accounts can explain the pattern of data from Englih and French, since the amount of resource in terms of linear distance is correlated with the depth-of-embedding. In this paper, Japanese, an SOV language with prenominal relatives, was tested in order to differentiate these two cost metrics. The resource theory predicts object-gap RCs to be easier due to fewer intervening words between the gap and the head noun, whereas the depth-of-embedding theory predicts subject-gap RCs to be easier. The results show that subject-gap RCs are easier than object-gap RCs in Japanese. The results are compatible with the depth-of-embedding theory. Other possible accounts are considered besides the depth-of embedding theory.

1. INTRODUCTION

One of the well-established complexity phenomena in sentence processing is the higher processing difficulity of object-gap relative clauses (RCs) such as (1) compared with subject-gap RCs such as (2) in Subject-Verb-Object languages, including English (e.g. King and Just 1991, Gibson 1998, 2000) and French (Holmes and O'Regan 1981).

(1) English object-gap RC
The reporter [who the senator attacked *e*] had a bad reputation.
(2) English subject-gap RC
The reporter [who *e* attacked the senator] had a bad reputation.

In (1), the relative pronoun 'who' is extracted from the object position of the RC, leaving a gap in its extraction site, whereas the same

[1] I would like to thank Edward Gibson and Kentaro Nakatani for help of conducting the research and for invaluable suggestions and feedback. I would also like to thank Carson Schütze and Paul Hagstrom for invaluable comments. My gratitude also goes to the viewers of the poster at the 16th CUNY Conference on Human Sentence Processing, who gave me insight and comments. All remaining errors are my own.

pronoun is extracted from the subject position in (2). The object-gap RC has been shown to be more complex by various measures including phoneme-monitoring, on-line lexical decision, reading times, and response-accuracy to probe questions (Holmes 1973, Hakes et al., 1976, Wanner and Maratsos 1978, Holmes and O'Regan 1981, Ford, 1983, Waters et al. 1987, King and Just 1991, discussed in Gibson 1998, 2000). The complexity of (1) cannot be explained by lexical information (e.g., word frequencies), by the real-world plausibility of the meaning of the sentence, or by the complexity of the discourse context, because all of these factors are the same in sentence (2), yet this sentence is less complex to process. As a result, the complexity of a sentence like (1) must be due to properties of its syntax: the integration of a dependent to a head occurs.

2. Literature Review and Research Question

Assuming that the increased complexity involved in object-gap RCs in English is due to an increase in distance over which the extraction occurs, there are two clear alternative hypotheses concerning how this distance is measured. One is that distance is measured in terms of resources, either for temporary storage or for syntactic integration (Chomsky and Miller, 1963, Gibson 1998, 2000, Lewis 1996). For example, according to the Dependency Locality Theory (DLT) proposed by Gibson (2000), sentence comprehension involves two components of computational resource use: (1) storage resources: keeping track of the incomplete structural dependencies in the current structure and (2) integration resources: connecting an incoming word into the current structure. As an explanation for the dependence on distance in integration-based complexity, Gibson (2000) proposes that the process of sentence comprehension is activation-based. Integrating a newly input maximal projection with a previous syntactic category, headed by *h1,* involves retrieving aspects of *h1* from memory. In an activation-based framework, this process involves re-activating *h1* to a target threshold of activation. Due to a limited quantity of activation in the system, *h1*'s activation will decay as intervening words are processed and integrated into the structure for the input. Thus, the difficulty of the structural integration depends upon the complexity of all aspects of the integrations that have taken place in the interim since *h1* was last highly activated[2]. The integration cost, which consists of discourse processing cost and structural integration cost, is quantified as follows:(pp.104-105)

[2] I assume that 'being highly activated' occurs when h1 was newly processed and integrated to a previously constructed syntactic structure. Further, an empty category is 'highly activated' when a reader detects the presence of an empty category and integrates it to the syntactic structure they have already constructed.

DLT *discourse processing cost* (the cost associated with accessing or constructing the discourse structure for the maximal projection of the input word head *h2*)
1 energy unit (EU) is consumed if *h2* is the head of a new discourse referent; 0EUs otherwise.

DLT *structural integration cost*
The structural integration cost associated with connecting the syntactic structure for a newly input head *h2* to a projection of a head *h1* that is part of the current structure for the input is dependent on the complexity of the computations that took place between *h1* and *h2*. For simplicity, it is assumed that 1EU is consumed for each new discourse referent in the intervening region.

According to the DLT, the maximal integration cost incurred during the processing of (1) occurs at the point of processing *attacked* and *had*. At the point of *attacked*, the cost is 3 EUs: 1EU for the construction of the new discourse referent, the event verb *attacked*, 0EUs for integrating the object-empty category to the verb *attacked* and 2EUs for the structural integration of the object-position empty category to the preceding RC pronoun *who.* Two discourse referents were introduced in the interim —the NP *the senator* and the event referent *attacked* — leading to an integration cost of 2EUs. At the point of *had*, the cost is also 3EUs: 1EU for the construction of the new discourse referent *had*, and 2EUs for the structural integration of the verb *had* to the subject NP *the reporter*[3]. Two new discourse referents — *the senator* and *attacked* — intervene, leading to a cost of 2EUs.

As for the subject-gap RC (2), the maximal integration cost occurs at only one point, the point of processing *had*: 3EUs. At the point of *attacked*, where the maximal cost occurs at (1), the cost is only 1EU in (2): for the construction of the new discourse referent, the event verb *attacked*. No structural integration cost needs to be consumed because the integration of the subject-position empty category to the preceding RC pronoun *who* is local**;** there is no new discourse referent in the interim. Thus, comparing the maximal integration costs between the two constructions, the DLT predicts that object-gap RCs are more difficult to process than subject-gap RCs in English.

