

The Effects of Givenness and Heaviness on VP-internal Scrambling and VP-external Scrambling in Japanese*

Satoshi Imamura
University of Oxford

The purpose of this article is to study the interactions between givenness, heaviness, and scrambling in the Japanese language. It is found that, first, the effects of givenness are stronger in VP-external scrambling than in VP-internal scrambling. This can be explained by the left periphery. Since the sentence initial position is relevant to information structure, VP-external scrambling is strongly influenced by discourse context. In contrast, VP-internal scrambling is unrelated to discourse context because there are no discourse-related projections within the VP. Second, it is shown that heaviness has the same effects both on VP-external scrambling and VP-internal scrambling. Third, there is no interaction between givenness and heaviness: both concepts independently affected word order choices.

Keywords: Scrambling; Givenness; Heaviness; Corpus Analysis; Left Periphery

1. Introduction

It is a universal property of natural language to have many options for expressing the same proposition. In Japanese, different word order permutations may be used to convey the same propositional content. Hence, both SOV and OSV orders are possible, as shown in (1). Saito (1985) claims that the OSV ordering is derived from the SOV one by moving the direct object toward the initial position in the sentence. This operation is called ‘scrambling’.

- (1) a. SOV
Taro-ga hon-o yonda.
Taro-NOM book-ACC read
‘Taro read a book.’
b. OSV
Hon-o Taro-ga yonda.
book-ACC Taro-NOM read
‘Taro read a book.’

* I am grateful to every person who supported my research. In particular, Samuel Hughes has helped to increase the articulateness of the texts in this paper.

What is the motivation for scrambling? Since it does not change the grammatical relationships between constituents, in the 1980s, Saito (1985), Kuroda (1988), and Fukui (1986) stated that scrambling is a semantically vacuous movement. However, as Kuno (1978) points out, it is conceivable that scrambling is influenced by information structure, because SOV and OSV orderings differ not in what they say about the world but in how they are conveyed to the hearer (Chafe 1976; Lambrecht 1996; Vallduvi and Engdahl 1996). Given this analysis, it is proposed that information structure is at work here. In order to examine the effects of information structure, the present study employed a quantitative Givōnian approach. Another motivation for scrambling is heaviness (Hawkins 1994; Yamashita 2002; Yamashita & Chang 2001). There is a general consensus that a complex structure involves higher processing costs than a simple structure. Grodner and Gibson (2005) stress that the limitation of a capacity for working memory is the main source of difficulty in processing complex structures. Since scrambled sentences are syntactically more complex than their canonical counterparts, scrambled sentences are in general more difficult to process than canonical sentences (Koizumi and Tamaoka 2010). Yet, when a scrambled constituent is a very long one, scrambled sentences are easier to process than canonical sentences. In a canonical order sentence (2a), for instance, the parser must keep the subject *keiji* ‘detective’ until it encounters the predicate *oikaketa* ‘chased’ in order to form a nexus between the subject and the predicate. During this process, the parser must store *keiji* in his or her working memory. Hence, the processing cost is high. On the other hand, in a scrambled sentence (2b), the parser does not need to remember *keiji* for a long time because it is locally linked with the predicate *oikaketa*. Thus, the processing cost is low. It has been said that heavy constituents are scrambled in order to reduce the processing cost.

- (2) a. *keiji-ga se-ga takaku-te gashiri-shita hannin-o*
 detective-NOM height-NOM tall-and big-boned suspect-ACC
oikaketa
 chased
 ‘The detective chased the suspect who is tall and big-boned.’
- b. *se-ga takaku-te gashiri-shita hannin-o keiji-ga*
 height-NOM tall-and big-boned suspect-ACC detective-NOM
oikaketa
 chased
 ‘The detective chased the suspect who is tall and big-boned.’

(adapted from Yamashita and Chang 2001: B47)

The aim of this study is to investigate the effects of givenness and heaviness on VP-internal and VP-external scrambling, using a corpus analysis. This paper is organized in the following way. Section 2 reviews previous studies about scrambling and the Givōnian approach. Section 3 presents my corpus analysis of scrambling in terms of givenness and heaviness. Section 4 explains the results of my analysis on the basis of

cartography. I show that the impact of givenness is stronger in VP-external scrambling than in VP-internal scrambling. On the other hand, there is no difference between the effects of heaviness on VP-internal scrambling and VP-external scrambling; I demonstrate, moreover, that givenness and heaviness affect scrambling independently. It is thus reasonable to suppose that both concepts are needed in order to analyze the functional aspects of scrambling. Section 5 is devoted to the conclusion and further studies.

