Markedness and Argument Structure

Markedness and Argument Structure

Satoshi IMAMURA

University of Oxford

1. Introduction

It is a well-known fact that the Japanese is a relatively free word order language. Since grammatical relations are expressed by case markers and major constituents, except for verbs, words can be freely ordered. Thus, both SOV and OSV can convey the same propositional meaning as shown in (1).

(1) a. SOV

Taro-ga	Jiro-o	oikaketa
Taro-NOM	Jiro-ACC	chased

'Taro chased Jiro.'

b. OSV

Jiro-o	Taro-ga	oikaketa
Jiro-ACC	Taro-NOM	chased

Theoretically, the direct object in OSV is considered to be moved from the VP-internal position to the initial position of the sentence (Miyagawa 2001, 2003, 2010; Saito 1985, 2009; Saito and Hoji 1983). This process is called 'scrambling'. In terms of frequency, Kuno (1973:4) estimated that the ratio of SOV to OSV orders appearing in newspaper articles is 17:1. This poses the question: why is the frequency of OSV lower than that of SOV? I will seek to answer this question on the basis of *Preferred Argument Structure* proposed by Du Bois (1987, 2003).

The paper is organized as it follows: Sections 2 provides an overview to previous studies on word order and information structure. Section 3 reanalyzes the results of Imamura and Koizumi (2011). In section 4, I am going to explain the distributions of marked and marked transitive sentences on the basis of *Preferred Argument Structure*. Section 5 will conclude and discuss the prospects for future studies in this area.

2. Previous Studies

It has long been claimed that in formation of sentences given information tend to appear before new information. This tendency is called 'given-new ordering'. The reason for this generalization is that it is easy to process sentences that begins with information that is already present in hearer's mind. Numerous studies have observed that given-new ordering has a great influence on the choice of word orders in many languages, as shown in Birner, Kaplan, and Ward (2007) for English, Birner and Mabootian (1996) for Farsi, Kaiser and Trueswell (2004) for Finnish, Kuno (1978) for Japanese, and Rambow (1993) for German. In particular, Kuno (1978: 54) observed that native-Japanese speakers tend to select OSV when the scrambled object is older than the subject. This observation has been supported by various studies. Imamura (2014) has demonstrated using a corpus analysis that scrambled objects are apt for providing information. Imamura, Sato, Koizumi (2014) has revealed using a sentence comprehension experiment that scrambled sentences were processed faster in the cases where they follow given-new ordering than the cases where they violate given-new ordering. Therefore, the finding shows that word orders are strongly influenced by information structure.

In terms of frequency, largely based on the study conducted by National Language Research Institute, Kuno (1973:4) reported that the ratio of frequencies of occurrences between the SOV and OSV is 17:1 Moreover, Miyajima (1964) collected 1365 clauses from magazines and observed that the ratios SOV and OSV word orders is approximately 22.5:1. In other words, 1307 examples were SOV and 58 examples were OSV. If the choice between SOV and OSV is not random, there must be legitimate reasons for this statistical bias. Why do the native-Japanese speakers prefer SOV to OSV? In my opinion, this fact can be explained by the interaction between grammar and discourse, especially by Du Bois's *Preferred Argument Structure*.

What is *Preferred Argument Structure*? Du Bois (1987, 2003) proposes that there are preferable relationships between argument structure and discourse. He hypothesizes that certain configurations of arguments are preferred over other grammatically possible alternations. Under these constraints, the speaker/writer should avoid more than one new core argument and new agent. Applying these constraints to the transitive construction, transitive subject position is undesirable for new information, but there is no such a constraint for object and intransitive subject positions. In sum, transitive subject is considered to be a position to express given information.

In order to examine the interaction between the syntactic structure and information structure in Japanese, I am going to reanalyze the results of Imamura and Koizumi (2011), which assembled SOV and OSV sentences from novels. Since Imamura and Koizumi (2011)

take topic marker *WA* into consideration, I am going to investigate the interaction between word orders (SOV vs. OSV) and particles (topic marker *WA* vs. case markers *GA* and *O*)

3. Reanalysis of Imamura and Koizumi (2011)

In this section, I am going to reanalyze the results of Imamura and Koizumi (2011) in terms of frequency.

3.1. Data and Materials

$S_{NOM}O_{ACC}V$, $S_{TOP}O_{ACC}V$, $O_{ACC}S_{NOM}V$, and $O_{TOP}S_{NOM}V$ sentences were collected from *Aozora Bunko*, which is a database of Japanese novels. They were all simple transitive sentences as shown in (2). What I should note here is that they belong to main clauses.

(2) Sentence Types

a. $S_{NOM}O_{ACC}V$

Taro-ga Hanako-o mitsuke-ta.
Taro-NOM Hanako-ACC find-PAST
'Taro found Hanako.'

b. $S_{TOP}O_{ACC}V$

Taro-wa Hanako-o mitsuke-ta.
Taro-TOP Hanako-ACC find-PAST
'As for Taro, he found Hanako.'

c. $O_{ACC}S_{NOM}V$

Hanako-o Taro-ga mitsuke-ta.
Hanako-ACC Taro-NOM find-PAST
'Taro found Hanako.'

d. $O_{TOP}S_{NOM}V$

Hanako-wa Taro-ga mitsuke-ta.
Hanako-TOP Taro-NOM find-PAST
'As for Hanako, Taro found her.'