[3] The subject NP 'the reporter' was re-activated when RC pronoun 'who' was integrated to the constructed structure. However, I assume that the last time the subject NP 'the reporter' was *highly* activated was the time it was newly input.

The DLT also predicts object-gap RCs to be more difficult to process than subject-gap RCs with respect to the storage cost. There is a greater storage cost in processing the object-gap RC in (1) than the subject-gap RC in (2) as soon as the first word following 'who' is processed in each. Four syntactic heads are expected after processing 'the reporter who the' in (1): a noun for the determiner 'the', a verb for the matrix clause, a verb for the relative clause, and an empty noun element associated with the operator. On the other hand, only two heads are necessary after processing the words 'the reporter who attacked' in (2): a noun for the object of 'attacked', and a verb for the matrix clause. The predictions that the DLT made by using the integration cost and by using the storage cost are the same, and Gibson (2000) notes that many resource complexity effects can be accounted for using the integration cost alone (p.102).

The other alternative hypothesis for how distance is measured for the purposes of computing complexity is in terms of depth of embedding of the extracted head (O'Grady 1987, O'Grady et al. 2000, Keenan and Comrie[4] 1977, Keenan and Hawkins 1987). For example, O'Grady et al. (2000) propose a depth-of-embedding-based theory, which they call Structural Distance Hypothesis (SDH). This hypothesis is defined as follows (p.2):

The Structural Distance Hypothesis

> The distance traversed by a syntactic operation, calculated in terms of the number of nodes crossed, determines a structure's relative complexity.

According to the SDH, the number of nodes between the gap and the head for (1) is two (S and VP). In contrast, the number of nodes between the gap and the head for (2) is one (S). Therefore, the prediction of SDH is that it is harder to process (1), the object-gap RC.

---

[4] Keenan and Comrie (1977) propose that languages vary according to an Accessibility Hierarchy (AH) with respect to which NP positions can be relativized, where the relativizability of each position implies the relativizability of all higher positions, an implicational universal.

Accessibility Hierarchy

Subject > Direct Object > Indirect Object > Major Oblique case NP > Genitive > Object of Comparison

Here, ">" means 'is more accessible than.' (P.66)

Keenan and Comrie claim that the AH universally determines the degree of accessibility to RC formation. So the possibility of relativizing decreases as we go down the AH, and in that sense the further we descend the AH the harder it is to relativize (p.68). Therefore, according to the AH, object-gap RCs are more marked than subject-gap RCs.

Both the resource theory and the depth-of-embedding theory can explain the patterns of data from the languages that have been the main focus of the attention in the processing literature to date. We cannot determine which of these accounts is correct by examining head-initial SVO languages like English, where the distance in terms of resources and in terms of depth-embedding are correlated. However, by examining typologically different languages such as Japanese, we can differentiate these two distance metrics.

The question being addressed in this paper is: what makes some extractions more difficult than others? In order to answer this question, this paper examines on-line processing data from Japanese, where the theories outlined above make different predictions. Japanese is a head-final, Subject-Object-Verb language. RCs precede their head nouns, unlike RCs in English and French, which follow their head noun. Therefore, the depth of embedding and the linear distance between the gap and the head are not correlated in Japanese.

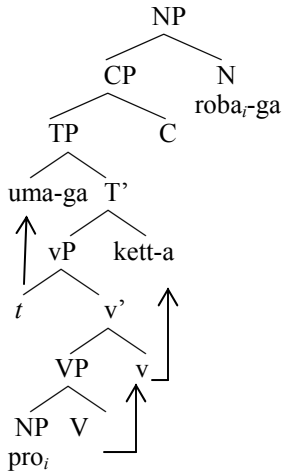
3. Syntax of Japanese Relative Clauses

Japanese relative clauses have been studied extensively (cf. Kuno 1973, McCawley 1976, etc) and various characteristics specific to Japanese have been observed, such as a lack of relative pronoun, an insensitivity to island effects and the existence of gapless relative clauses (see details for Kuno 1973, and Murasugi 2000). In order to account for these empirical observations, it has been argued that Japanese relative clauses do not involve movement but have a null pronoun in the position of the gap coindexed with the head noun (Perlmutter 1972, Hoshi1995, Murasugi 2000, Fukui and Takano 2000, etc). Following their insights, I assume throughout this paper that Japanese relative clauses do not involve movement but rather have a null pronoun co-indexed with the relative head noun.

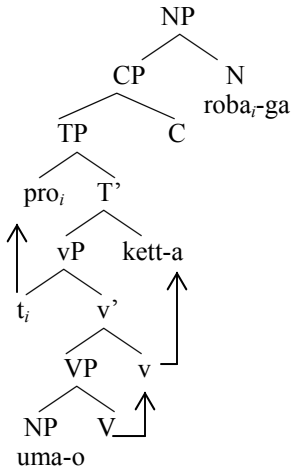
The following are examples of Japanese relative clauses and their syntactic representations.[5] For notational purposes, the null pronouns in the subject and object position are notated as empty categories, "e" for short.

(3) Object-gap RC: [uma-ga e_i ketta] roba$_i$-ga sinda.
horse-nom kicked mule-nom died
'The mule that the horse kicked died.'

[5] Due to the lack of determiners, it is often assumed in the literature that Japanese does not have a D projection. See Fukui and Takano (2000) for details.



(4) Subject-gap RC: [e_i uma-o ketta] roba$_i$-ga sinda.
horse-acc kicked mule-nom died
'The mule that kicked the horse died.'