2. Previous Studies

2.1. Scrambling

2.1.1. VP-external Scrambling

It is well known that Japanese is a relatively free word order language. Syntactically, it has been assumed in general that the direct object in OSV ordering is moved from the VP-internal position to the sentence-initial position (Miyagawa 2001, 2003, 2010; Saito 1985, 2009; Saito and Hoji 1983). This type of word order change is called ‘VP-external scrambling’ because the direct object is moved to a VP-external position. It should be noted that scrambling does not affect the grammatical relations between constituents. Both SOV and OSV orders can convey the same propositional meaning. Thus, scrambling appears to be a semantically vacuous movement. Saito (2012) observed, however, that scrambling has implications for the interpretation of topic marker *wa*. In SOV sentences like (3a), as Kuno (1973) notes, only sentence-initial *wa* can be interpreted as a thematic topic. However, in OSV sentences like (3b), when the *wa*-marked direct object is scrambled to the sentence initial position, both the subject and the direct object can have a thematic interpretation. Consequently, Saito maintains that scrambling can count as topicalization.

- (3) a. Taro-wa (kyonen) sono hon-wa katta.
 Taro-TOP last year that book-TOP bought
 A. ‘Speaking of Taro, he bought that book (last year), but I don’t know about other books’ (Taro=Thematic, that book=contrastive)
 B. ‘Taro bought that book (last year), but I don’t know about other people and other books’ (Taro=contrastive, that book=contrastive)
- b. sono hon-wa Taro-wa (kyonen) kat-ta.
 that book-TOP Taro-TOP last year bought
 A. ‘Speaking of that book, Taro bought it (last year), but I don’t know about other people’ (that book=thematic, Taro=contrastive)
 B. ‘Speaking of Taro, he bought that book (last year), but I don’t know about other books’ (Taro=thematic, that book=contrastive)
 C. ‘Speaking of that book and speaking of Taro, he bought it (last year)’ (that book=thematic, Taro=thematic)
 D. ‘Taro bought that book (last year), but I don’t know about other

books and other people’ (that book=contrastive, Taro=contrastive)
(Saito 2012: 158–9)

Moreover, Kuno (1978: 54) observed that Japanese word orders are determined by given-new ordering, which means that given information is mentioned early and new information later. In other words, scrambling is an operation to place given information in an earlier position compared to new information. In addition, Imamura, Sato, and Koizumi (2014) showed using a sentence comprehension experiment that OSV sentences were processed faster in given-new ordering than in new-given ordering. All of these studies support the idea that scrambling is discourse-driven.

Some scholars claim that scrambling is motivated by heaviness (Hawkins 1994; Yamashita 2002; Yamashita & Chang 2001). Hawkins (1994) states that heavy constituents are scrambled to facilitate the processing. Specifically, Yamashita and Chang (2001) demonstrated that native Japanese speakers were prone to move long constituents to earlier positions than short constituents in sentence production.

Summing up, two factors as explanations for the use of VP external-scrambling: information structure and heaviness. Since information structure is a broad concept, this study focusses on the effects of givenness, using a Givónian approach.

2.1.2. VP-internal scrambling

According to Hoji (1985), S-DO-IO-V sentences like (4b) are derived from S-IO-DO-V sentences like (4a). This word order change is called ‘VP-internal scrambling’ because the exchange between the indirect object and the direct object arises within VP. However, Miyagawa (1997) argues that both S-IO-DO-V and S-DO-IO-V orders are base-generated, while Matsuoka (2003) claims that which word order is canonical depends on the verb type. *Watasu*-type ‘to pass’ verbs take S-DO-IO-V as basic word order whereas *miseru*-type ‘to show’ verbs have S-IO-DO-V as canonical word order. The validity of these hypotheses was tested by Koizumi and Tamaoka (2004), who reported on the basis of a sentence comprehension experiment that both show-type and pass-type sentences with the S-IO-DO-V order were processed faster than with the S-DO-IO-V order. This result indicates that S-IO-DO-V ordering is canonical and that S-DO-IO-V ordering is scrambled, because scrambled sentences take more processing time than their canonical counterparts, generally due to the difference in syntactic complexity (Tamaoka et al. 2005). Koizumi and Tamaoka’s findings thus support the traditional analysis proposed by Hoji (1985) rather than the more recent proposals of Miyagawa (1997) and Matsuoka (2003). This paper accordingly presupposes that S-IO-DO-V ordering is canonical and S-DO-IO-V ordering is scrambled.

- (4) a. S-IO-DO-V
 Taro-ga Hanako-ni present-o ageta.
 Taro-NOM Hanako-DAT present-ACC gave
 ‘Taro gave Hanako a present.’

b. S-DO-IO-V

Taro-ga present-o Hanako-ni ageta.
 Taro-NOM present-ACC Hanako-DAT gave
 ‘Taro gave Hanako a present.’