3.2. Procedure

First, examples were accumulated using regular expressions from *Aozora Bunko*. With the second, *noda*-construction samples, an embedded clause, or a double object construction were eliminated from the analysis. Third, the frequencies of each condition were measured.

3.3. Data Analysis

A series of Chi-square tests were conducted for each conditions.

3.4. Results

The summary of observed frequencies for each conditions are shown in the Table 1. Series of Chi-square tests were carried out to see if there are significant differences among them. First, it has been revealed that there were a significant difference in the overall data ($\chi^2(3)=5550.447, p<.001$). Second, the frequency of SOV was significantly higher than that of OSV ($\chi^2(1)=2733.224, p<.001$). Third, there were significant differences between the frequencies of $S_{TOP}O_{ACC}V$ and $S_{NOM}O_{ACC}V$ ($\chi^2(1)=1448.177, p<.001$), $S_{NOM}O_{ACC}V$ and $O_{TOP}S_{NOM}V$ ($\chi^2(1)=311.143, p<.001$), and $O_{TOP}S_{NOM}V$ and $O_{ACC}S_{NOM}V$ ($\chi^2(1)=6.86, p<.001$).The result shows, $S_{TOP}O_{ACC}V$ occurred more frequently than $S_{NOM}O_{ACC}V$, which in turn rose more frequently than $O_{TOP}S_{NOM}V$, whose frequency was higher than that of $O_{ACC}S_{NOM}V$.

Table 1. Summary of Observed Frequencies for Sentence Types

Sentence Types	Number
$S_{NOM}O_{ACC}V$	450
$S_{TOP}O_{ACC}V$	2526
$O_{ACC}S_{NOM}V$	30
$O_{TOP}S_{NOM}V$	54
Total	3060

4. Discussion

The results of the analysis shows that SOV occurred more frequently than OSV. This supports the thesis of Kuno (1973) and Miyajima (1964) in that the frequency of OSV is much lower than that of SOV. This fact can be explained by *Preferred Argument Structure*. Note that subjects in OSV are apt to be newer than their objects (Kuno 1978). Therefore, OSV sentences in Japanese tends to violate the notion of *Preferred Argument Structure,* in that the topical referent is realized as object, not as subject, and thus it is a marked option. This may be the reason why the frequency of OSV is low.

Next, it has been revealed that $S_{TOP}O_{ACC}V$ (2526 examples) occurred more frequently than $S_{NOM}O_{ACC}V$ (450 examples) but that there were no significant differences between the frequency of $O_{ACC}S_{NOM}V$ (30 examples) and $O_{TOP}S_{NOM}V$ (54 examples) in Japanese novels. The ratio of $S_{NOM}O_{ACC}V$ to $S_{TOP}O_{ACC}V$ was 1:5.6, whereas that of $O_{ACC}S_{NOM}V$ to $O_{TOP}S_{NOM}V$ was 1:1.8. This fact also coincides with the notion of information structure and argument structure. First, *Preferred Argument Structure* by Du Bois (1987, 2003) anticipates that transitive subjects are given information. In addition, Lambrecht (1996:131) points out that there is a correlation between subjectivity and topicality. Moreover, Mak, Wonk, and

Schriefes (2008) claim that topical referents tend to be recognized as subject. Taking their observation into account, *WA*-marked subject is more preferable to *GA*-marked subject because subject is a desirable and natural position for topical referents. Therefore, *GA*-marked subject is more marked than *WA*-marked subject. This may be the reason why the frequency of $S_{TOP}O_{ACC}V$ is much higher than $S_{NOM}O_{ACC}V$. Second, the frequency of $O_{ACC}S_{NOM}V$ is not so different from that of $O_{TOP}S_{NOM}V$. Recall that the object position is not strongly constrained by the information structure (Du Bois 1987, 2003). Thus, it does not matter whether the object is *WA*-marked or *O*-marked. This possibly explains the fact that there were no drastic differences between the frequency of $O_{ACC}S_{NOM}V$ and $O_{TOP}S_{NOM}V$.

However we must note that intransitivity has not been taken into account in the analysis. In contrast to *Preferred Argument Structure*, it is conceivable for both intransitive and transitive subjects as a whole tend to provide given information. For this reason, it is not clear whether only transitive subjects are constrained by givenness or both transitive and intransitive subjects are strongly influenced by givenness. The former prediction relies on Du Bois (1987, 2003) and the latter prediction counts on Lambrecht (1996) and Mak, Wonk, and Schriefes (2008). Therefore, further studies are needed in elucidating the issue.