There is one discourse referent between the gap and the antecedent in the object-gap RC (3). In contrast, there are two discourse referents between the gap and the antecedent in the subject-gap RC (4). Since the subject-gap RC has more intervening discourse referents between

the gap and the antecedent, the DLT predicts that subject-gap RC is more difficult to process than object-gap RC in terms of integration in Japanese. Note that the DLT is a theory to account for the cost of integrating two elements, not only for a dependent to a head, but for a pronominal referent to its antecedent (p.95). The same result is predicted in terms of storage cost. At the point where the first word is processed, there is a greater storage cost for the subject-gap RC than the object-gap RC. Specifically, the object-gap RC (3) requires one head: a verb to form a grammatical sentence. In contrast, the subject gap RC (4) requires two heads: a subject and a verb to form a grammatical sentence[6]. Since the rest of the storage cost required for both constructions is the same, the overall storage cost is greater for the subject-gap RC. Gibson (2000:102) notes that "… it turns out that many resource complexity effects can be explained using integration cost alone." Thus, I will mainly focus on the integration cost when discussing the predictions made by the DLT in the following sections. Note that the predicted patterns are similar when storage cost is also taken into consideration (p. 105).

The SDH, on the other hand, predicts that the object-gap RC (3) is more difficult than the subject-gap RC (4). As shown in the above syntactic representations, objects are more deeply embedded than subjects. Assuming that the prediction of the SDH is applicable to the dependencies between a pronominal referent and its antecedent, associating a more deeply embedded pronoun to its antecedent is more costly than associating a less deeply embedded pronoun to its antecedent.

In summary, the two theories, the DLT and the SDH, make opposite predictions for processing of Japanese RCs. The DLT predicts subject-gap RCs to be more difficult to process, whereas the SDH predicts object-gap RCs to be more difficult to process in Japanese.

4. EXPERIMENT

4.1. *Conditions*

A self-paced, word-by-word reading experiment was conducted to compare the processing complexity between object-gap RCs and subject-gap RCs in Japanese[7]. Two pairs of conditions were tested, as

[6] Although Japanese allows topic drop and it is possible for readers to process the lack of subject as null subject, such a possibility is not considered here because no items were presented with contexts. See footnote 19 for details.

[7] To be more precise, 'word' here means a content word followed by particles such as case-markers and inflectional morphemes.

exemplified in (5) and (6). The RCs to be compared in (5) were singly embedded, whereas the RCs in (6) were doubly embedded.

(5) a. Japanese singly embedded object-gap RC
[S N2-nom [VP $\boldsymbol{e}_i$ V2]] N3-dat-top$_i$
resources: 2EUs
depth of embedding: two nodes(S &VP)

b. Japanese singly embedded subject-gap RC
[S $\boldsymbol{e}_i$ N2-acc V2] N3-dat-top$_i$
resources: 3EUs
depth of embedding: one node(S)

(6) a. Japanese doubly embedded object-gap RC
[S[S N1-nom [VP $\boldsymbol{e}_i$ V1]]N2-nom$_i$ [VP $\boldsymbol{e}_j$ V2]] N3-dat-top$_j$
outer RC: resources: 2EUs
embedding: two nodes(S & VP)

b. Japanese doubly embedded subject-gap RC
[S $\boldsymbol{e}_j$ [S $\boldsymbol{e}_i$ N1-acc V1]N2-acc$_i$ V2] N3-dat-top$_j$
outer RC: resources: 5EUs
embedding: one node(S)

Doubly embedded versions were tested in addition to singly embedded versions in order to magnify the differences to a measurable level, as the predicted effects might be difficult to measure in singly embedded versions.

Effects of perspective shift (MacWhinney 1977, 1982, MacWhinney and Pleh 1988) and case-matching (Sauerland and Gibson 1999) are controlled for in the above conditions. MacWhinney and Pleh (1988) claim that processing resources are required to shift the perspective of a clause, where the perspective of a clause is taken from the subject of the clause. According to this theory, a subject-modifying object-gap RC as in (1) requires two perspective shifts: (1) from the perspective of the matrix subject to the subject of the RC and (2) from the perspective of the subject of the RC back to the matrix subject, after the RC is processed. Processing the subject-gap RC in (2) requires no perspective shifts, because the matrix subject is also the subject of the RC, so that both clauses come from the same perspective. Thus, the object-gap RC is more complex than the subject-gap RC.

Similarly, Sauerland and Gibson (1999) propose a case matching principle, which predicts an increase in processing complexity when the case of a relative pronoun and its associated head noun do not match. In order to control for these effects, all of the above conditions are constructed so that the extracted element does not match the case of

its head noun for either the nominative or the accusative gap-site ("case clash" conditions).

4.2. *Predictions*

Resource theories predict higher complexity for subject-gap RCs, (5b) and (6b), as compared with object-gap RCs, (5a) and (6a). The following tables give the word by word predictions of the DLT for the test items.

Table 1. Word by word predictions of the DLT: (5a) singly embedded object-gap RC

	Input word		
Cost type	N2-nom	V2	N3-dat-top
New discourse referent	1	1	1
Structural integration	0	0[8]	1
Total (in EUs)	1	1	2

According to the DLT, the maximal integration cost incurred during the processing of (5a) is 2EUs, occurring at the point of processing N3.

- 1EU for the construction of N3.
- 0EUs for the structural integration of the object N3 to V2. No new discourse referents intervene.
- 1EU for structural integration coindexing the object null pronoun with the antecedent head noun[9]. One new referent (V2) intervenes.

Table 2. Word by word predictions of the DLT: (5b) singly embedded subject-gap RC

	N2-acc	V2	N3-dat-top
New discourse referent	1	1	1
Structural integration	0[10]	1[11]	2
Total (in EUs)	1	2	3

[8] As soon as a reader encounters the transitive verb V2, s/he detects the lack of object and creates an empty category for it. The integration of an object empty category to N2-nom costs 0EUs due to the local integration. Integrating an object-position empty category to V2 also costs 0EUs because the integration is local. Integrating V2 to N2-nom also costs 0EUs because no new discourse referent intervenes. Therefore, the structural integration cost at V2 adds up to 0EUs.