Assessing the functional significance of scrambling, Yamashita and Chang (2001) observed that heaviness affected the choice between S-IO-DO-V and S-DO-IO-V orderings in sentence production. Participants produced S-DO-IO-V ordering more often in long-DO conditions than in short-DO conditions. Regarding information structure, Ferreira and Yoshita (2003) investigated the effects of given-new ordering on VP-internal scrambling. Japanese speakers were asked to answer questions like (5) by using a canonical sentence (S-IO-DO-V ordering) like (6a) or a scrambled sentence (S-DO-IO-V ordering) like (6b). Question sentences were used in order to establish a dative or accusative argument in an answer sentence as given information. Thus, *otetsudaisan* ‘housekeeper’ is given information in (6) because she is referred to once in (5), resulting in given-new order in (6a) and new-given order in (6b). The results showed no interaction between givenness and VP-internal scrambling.

(5) Question:

okusan-ga otetsudaisan-ni kansyashiteita. sorekara dōshita.
 housewife-NOM housekeeper-DAT thanked then what.happened
 ‘The housewife thanked the housekeeper. What happened next?’

(6) Answers:

- a. okusan-ga otetsudaisan-ni present-o okutta.
 housewife-NOM housekeeper-DAT present-ACC sent
 ‘The housewife sent the housekeeper a present.’
- b. okusan-ga present-o otetsudaisa-ni okutta.
 housewife-NOM present-ACC housekeeper-DAT sent
 ‘The housewife sent the present to the housekeeper.’

Taken together, previous studies have reported that VP-internal scrambling is influenced by heaviness but not by givenness, as opposed to VP-external scrambling, where both heaviness and givenness seem at work. What causes this difference? Or, is it a coincidence? One purpose of this study is to disentangle this issue.

2.2. Givónian Givenness

Since ‘givenness’ is a vague concept, this study employs Givónian givenness, which Givón (1983, 1994) proposed as a quantitative approach for calculating the topicality of referents. In this paper, Givónian givenness is determined by Referential Distance (RD). The metric of RD measures the gap between a referent in the current clause and its antecedent, using clause boundaries as units. If there is no antecedent in the previous clauses, the RD is assigned a value of 21 because without some limitation

it would be infinite.¹ Hence the RD is expressed by some number of clauses from 1 to 21. What should be emphasized here is that RD is not a binary concept, and can express multiple degrees of givenness. That is, it is possible to state that some referent is older than other referents. This is illustrated by figure 1.

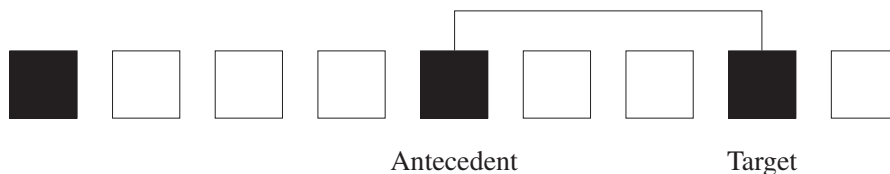


Figure 1. Image illustrating Referential Distance

Suppose that each square represents a clause and that black squares include the same referent. In order to measure the RD of the referent in the black square furthest to the right (target clause), one must return to the middle black square. Since there are three clause boundaries between the referent in the middle clause and the one in the right clause, the RD for the referent in the right black square is 3. Although the same referent is once mentioned in the black square furthest to the left, this has nothing to do with the RD of the referent in the black square furthest to the right. This is because RD is the value of the distance between the target referent and its nearest antecedent. RD is a well-recognized measurement that is easily implementable, and its employment renders the results of my analysis reproducible.

3. Corpus Analysis of Scrambling

3.1. Basic Purposes and Procedures

The purpose of this paper is to examine the interactions between scrambling types, heaviness, and givenness in Japanese. Previous studies have demonstrated that heaviness and givenness are important factors for the usage of scrambling. Yet no study has examined the interactions between them. It is conceivable that scrambling is selected only when a scrambled constituent is both heavy and given; alternatively, there may be a complementary distribution between heaviness and givenness. When there are two competing factors, the strength of one factor becomes greater when the other factor is weak (Arnold et al. 2000; Trueswell and Tanenhaus 1994). If givenness competes with heaviness, scrambling will be chosen when it meets only one factor. Moreover, previous studies do not distinguish VP-external scrambling from VP-internal scrambling.

¹ The limitation of RD is rather arbitrary. For example, Givón (1994) proposed that it should be 3 and Cooreman (1992) suggested that it should be 15 because there was no example with an RD higher than 15. However, I observed sentences with RDs higher than 16, so I followed the criteria of Givón (1983) and Shimojō (2005).

Recall that several studies observed a correlation between givenness and OSV (Imamura et al. 2014; Kuno 1978), and no interrelation between givenness and S-DO-IO-V ordering (Ferreira and Yoshita 2003; Yamashita 2002). What makes this difference? There are two possibilities here. One possibility is that there are interrelations between givenness and scrambling types. If this were so, we should observe the interaction between givenness and scrambling in VP-external scrambling (OSV ordering) but not in VP-internal scrambling (S-DO-IO-V ordering). The other possibility is that the difference in modality is at work here. The reason that the two studies produced apparently contrasting results may be that Imamura et al. tested sentence comprehension, whereas Ferreira and Yoshita tested sentence production. In the present study, VP-internal scrambling is compared with VP-external scrambling in terms of givenness and heaviness. To this end, the RDs and the lengths of direct objects in OSV and S-DO-IO-V sentences are measured.

3.2. Method

3.2.1. Corpus Data

The Balanced Corpus of Contemporary Written Japanese (BCCWJ) was employed in order to assemble relevant data. BCCWJ is designed to be representative of contemporary written Japanese and thus includes 100 million words from well-balanced written materials covering books, magazines, newspapers, library books, bulletin boards, blogs, best-selling books, school textbooks, minutes of the National Diet, publicity of newsletters of local governments, laws, and poetry (see Maekawa et al. 2008 in detail).

3.2.2. Materials

OSV and S-DO-IO-V sentences were assembled from BCCWJ by using *Chunagon*, which is a web interface program. With regard to OSV ordering, the string [*o*(ACC)-noun-*ga*(NOM)] was used to extract OSV examples. The reason for using only strings with subject NPs of minimal length is that the left boundaries of NPs are not marked in the corpus. The limitation of my design is that it cannot pick out complex subjects completely. Complex subjects modified by a relative clause in OSV sentences (e.g. [[noun-*ga*-verb]-noun-*ga*]) were eliminated by hand in order to control the data. Thus, the scope of the OSV-ordering string includes only a simple (non-branching) noun subject.

With regard to S-DO-IO-V, the string [*o*(ACC)-noun-*ni*-(DAT)-verb] was utilized. Since there were too many unrelated examples, they were narrowed down on the basis of the verb meaning. Specifically, the GIVE-schema proposed by Shibatani (1994) was employed and only examples that met (7) were analyzed.

(7) The GIVE-schema for Japanese

Structure: [NP₁ NP₂ NP₃ GIVE]

NP₁ = coded as a subject NP₂ = coded as an indirect object

NP_3 = coded as a direct object
 Semantics: NP_1 CAUSES NP_2 TO HAVE NP_3 ; i.e.
 NP_1 = human agent, NP_2 = human goal, NP_3 = object theme
 NP_2 exercises possessive control over NP_3 ,
 NP_1 creates the possessive situation on behalf of NP_2
 (Shibatani 1994: 45)

3.2.3. Calculation of Heaviness

In order to measure the lengths of direct objects, the *bunsetsu* of direct objects were counted. A *bunsetsu* is a basic linguistic unit in Japanese linguistics, consisting of one or more content words followed by zero or more functional words. Generally speaking, *bunsetsu* correspond to a phrase. The reason why *bunsetsu* is chosen to be counted is that the length of the subjects in OSV orderings and the indirect objects in S-DO-IO-V orderings is controlled in terms of *bunsetsu*. The *bunsetsu*-count of the subjects in OSV ordering and the indirect objects in S-DO-IO-V ordering is always 1 in my data because they are a single noun plus a case particle *GA* or *NI*. In (8), for example, *Dan-ni* ‘to Dan’ forms a single *bunsetsu* because it is a content word *Dan* followed by a functional word *NI*.

- (8) *watashi-wa genkō-o Dan-ni watashita.*
 I-TOP draft-ACC Dan-ACC passed
 ‘I passed the draft to Dan.’

(BCCWJ)

3.2.4. Calculation of RD

The criteria of RD calculation are partly drawn from Shimojō (2005), specifically the criteria of back-channel feedback and adjacent predicates. Other criteria are my own. The following sections will explain the details of these criteria.

3.2.4.1. Bridging Relationships

Some kinds of inferable information are categorized as discourse-old information. In particular, bridging relations are taken into consideration. Bridging is an inference from a referent that is explicitly mentioned in the preceding discourse. In (9), the hearer must suppose that *ringo* ‘apple’ is a part of *kudamono* ‘fruit’. This is a bridging relation. Though *ringo* ‘apple’ is not directly referred to in (9a), its RD is 1 because *kudamono* ‘fruit’ can be regarded as the antecedent.

- (9) a. *Taro-wa kinō kudamono-o katta*
 Taro-TOP yesterday fruit-ACC bought
 ‘Taro bought fruit yesterday.’
 b. *shikashi, ringo-wa kusatteita.*
 but apple-TOP were.rotten

‘But the apples were rotten’

3.2.4.2. Complex Clause

Complex clauses are divided into separate clauses based on predicates. Subordinate clauses are therefore regarded as independent clauses. For example, the complex clause (10) is divided into two clauses because it contains the two predicates; *nagameru* ‘look at’ and *watasu* ‘pass’. In order to illustrate the process of calculation of RD, let us measure the RD of *sore* ‘it’. The first step is to check the antecedent of *sore*. Here, it is *syashin* ‘photo’ because it refers to the same object that *sore* does. The second step is to calculate the clause boundaries between the target referent *sore* and its antecedent *shashin*. Since there is only one clause boundary between them, the RD of *sore* is 1.