[9] I assume that the object-position empty category is created before the integration of V2 to the previously constructed structure occurs. Therefore, V2 is considered to be an intervening referent.

[10] As soon as a reader encounters the first word, N2-acc, s/he detects the lack of subject and creates an empty category for it. Then N2-acc is integrated to the subject-position empty category, which cost 0EUs because the integration is local.

[11] 0EUs for the local integration of N2-acc to V2. 1EU for integrating the subject-position empty category to V2: 1 new discourse referent (N2-Acc) intervenes.

The maximal integration cost incurred during the processing of (5b) is 3EUs, occuring at the point of processing N3.

- 1EU for the construction of N3.
- 0EUs for the integration of N3 to V2.
- 2EUs for the integration coindexing the subject-position empty category with the antecedent head noun. Two new referents (N2, V2) intervene.

Table 3. Word by word predictions of the DLT: (6a) doubly embedded object-gap RC

	N1-nom	V1	N2-nom	V2	N3-dat-top
New discourse referent	1	1	1	1	1
Structural integration	0	0[12]	1	0	1
Total (in EUs)	1	1	2	1	2

In (6a), the maximal integration cost is 2EUs, occurring at the point of processing N2 and N3.

- 1EU for the construction of N2 or N3.
- 0EUs for the integration of N2 to V1, or N3 to V2.
- 1EU for the integration coindexing each object null pronoun with its antecedent head noun. One new referent (V1 or V2, respectively) intervenes.

Table 4. Word by word predictions of the DLT: (6b) doubly embedded subject-gap RC

	N1-acc	V1	N2-acc	V2	N3-dat-top
New discourse referent	1	1	1	1	1
Structural integration	0	1	0[13]	1[14]	2
Total (in EUs)	1	2	1	2	3

In (6b) the maximal integration cost for the outer clause is 3EUs, occurring at the point of processing N3.

- 1EU for the construction of N3.
- 0EUs for the integration of N3 to V2.
- 2EUs for the integration coindexing the subject null pronoun with the antecedent head noun. The matrix clause subject-position empty category is created as soon as a reader

[12] Again I assume that the object-position empty categories are created right after a reader encounters V1 or V2 respectively, and the integrations of V1 and V2 take place after the creation of the empty categories.

[13] As soon as a reader encounters N2-acc, s/he notices the lack of matrix clause subject and creates a matrix clause subject-position empty category.

[14] Integrating V2 to the matrix clause subject-position empty category costs 1EU, since one new discourse referent intervenes: N2-Acc.

encounters N2-acc (see footnote 15). Therefore, two new referents (N2 and V2) intervene.

In contrast, the depth-of-embedding hypothesis makes the reverse prediction since an object is always more deeply embedded than a subject regardless of word order. Namely, it predicts (5a) is more difficult to process than (5b), and (6a) is more difficult to process than (6b).

Two control conditions for (6) were also tested in order to see the effects of case-matching and perspective shifts in the most embedded clauses. In the following control conditions (7), null pronouns match the case of their antecedents for both the nominative and the accusative gap-site in the most embedded clauses, unlike the conditions in (6). By comparing the reading time at region N2 between (7a) and (7b), we can see whether the results gained from the above conditions are attributed to the case matching/clashing conditions.

(7) a. Japanese case matching subject-gap RCs
[S $\boldsymbol{e_i}$ N1-acc V1] N2-nom$_i$ V2 N3-dat-top
inner RC: resources: 3EUs
embedding: one node(S)

b. Japanese case matching object-gap RCs
[S N1-nom [VP $\boldsymbol{e_i}$ V1]] N2-acc$_i$ V2 N3-dat-top
inner RC: resources: 2EUs
embedding: two nodes (S &VP)

The DLT predicts (5a) to be processed faster than (5b) during the processing of N2. On the other hand, the SDH predicts the reverse results.

Furthermore, by comparing (7a) and (6b), and (7b) and (6a), we can see whether case clash conditions are more difficult to process than case match conditions. As stated above, the Perspective Shift Theory and the Case Matching Principle predict that (7a) and (7b) should be easier than (6b) and (6a).

4.3. *Participants*

Forty-four native speakers of Japanese were paid to participate in the experiment. They were residents of the Boston area at the time of the experiment and were naïve as to the purposes of the study.

4.4. *Materials*

Twenty-four sets of sentences like the following were constructed, each with two singly embedded conditions, two doubly-embedded conditions and two case matching.

(8) a. Singly embedded object-gap RC :
[kuruma-ga e_i tuisekisita] ootobai-ni-wa$_i$
car-nom e_i *chased* *motorbike-dat-top*$_i$
kookoosei-ga notteita.
high_school_student-nom rode.
'A high school student was on the motorbike which the car chased.'

b. Singly embedded subject-gap RC :
[e_i kuruma-o tuisekisita] ootobai-ni-wa$_i$
e_i *car-acc* *chased* *motorbike-dat-top*$_i$
kookoosei-ga notteita.
high_school_student-nom rode.
'A high school student was on the motorbike which chased the car.'

(9) a. Doubly embedded object-gap RC :
[[torakku-ga e_i oikosita] kuruma-ga$_i$ e_j tuisekisita]
truck-nom e_i *passed* *car-nom*$_i$ e_j *chased*
ootobai-ni-wa$_j$ kookoosei-ga notteita.
motorbike-dat-top$_j$ *high_school_student-nom rode.*
'A high school student was on the motorbike which the car which the truck passed chased.'

b. Doubly embedded subject-gap RC :
[e_j [e_i torakku-o oikosita] kuruma-o$_i$ tuisekisita] ootobai-ni-wa$_j$
e_j e_i *truck-acc passed car-acc*$_i$ *chased motorbike-dat-top*$_j$
kookoosei-ga notteita.
high_school_student-nom rode.
'A high school student was on the motorbike which chased the car which passed the truck.'