- (10) [₁Martin-wa syashin-o nagame,] [₂(S) sore-o Mortimer-ni watashita]
 Martin-TOP photo-ACC look.at (he) it-ACC Mortimer-DAT passed
 ‘Martin looked at the photo and passed it to Mortimer.’

(BCCWJ)

3.2.4.3. Adjacent Predicates

The V_1 -*te*- V_2 form is normally placed in the same clause, but when V_1 and V_2 have different subjects, each verb is regarded as belonging to an independent clause (Shimojō 2005: 57–8). For example, in (11a), the linked verb *kat-te-kite* ‘buy-TE-come-and’ share the zero subject ‘I’. The V_1 -*te*- V_2 form thus belongs to the same clause. In contrast, in (11b), V_1 and V_2 have different subjects. In other words, V_1 *motte* ‘have’ forms a nexus with *tomodachi* ‘friend’ and V_2 *karite* ‘borrow’ forms a nexus with the zero subject ‘I’. In this case, both V_1 and V_2 constitute an independent clause because they do not share the same subject.

- (11) a. kyanberu-no suupukan kat-te-ki-te
 Campbell-LK soup.can buy-TE-come-and
 ‘(I) bought a Campbell soup can (and came).’
 b. dorai-no-yatsu-o tomodachi-ga motte-te (S) (O) karite
 dry-LK-one-ACC friends-NOM have-TE (I) (it) borrow-and
 ‘A friend had dry (basil) and (I) borrowed (it).’

(Shimojō 2005: 57–8)

3.2.4.4. Back-channel feedback

Generally speaking, back-channel feedback such as *soo* ‘indeed’ and *un* ‘yeah’ are propositionally empty and are given by the hearer while speaker is holding the conversational turn (Shimojō 2005: 58). They are considered to be dependent on another clause and do not form an independent clause.

3.2.4.5. Copula

Copula expressions such as *da* and *dearu* are regarded as predicates and hence they head independent clauses. Thus, in (12), the copula *dearu* identifies an independent clause because clause boundaries are based on predicates in this study.

- (12) waga-kuni-wa kokudo-no sanbun-no-ni-o shinrin-ga
 our.country-TOP land-GEN thirds-of-two-ACC forest-NOM
 simeru-hodo shinrinshigen-ni megumareta-kuni-deari...
 account.for-about forest.resources-DAT be.blessed-country-COP
 ‘Our country is a country that is blessed with forest resources such that two-thirds of the land in our nation is covered with forests...’

(BCCWJ)

3.2.4.6. Proposition

The method for determining RD has been developed for calculating the discourse status of a referent (Givōn 1983, 1994). Propositions are not included in this method because they are not referents themselves, but relationships between referents. Instead of directly calculating the RD of a proposition, the RDs of the related referents were counted. In my approach, the RD of the proposition is the smallest value of the referents pertinent to that proposition. For instance, in (13b), the scrambled object is the proposition *Hänsel-ga naka-ni hai-routosuru* ‘that Hänsel is trying to come in it’, which includes the referents *Hänsel* and *candy house*. Therefore, this proposition has the two related referents, *Hänsel* and *candy house*. In this study, the RDs of both *Hänsel* and *candy house* are calculated. Note that the head of the scrambled object is nominalizer *no*, but that it is anchored by *Hänsel* and *candy house*. Thus the RD of the scrambled object is replaced by the anchoring expression’s RD, which is 1.

- (13) a. okashinoie-ga aru-node hutari-wa hidoku bikkurisuru
 candy.house-NOM be-because two.person-TOP very surprised
 ‘Since there is a candy house, the two are very surprised.’
 b. Hänsel-ga naka-ni hai-routosuru-no-o Gretel-ga
 Hänsel-NOM inside-LOC come-try.to.do-NMZ-ACC Gretel-NOM
 togameru
 blame.for
 ‘Gretel berates Hänsel for trying to enter.’

(BCCWJ)

3.3. Results

This study analyzed 3273 examples of OSV clauses and 2229 examples of S-DO-IO-V clauses from BCCWJ, and measured the RD value and the length for the scrambled direct object in each example. Table 1 summarizes the distributions of scrambled objects in OSV and S-DO-IO-V orderings in terms of RD. It reveals that 2681 scram-

bled objects (81.91%) in OSV orderings and 1463 scrambled objects in S-DO-IO-V orderings (65.63%) have an antecedent.

Table 1. Tokens distribution of O in OSV and S-DO-IO-V in terms of RD

RD	Number of O in OSV (%)	Number of O in S-DO-IO-V (%)
1	1724 (52.67%)	732 (32.84%)
2	368 (11.24%)	222 (9.96%)
3	194 (5.93%)	131 (5.88%)
4	102 (3.12%)	74 (3.32%)
5	61 (1.86%)	60 (2.69%)
6	49 (1.50%)	39 (1.75%)
7	34 (1.04%)	45 (2.02%)
8	37 (1.13%)	24 (1.08%)
9	19 (0.58%)	23 (1.03%)
10	12 (0.37%)	20 (0.90%)
11	14 (0.43%)	17 (0.76%)
12	15 (0.46%)	7 (0.31%)
13	8 (0.24%)	9 (0.40%)
14	5 (0.15%)	7 (0.31%)
15	5 (0.15%)	11 (0.49%)
16	4 (0.12%)	7 (0.31%)
17	5 (0.15%)	9 (0.40%)
18	10 (0.31%)	3 (0.13%)
19	10 (0.31%)	7 (0.31%)
20+ ²	5 (0.15%)	16 (0.72%)

² 20+ includes the examples whose RD are more than 20. However, their values are substituted for 20.

NPM ³	592 (18.09%)	766 (34.37%)
Total	3273 (100%)	2229 (100%)

Table 2 is the summary of the distribution of scrambled objects in terms of *bunsetsu*. Note that the subjects in OSV orderings and the indirect objects in S-DO-IO-V orderings are always 1 *bunsetsu* due to my design. Hence, more than one *bunsetsu* in Table 2 means that the scrambled direct object is longer than the subject in OSV or the indirect object in S-DO-IO-V orderings in terms of *bunsetsu*. Thus, heaviness correlates with scrambled objects in about 75.95% of examples in OSV and in about 69.85% of examples in S-DO-IO-V, where the direct object is longer than one *bunsetsu*.

Table 2. The lengths of scrambled direct objects in terms of *Bunsetsu*

Bunsetsu	Number of O in OSV (%)	Number of O in S-DO-IO-V (%)
1	787 (24.05%)	672 (30.15%)
2	1028 (31.41%)	818 (36.70%)
3	564 (17.23%)	318 (14.27%)
4	379 (11.58%)	159 (7.13%)
5	230 (7.03%)	108 (4.85%)
6	109 (3.33%)	55 (2.47%)
7	61 (1.86%)	38 (1.70%)
8	44 (1.34%)	21 (0.94%)
9	16 (0.49%)	14 (0.63%)
10+ ⁴	55 (1.68%)	26 (1.17%)
Total	3273	2229

Next, a Pearson correlation test was conducted between RD and *bunsetsu* in order to see if there is a correlation between givenness and heaviness. This revealed that there is no correlation between givenness and heaviness in either OSV orderings ($r = -.09$, $p < .01$) or S-DO-IO-V orderings ($r = -.014$, $p < .01$). Therefore, the RD of the scrambled object is independent of its length.

³ NPM is an abbreviation of No Previous Mention. Following Shimojō (2005: 74), NPM is counted as 21.

⁴ 10+ includes 10 and more than 10 *bunsetsu*.

4. Discussion

The aim of this analysis is to explore the relationships between givenness, heaviness, and scrambling types. First, givenness is likely to have a larger effect on VP-external scrambling than on VP-internal scrambling. In fact, 81.91% of scrambled objects are given information in OSV while only 65.63% of scrambled objects in S-DO-IO-V sentences are given information. Why is there such a difference between VP-external scrambling and VP-internal scrambling? One explanation is based on the mapping between information structure and syntactic structure. According to Rizzi (1997), the sentence-initial position is related to discourse function. He presupposes that the left periphery (CP-domain) is used in order to express the interfaces between syntactic structure and information structure.

(14) Split-CP

[ForP For° [Top* Top° [FocP Foc° [TopP* Top° [FinP Fin° [TP...

As shown in (14), the left periphery consists of many kinds of discourse-related projections. The CP shell is divided into a Force layer, a Higher Top layer, a Focus layer, a Lower Top layer, and a Finiteness layer. Considering scrambling in terms of the left periphery, it may be seen that VP-external scrambling is motivated by givenness because the sentence-initial position can have a relationship with discourse-related projections in the CP-domain. In contrast, VP-internal scrambling is not strongly influenced by givenness because there are no projections for information structure within the VP-domain. On top of that, Diesing (1992) stated that discourse-anaphoric arguments must move from the VP-domain to the old information domain. This analysis suggests that the VP-domain is a preferable place for new information. To summarize: givenness seems to have strong effects on VP-external scrambling but to have weak effects on VP-internal scrambling. This difference may be explained by the existence of discourse-related projections at the sentence-initial position.