(10) a. Subject-gap inner RC in case matching condition :
[e_i torakku-o oikosita] kuruma-ga$_i$ tuisekisita ootobai-ni-wa
e_i *truck-acc passed* *car-nom*$_i$ *chased* *motorbike-dat-top*
kookoosei-ga notteita.

high_school_student-nom rode.
'A high school student was on the motorbike which the car which passed the truck chased.'

b. Object-gap inner RC in case matching condition :
[torakku-ga e_i oikosita] kuruma-o$_i$ tuisekisita ootobai-ni-wa
truck-nom e_i *passed car-acc*$_i$ *chased motorbike-dat-top*
kookoosei-ga notteita.
high_school_student-nom rode.
'A high school student was on the motorbike which chased the car which the truck passed.'

The critical region of comparison in the singly embedded versions in (8) consists of the first three words: N1 V1 N2. The critical region in the doubly embedded versions in (9) consists of the first five words: N1 V1 N2 V2 N3. The critical region in the control case-matching versions in (10) consists of the first three words: N1 V1 N2. Reversible RCs were chosen and each of these comparisons involves the same words in a different order, so lexical frequency is controlled overall.

A plausibility norming survey was also conducted in order to control the difference with respect to the real-world plausibility of the sentences. The target sentences were split into six lists balancing all factors in a Latin-Square design. Each list was combined with sixty-five fillers of various types. In order to control the accessibility of the referents, only definite descriptions are used in the sentences[15] (Warren and Gibson 1999). The stimuli were pseudo-randomized separately for each participant so that at least one filler item intervened between any two targets.

4.5. *Procedure*

The task was word-by-word self-paced non-cumulative reading using a moving window display (Just, Carpenter & Wooley 1982). Linger 1.7 by Doug Rohde was the software used to run the experiments. Sentences were shown using Japanese characters. Stimulus segments initially appeared masked with dashes, and participants pressed the space bar to reveal each subsequent region of the sentence causing all

[15] It has been established in the discourse processing literature that the difficulty of processing an NP depends on the lexical accessibility of the referent of the NP. The less accessible (i.e., more infrequent or less recently mentioned) the referent of an NP is in the discourse, the more resources are required to find or construct it (Warren & Gibson 1999). Entities or individuals that serve as topics, which are usually referred to with pronouns, are highly accessible, so they require a small amount of resources to access. New or unfamiliar entities or individuals, such as proper names and definite descriptions, require more resources to access.

other regions to revert to dashes. The regions consist of *bunsetsu* (a content word followed in some cases by inflectional morphemes and case particles). The segmentation in sentences (8)-(10), indicated with spaces, was the actual segmentation used in the experiment. The reading time (RT) the participant spent on each word was recorded as the time between key-presses.

At the end of each sentence, a yes-no question asking about information contained in the preceding sentence appeared on a new screen. When the participants gave an incorrect answer, a sentence in Japanese meaning "Sorry, your answer was incorrect" flashed briefly on the screen. No feedback was provided for correct responses. Participants were asked to read sentences at a natural rate and to be sure that they understood what they read.

The experimental trials were preceded by a screen of instructions and seven practice trials. All sentences were presented on a single line. The experiment took participants approximately 20 minutes.

4.6. *Plausibility Norming Survey*

A questionnaire study was also conducted in order to control for potential plausibility differences between the three pairs of conditions. Twenty five native Japanese-speaking participants who did not take part in the self-paced reading experiment completed the survey. The items tested in this questionnaire consisted of the simple transitive clauses that made up each RC. For the singly embedded version, the materials consisted of one simple SOV clause in each version. For the doubly embedded versions, there were two simple clauses for each item. One of them was same as the SOV clause in the singly embedded version. Therefore, for each set the following four simple clauses were rated.

(11) a. <u>One clause object-gap control</u>:
kuruma-ga ootobai-o tuisekisita
car-nom motorcycle-acc chased
'The car chased the motorcycle.'

b. <u>One clause subject-gap control</u>:
ootobai-ga kuruma-o tuisekisita
motorcycle-acc car-acc chased
'The motorcycle chased the car.'

(12) a. <u>Two clause object-gap control</u>:
torakku-ga kuruma-o oikosita

truck-nom car-acc passed
'The truck passed the car.'

b. Two clause subject-gap control:
kuruma-ga torakku-o oikosita
car-nom truck-acc passed
'The car passed the truck.'

Participants rated the plausibility of these sentences on a scale of 1 (natural) to 5 (unnatural). They were asked to judge the naturalness in the real world of the events described in the sentences, that is, how likely they were to occur.

The results of the survey were that four of the twenty-four items were found to be significantly more plausible ($p<.05$ by t-test) in one version. These four items were therefore omitted from the reading time analyses. The remaining twenty items were matched for plausibility across all versions (means: 2.18 for (9a), 2.05 for (9b), 2.16 for (10a), 2.05 for (10b)).

4.7. *Data Analysis*

The results were analyzed using Lingalyzer 1.1, an analysis program written by Doug Rohde. The four items for which one version was less plausible than another were omitted from analyses, leaving 20 items to be analyzed. Analyses were conducted on comprehension question response accuracy and reading times.

Residual reading times per region (Ferreira & Clifton, 1986) were derived by subtracting from raw reading times each participant's predicted time to read regions of the same length (measured in number of characters), which in turn was calculated from a linear regression across all of a participant's sentences in the experiment. Thus a typical word will be read at 0 ms of normalized reading time, whereas words read quickly will have negative residual reading times, and words read slowly will have positive residual reading times.