It is possible however that both VP-external scrambling and VP-internal scrambling are motivated by givenness. Matsuoka (2003) contends that S-DO-IO-V ordering is base-generated in sentences with a *Watasu*-type ‘to pass’ verb. Although it is unclear whether Matsuoka’s classification is on the right track, if S-DO-IO-V sentences are base-generated, they do not have to be motivated because there is no movement operation, leading to the conclusion that givenness makes no impression on some S-DO-IO-V sentences. This may be the reason why givenness generally has a weaker impact on S-DO-IO-V sentences than OSV ones. If so, it is technically feasible to maintain that other S-DO-IO-V sentences are motivated by givenness. In addition, Miyagawa and Tsujioka (2004) found that some idiomatic S-DO-IO-V sentences are base-generated, as shown in (15). The reversed word order IO-DO leads to disappearance of the idiomatic meaning *te-ni-ireru* ‘acquire’. However, it is not obvious whether there are verb classes that prefer S-DO-IO-V as a canonical word order. Further studies are needed in order

to answer the question as to whether S-DO-IO-V sentences with a certain verb type are base-generated or not.

- (15) Taroo-wa kuruma-o te-ni ireta.
 Taro-TOP car-ACC hand-in put.in
 ‘Taro acquired a car.’

(Miyagawa and Tsujioka 2004: 21)

Second, heaviness has essentially the same effect on VP-internal scrambling as it does on VP-external scrambling. In fact, heaviness correlates with 69.85% of VP-internal scrambled objects and 75.95% of VP-external scrambled objects. The difference is only 6.1%. The effects of heaviness do not vary greatly between scrambling types. When scrambling is chosen because of heaviness, the motivation for this grammatical operation is to reduce processing cost. It does not matter whether scrambled constituents are given information or not. There is no reason for heaviness to distinguish VP-internal scrambling from VP-external scrambling. This may be the reason why heaviness has the same effect both on VP-internal scrambling and VP-external scrambling. Consequently, scrambling-type distinction is not important for heaviness.

Third, it has been shown that there is no correlation between heaviness and givenness in either VP-internal or VP-external scrambling. This fact means that heaviness and givenness independently influence word order choices. Thus, it is necessary to take both concepts into account in order to explain the function of scrambling. This fact implies that there may not be a simple motivation for scrambling. Scrambling may be an amalgam of operations motivated by several factors. Other factors may therefore be needed in order to explain the usage of scrambling with precision. Specifically, an alternative approach is indispensable to deal with qualitative data. As a quantitative approach, the Givōnian approach concentrates only on linguistic data and excludes other factors such as the speaker’s point of view and the relationship between the speaker and the hearer. In order to complement the Givōnian approach, future studies will need to take other factors into consideration.

5. Conclusion

The aim of this paper has been to probe the relationships between givenness, heaviness, and scrambling more closely. First, it has been demonstrated that givenness has a greater impact on VP-external scrambling than on VP-internal scrambling. This can be explained by the left periphery. Second, there is no difference in heaviness between VP-external scrambling and VP-internal scrambling. This can be accounted for by presuming that there is no need to tell VP-external scrambling from VP-internal scrambling in terms of processing cost. Third, there was no direct interaction between givenness and heaviness. This signifies that both givenness and heaviness have independent effects on scrambling. Though this study has focussed on givenness and heaviness, there

may be other factors that have a great influence on word order changes. However, it has been demonstrated that information structure affects scrambling. I therefore propose that a fruitful avenue of future research is to survey the functions of scrambling at the discourse level.