The residual reading times were trimmed so that data points beyond three standard deviations from the relevant condition x region cell mean of all the subjects were discarded, which corresponded to less

than 2.5% of the test data. The means and analyses presented below are based on the trimmed residual reading times[16].

5. RESULTS

5.1. *Comprehension Question Performance*

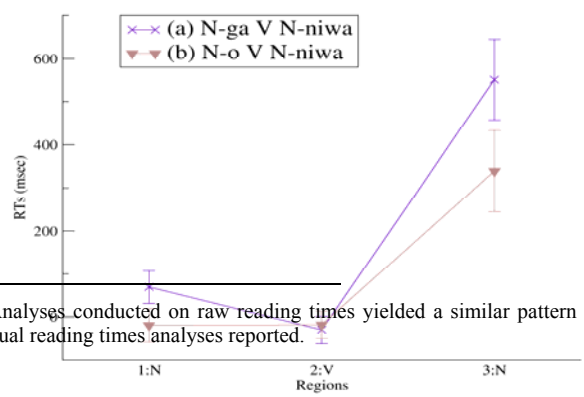
The percentages of correct answers for each condition are presented in Table 6. Although comprehension question performance tended to be better in the singly embedded subject-gap sentences than in the singly embedded object-gap sentences, this difference did not reach significance (Fs< 2.10, Ps>0.16). In the doubly embedded RCs, performance tended to be better in the subject-gap versions, but this effect did not reach significance either (Fs<2.0, Ps>0.16). Comprehension performance for the target items was relatively low (71.1% overall), probably due to (1) the difficulty of the object-gap versions, and (2) the complexity of the discourse contexts in doubly embedded versions. Mean performance on the filler items was 87.0%, thus the participants were paying attention in the task.

Table 2. Mean (standard error) comprehension question performance in percent correct by condition

	Singly embedded RC	Doubly embedded
Object extraction	71.3%(3.7)	65.8%(3.9)
Subject extraction	78.5%(3.4)	71.7%(3.7)

5.2. *Reading Times*

Due to the complexity of the target items, all items were analyzed independent of whether the comprehension questions were answered correctly. Figure 1 plots mean residual reading times per region in the singly-embedded RCs in (5).



[16] Analyses conducted on raw reading times yielded a similar pattern of results as the residual reading times analyses reported.

Figure 1. Plot of mean (standard error) residual reading times per word for the singly embedded conditions in (5a) and (5b)

An ANOVA for region 1N revealed a near significant difference ($F_1(1,44) = 5.60$ $p < 0.03$; $F_2(1,19) = 3.90$ $p < 0.06$). In region 2V, no significant difference was detected (Fs < 0.5). In region 3N, the extracted head noun, subject-gap RCs were processed faster than object-gap RCs, but this effect was significant only in the item analysis ($F_1(1,44) = 2.58$ $p = 0.12$; $F_2(1,19) = 4.60$ $p < 0.05$).

Figure 2 plots mean residual reading times per word by region by participants in the doubly embedded conditions in (6a) and (6b).

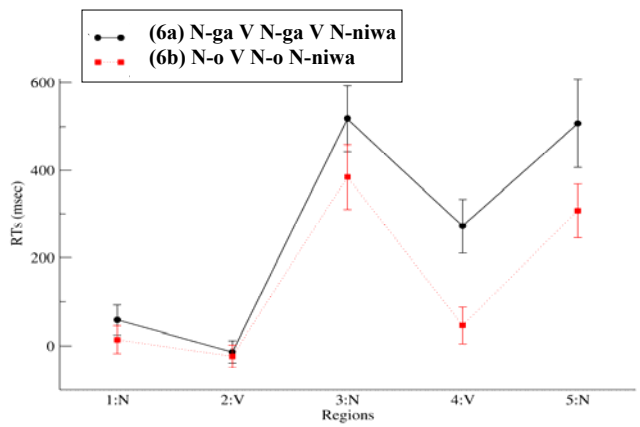


Figure 2. Plot of mean (standard error) raw reading times per word for the doubly embedded conditions in (6).

For the doubly-embedded structures like (6), the same results obtained. In region 1N, there was no significant difference (Fs <1.22, Ps > 0.28). In region 2V, there was also no significant difference (Fs < 0.05, Ps > 0.8). In region 3N, the subject-gap RCs tended to be read faster than object extracted RCs, but this effect did not reach significance ($F_1(1,44) = 2.81$, $p = 0.10$; $F_2(1,19) = 0.94$, $p = 0.35$). In region 4V and 5N combined as one region, an ANOVA revealed that the subject-gap RCs were processed faster than the object-gap RCs for both participants and items analyses ($F_1(1,44) = 8.778$, $p = 0.005$; $F_2(1,19) = 12.186$, $p = 0.002$).

Figure 3 plots mean residual reading times per word by region in the case matching conditions exemplified in (7). In the case matching conditions, there was no significant difference on the first word 1N (Fs<0.38, Ps> 0.54). In region 2V, there was a nearly significant difference ($F_1(1,44) = 1.55$, $p = 0.22$; $F_2(1,19) = 4.23$, $p = 0.054$). In region 3N, the extracted head noun, the subject-gap RCs tended to be read faster than object-gap RCs, but this effect did not reach significance ($F_1(1,44) = 2.81$, $p = 0.10$; $F_2(1,19) = 0.94$, $p = 0.35$).