References

- Arnold, J. E., Losongco, A., Wasow, T. and Ginstrom, R. 2000. "Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering." *Language* 76:1, 28-55.
- Chafe, W. 1976. "Givenness, Contrastiveness, Definiteness, Subjects, Topics and Points of View." In N. Charles (ed.), *Subjects and Topic*, 25-55. New York: Academic Press.
- Cooreman, A. 1992. "The pragmatics of word order variation in Chamorro narrative text." In P. Doris, P. (ed.), *Pragmatics of word order flexibility*, 243-64. Oregon: John Benjamins Publishing.
- Diesing, M. 1992. *Indefinites*. Cambridge, Mass: MIT Press.
- Ferreira, V. S. and Yoshita, H. 2003. "Given-new ordering effects on the production of scrambled sentences in Japanese." *Journal of psycholinguistic research* 32:6, 669-692.
- Fukui, N. 1986. *A theory of category projection and its implications*. Doctoral dissertation, MIT.
- Givón, T. 1983. "Topic continuity in discourse: An introduction." In T. Givón (ed.), *Topic Continuity in Discourse*, 5-41. Amsterdam: John Benjamins Publishing.
- Givón, T. 1994. "The pragmatics of de-transitive voice: Functional and typological aspects of inversion." In T. Givón (ed.), *Voice and inversion*, 3-47. Amsterdam: John Benjamins Publishing.
- Grodner, D. and Gibson, E. 2005. "Consequences of the serial nature of linguistic input for sentential complexity." *Cognitive Science*: 29, 261-290.
- Hawkins, J. A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge.
- Hoji, H. 1985. *Logical Form Constraints and Configurational Structures in Japanese*. Doctoral dissertation, University of Washington, Seattle.
- Imamura, S., Sato, Y. and Koizumi, M. 2014. "Influence of Information Structure on Word Order Change and Topic Marker *WA* in Japanese." In W. Aroonmanakun, P. Boonkwan, and T. Supnithi (eds.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation*, 432-441. Bangkok: Chulalongkon University.
- Koizumi, M. and Tamaoka, K. 2004. "Cognitive processing of Japanese sentences with ditransitive verbs." *Gengo Kenkyū* 125: 173-190.
- Koizumi, M. and Tamaoka, K. 2010. "Psycholinguistic evidence for the VP-internal subject position in Japanese." *Linguistic Inquiry* 41: 663-680.
- Kuno, S. 1973. *The structure of the Japanese language*. Cambridge, Mass: MIT Press.
- Kuno, S. 1978. *Danwa-no Bumpō* (Grammar of Discourse). Tokyo: Taishūkan.
- Kuroda, S.-Y. 1988. "Whether we agree or not." In W. Poser (ed.), *Papers from the Second International Workshop on Japanese Syntax*, 103-143. Stanford, CA: CSLI Publications.
- Lambrecht, K. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Publishing.
- Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koizo, H., Yamaguchi, M., Tanaka, M. and Den, Y. (2008). Balanced corpus of contemporary written Japa-

- nese. *Language Resources and Evaluation* 48: 345–71.
- Matsuoka, M. 2003. “Two Types of Ditransitive Constructions in Japanese.” *Journal of East Asian Linguistics* 12:2, 171–203.
- Miyagawa, S. 1997. “Against optional scrambling.” *Linguistic inquiry* 28:1, 1–25.
- Miyagawa, S. 2001. “The EPP, scrambling and *wh-in-situ*.” In M. Kenstowicz (ed.), *Ken Hale: A life in language*, 293–338. Cambridge, Mass: MIT Press.
- Miyagawa, S. 2003. “A-movement scrambling and options without optionality.” In S. Karimi (ed.), *Word order and scrambling*, 177–200. Oxford: Blackwell Publishers.
- Miyagawa, S. 2010. *Why Agree? Why Move? Unifying Agreement-based and Discourse-configurational Languages*. Cambridge: MIT Press.
- Miyagawa, S. and Tsujioka, T. 2004. “Argument structure and ditransitive verbs in Japanese.” *Journal of East Asian Linguistics* 13:1, 1–38.
- Rizzi, L. 1997. “The fine structure of the left periphery.” In L. Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer Academic Publisher.
- Saito, M. 1985. *Some Asymmetries in Japanese and their Theoretical Implications*. Doctoral dissertation, MIT.
- Saito, M. 2012. “Semantic and Discourse Effects of Scrambling.” In B. Frellesvig, J. Kiaer and J. Wrona (eds.), *Studies in Japanese and Korean Linguistics*, 146–72. München: Lincom GmbH.
- Saito, M. and Hoji, H. 1983. “Weak cross over and move α in Japanese.” *Natural Language and Linguistic Theory* 1:2, 245–259.
- Shibatani, M. 1994. “Benefactive constructions: A Japanese-Korean comparative perspective.” *Japanese/Korean Linguistics* 4, 39–74.
- Shimojō, M. 2005. *Argument encoding in Japanese conversation*. Hampshire and New York: Palgrave Macmillan.
- Tamaoka, K., Sakai, H., Kawahara, J., Miyaoka, Y., Lim, H. and Koizumi, M. (2005). Priority information used for the processing of Japanese sentences: Thematic roles, case particles or grammatical functions? *Journal of Psycholinguistic Research*, 34: 281–332.
- Trueswell, J. and Tanenhaus, M. 1994. “Toward a lexical framework of constraint-based syntactic ambiguity resolution.” In C. Clifton, L. Frazier and K. Rayner (eds.), *Perspectives on sentence processing*, 155–179. Psychology Press.
- Vallduví, E. and Engdahl, E. 1996. “The linguistic realization of information packaging.” *Linguistics*, 34:3, 459–520.
- Yamashita, H. and Chang, F. 2001. “Long before short preference in the production of a head-final language.” *Cognition* 81:2, B45–B55.
- Yamashita, H. 2002. “Scrambled sentences in Japanese: Linguistic properties and motivations for production.” *TEXT-THE HAGUE THEN AMSTERDAM THEN BERLIN* 22:4, 597–634.