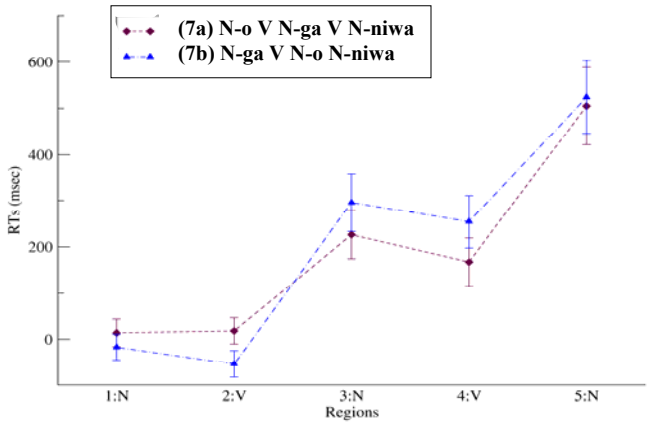


Figure 3. Mean residual reading times per word by region for the case matching conditions in (7)

Turning now to the case match vs. case clash conditions, Figure 4 plots mean Residual Reading Times per word by region for case matching and case clashing conditions in (6) and (7).

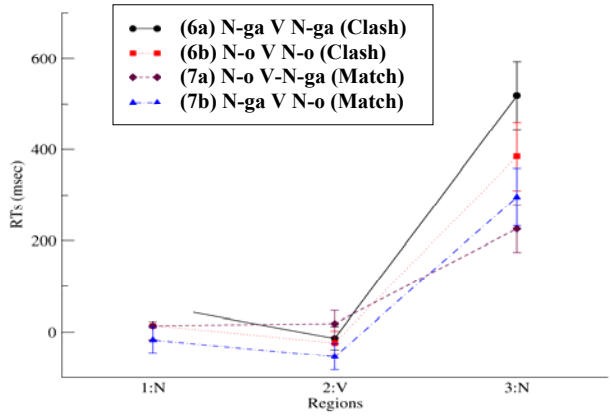


Figure 4. Mean residual reading times per word by region for the case matching and case clashing conditions in (6) and (7)

There was no significant difference at region 1N (Fs < 1.4, Ps > 0.2). In region 2V, there was also no significant difference (Fs < 0.11). In region 3N, case matching conditions were processed significantly faster than case-clashing conditions ($F_1(1,44) = 7.69$, $p < 0.01$; $F_2(1,19) = 6.93$, $p < 0.02$). We collapsed the object-gap RCs and subject-gap RCs together in order to compare the case-matching effects.

6. Discussion

The above results demonstrate that object-gap RCs are significantly more complex than subject-gap RCs in the item analysis for the singly embedded conditions. For the doubly embedded conditions, the object-gap RCs are significantly more complex than subject-gap RCs in both participant and item analyses[17]. Therefore, we can conclude that the results of this experiment show that object-gap RCs are more complex than subject-gap RCs in Japanese. The reaction time data in comparisons involving singly and doubly embedded conditions and case-matching conditions provided the evidence for this observation, with the response accuracy data providing some additional support.

These results support the depth-of-embedding hypothesis over the resource account. However, this runs counter to the results of a recent study of Chinese RCs by Hsiao and Gibson (2003), which show higher complexity for subject-gap RCs. Although Chinese is an SVO language, relative clauses precede their head nouns, like RCs in Japanese, as shown in the following examples (p. 6).

(13) a. <u>Chinese object-gap RC</u>
Fuhao yaoching e_i de guanyusan$_i$ shinhuaibugui danshi
tycoon invite gen official have_bad_intentions but
shanyu yintsang[18]
good_at hiding
'The official who the tycoon invited has bad intentions but is good at hiding them.'

b. <u>Chinese subject-gap RC</u>
e_i yaoching fuhao de guanyusan$_i$ shinhuaibugui danshi

[17] This significance was detected only at the relative head of the outer clause. In the region where the head noun of the inner clause is read, the subject-gap RCs tended to be read faster than object-gap RC, but this effect did not reach significance.
[18] *gen* is a genitive marker.

invite tycoon gen official have_bad_intentions but
shanyu yintsang
good_at hiding
'The official who invited the tycoon has bad intentions but is good at hiding them.'

The DLT predicts that subject-gap RCs should be harder in Chinese while the SDH predicts that object-gap RCs should be harder. Hsiao and Gibson's results cannot be explained by the depth-of-embedding hypothesis.

An alternative account can be provided in order to explain the data gained in the present experiment. There is always a temporary ambiguity such that the object-gap RCs can initially be interpreted as being a main clause in Japanese RCs, which requires reanalysis at some point. The higher complexity of object extraction might be attributed to this reanalysis effect. When a transitive verb (e.g., respect) is encountered, this main clause reading must be given up in favor of an RC analysis because of the absence of an object. On the other hand, the subject-gap RCs have no such ambiguity[19]. In subject-gap RCs, the first word the participants encounter is the object. A missing subject allows them to have a relative clause reading rather than a main clause reading from the very first word[20]. Therefore, it is highly possible that the above results are due to the confounding factor of this reanalysis effect. Namely, the higher complexity of subject extractions might be canceled out due to the reanalysis effect of the object extraction, which results in higher complexity of object extractions.

In order to control for the potential temporary ambiguity and the reanalysis effect, further research must be conducted. One way to avoid the potential temporary ambiguity is by setting up a context where the participants have relative clause readings available as soon as they encounter the first word for both conditions: subject-gap RCs and

[19] One might argue that it also possible to have a main clause reading for subject-gap RCs due to the topic drop property of Japanese. Japanese allows a null discourse-based pronominal to be used in argument position when the referent is well established in the context. However, in our test items, all sentences were presented in a null context, and none of the fillers contained any null discourse-based pronominals. Thus, it is unlikely that participants analyzed the target stimuli as containing such topic pronominals.

[20] It is also possible that readers still mainined the matrix clause reading after processing the sentence initial object, expecting to encouter a subject as a next word. If readers indeed processed an object-scrambling construction (O S V), the significant slow-down should be observed at region 2: the verb region in subject-gap RCs. However, there is no significant slow-down observed at that point (See Figure 1 and 2). Therefore, the possibility of processing scrambled constructions is not discussed here.

object-gap RCs. This experiment involving context will be pursued as a future study.

If the results of this context-embedded RC experiment still show that object-gap RCs are harder than subject-gap RCs, it will show that depth-of embedding hypothesis makes the right prediction, and suggest that the Chinese data must have some other confounding factor that makes subject-gap RCs harder than object-gap RCs.

Another important finding of this experiment is that when we collapse the object vs. subject-gap RCs and compare the case-matching with case-clashing conditions, the case-matching conditions were processed significantly faster than case-clashing conditions. In other words, case-clashing conditions were significantly more difficult than case-matching conditions. This means that we always have to control case conditions in order to gain reliable results.

7. Conclusion

The results of the present experiment demonstrate the following two things: Firstly, subject-gap RCs are easier to process than object-gap RCs in Japanese, especially in case-clashing conditions. This result is consistent with the depth-of-embedding hypothesis over the resource account. Secondly, the present results show that case-matching relative clauses are easier than case-clashing relative clauses, at least in Japanese. This suggests the necessity of controlling the case conditions for other languages in order to gain reliable results.

However, the experiment was not designed to control for the potential temporary ambiguity that increases the complexity of object-gap RCs. Therefore, a further experiment needs to be conducted in order to control for the potential temporary ambiguity and reanalysis effects by setting up a context which forces the participants to have a relative clause reading for both conditions.

REFERENCES

CHOMSKY, NOAM, and MILLER, GEORGE. 1963. Introduction to the formal analysis of natural languages. In Duncan Luce, Robert Bush and Eugene Galanter (eds.), *Handbook of Mathematical Psychology, Volume 2:* pp. 269-321. Wiley: New York.

FERREIRA, FERNANDA, and CHARLES CLIFTON JR. 1986. The independence of syntactic processing. *Journal of Memory and Language* 25, 348-368.

FUKUI, NAOKI. and YUJI TAKANO. 2000. Nominal Structure: An extension of the symmetry principle. In Peter Svenonius (ed.), *The derivation of VO and OV*, 219-254. Amsterdam/Philadelphia: John Benjamins, 2000.

GIBSON, EDWARD. 1998. Linguistic complexity: locality of syntactic dependencies, *Cognition, 68,* 1-76.

GIBSON, EDWARD. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Yasushi Miyashita, Alec Marantz and Wayne O'Neil (eds.), *Image, language, brain,* pp. 95-126. Cambridge, MA: MIT Press.

HOLMES, V. M, and O'REGAN, J. K. 1981. Eye fixation patterns during the reading of relative clause sentences. *Journal of Verbal Learning and Verbal Behavior, 20*, 417-430.

HSIAO, FRANNY. and GIBSON, EDWARD. 2003. Processing Relative Clauses in Chinese. *Cognition, 90,* 3–27.

HOSHI, KOJI. 1995. Structural and Interpretive Aspects of Head-Internal and Head-External Relative Clauses. Doctoral dissertation: University of Rochester.

JUST, MARCEL., CARPENTER, PATRICIA, and WOOLEY, J. D. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 3,* 228-238.

KING, JONATHAN, and JUST, MARCEL. 1991. Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language, 30,* 580-602.

KEENAN, EDWARD, and COMRIE, BERNARD. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry 8,* 1. 63-99.

KEENAN, EDWARD, and HAWKINS, SARAH. 1987. The psychological validity of the accessibility hierarchy. In Keenan Edward. (ed.). *Universal Grammar : 15 essays,* pp. 60-85. Routledge: London.

KUNO, SUSUMU. 1973. *The structure of the Japanese Language*. MIT Press, Cambridge, Mass.: MIT Press.

LEWIS, RICHARD. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research, 25,* 93-116.

MACWHINNEY, BRAIN. 1977. Starting points. *Language, 53*, 152-168.

MACWHINNEY, BRIAN. 1982. Basic syntactic processes. In Stan Kuczaj (ed.), *Language development: Volume 1, syntax and semantics* pp. 73-136. Erlbaum, Hillsdale, NJ.

MACWHINNEY, BRIAN, and PLÉH, CSABA. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition, 29,* 95-141.

MCCAWLEY, JAMES. 1976. Relativization. In Masayoshi Shibatani (ed.), *Syntax and Semantics 5: Japanese Generative Grammar*, pp. 295-306. California: Academic Press.

MURASUGI, KEIKO. 2000. An Antisymmetry analysis of Japanese relative clauses. In Artemis Alexiadou, Paul Law, Andre Meinunger, and Chris Wilder (eds.) *The syntax of Relative clauses*, pp. 231-263. Amsterdam: John Benjamins,

O'GRADY, WILLIAM. 1987. *Principles of grammar and learning.* Chicago: University of Chicago Press.

O'GRADY, WILLIAM., YAMASHITA, YOSHIE., LEE, MISEON., CHOO, MIHO., and CHO, SPPKEUN. 2000. Computational factors in the acquisition of relative clauses. A talk given at the International Conference on the Development of the Mind, Tokyo.

PERLMUTTER, DAVID M. 1972. Evidence for shadow pronouns in French relativization. In Paul M Peranteau et al. (eds.), *The Chicago Which Hunt*, pp. 73-105. Chicago: Chicago Linguistic Society.

SAUERLAND, ULI. and GIBSON, EDWARD. 1998. How to predict the relative clause attachment preference. Paper presented at the 11th CUNY sentence processing conference, Rutgers University: New Brunswick NJ.

WARREN, TESSA. and GIBSON, EDWARD. 2002. The influence of referential processing on sentence complexity. *Cognition, 85*, 79-112.