5. Conclusion

In this paper, I have investigated the relationship between grammar and discourse in terms of frequency. Consequently, it has been demonstrated that the frequency of $S_{TOP}O_{ACC}V$ was much higher than that of $S_{NOM}O_{ACC}V$, but the difference in frequency between $O_{ACC}S_{NOM}V$ and $O_{TOP}S_{NOM}V$ was not so significant in Japanese novels. This fact can be explained by *Preferred Argument Structure* (Du Bois 1987, 2003). In SOV, the transitive subject positon is desirable for topical referents. Thus, they tend to be realized as *WA*-marked referents. This is why $S_{TOP}O_{ACC}V$ occurred much more frequently than $S_{NOM}O_{ACC}V$. $S_{TOP}O_{ACC}V$ is unmarked because it follows *Preferred Argument Structure*. In OSV, on the other hand, topical referents are realized as objects and transitive subjects are non-topical referents. This tendency violates the concept of *Preferred Argument Structure*. This is why the frequency of OSV is much lower than that of SOV. In that meaning, OSV is more marked than SOV. Furthermore, comparted to the difference between $S_{TOP}O_{ACC}V$ and $S_{NOM}O_{ACC}V$, there were fewer differences between $O_{TOP}S_{NOM}V$ and $O_{ACC}S_{NOM}V$ in frequency. The reason behind this result is that the object position is not constrained by *Preferred Argument Structure* and thus it is not strongly affected by givenness.

Acknowledgement

This paper is devoted to Prof. Shinich Chigusa., who supervised me when I was an undergraduate and I am grateful to him for his support before leaving Japan. I am sure that my years of contact with him was the major factor in bringing me to my present position. Enlightened by his lectures, I have conducted literature reviews to acquire knowledge of pragmatics, which has become the basis of my research now. Next, I am very grateful for Kai Salisbury for proofreading this article. It has helped to increase the articulateness of the texts in this paper.

References

Birner, B. J, Kaplan, J. P., & Ward, G. (2007) Functional compositionality and the interaction of discourse constraints. *Language* 83: 317-343.

Birner, B. J., & Mahootian, S. (1996) Functional constraints on inversion in English and Farsi. *Language Sciences* 18, 127-138.

Du Bois, J. W. (1987) The discourse basis of ergativity. *Language* 63: 805-855.

Du Bois, J. W. (2003) Argument Structure. In Du Bois, J. W., Kumpf, L. E., & Ashby, W. J. (eds.), *Preferred argument structure: grammar as architecture for function*. John Benjamins Publishing.

Imamura, S. (2014) The Influence of givenness and heaviness on OSV in Japanese. In Aroonmanakun, W., Booonkwan, P., & Supnithi, T. (eds.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation*, 224-233. Bangkok: Chulalongkon University.

Imamura, S. & Koizumi, M. (2011) A centering analysis of word order in Japanese, *The Tohoku Linguistic Society* 20: 59-74.

Imamura, S., Sato, Y., & Koizumi, M. (2014) Influence of Information Structure on Word Order Change and Topic Marker *WA* in Japanese. In Aroonmanakun, W., Booonkwan, P, & Supnithi, T. (eds.), *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation*, 432-441. Bangkok: Chulalongkon University.

Kaiser, E., & Trueswell, J. C. (2004) The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94: 113-147.

Kuno, S. (1973) *The structure of the Japanese language*. Cambridge, Mass: MIT Press.

Kuno,S. (1978) *Danwa-no Bunpō* [Grammar of Discourse]. Tokyo: Taishūkan.

Lambrecht, K. (1996) *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Publishing.

Mak, W. M., Vonk, W. & Schriefers, H. (2008) Discourse structure and relative clause processing. *Memory & Cognition*, 36: 170-181.

Miyagawa, S. (2001) The EPP, scrambling, and *wh*-in-situ. *Current Studies in LInguistics Series*, 36: 293-338.

Miyagawa, S. (2003) *A-movement scrambling and options without optionality*. In Karimi, S. (ed), Word order and scrambling, 177-200. Blackwell Publishers.

Miyagawa, S. (2010) *Why Agree? Why Move? Unifying Agreement-based and Discourse-configurational Languages*. Cambridge: MIT Press.

Miyajima, T. (1964) *Joshi jodōshi no yōhō. Gendai zasshi kyūjyussyu no yōgo yōji* [The usages of articles and auxiliary verbs: Ninety-nine technical terms and characters in modern magazines],

Rambow, O. (1993) Pragmatic aspects of scrambling and topicalization in German. In Workshop on Centering Theory in Naturally-Occurring Discourse. Philadelphia, PA, May 1993.

Saito, M. (1985) *Some Asymmetries in Japanese and their Theoretical Implications*. Doctoral dissertation, MIT.

Saito, M. (2009). Optional A-scrambling. *Japanese/Korean Linguistics*, 16: 44-63.

Saito, M. & Hoji, H. (1983) Weak cross over and move α in Japanese. *Natural Language and Linguistic Theory*, 1: 245-259.

有標性と項構造

今村怜

和文要旨

「SはOをV」と「SがOをV」では頻度において大きな乖離があったのに対し、「OはSがV」と「OをSがV」の間には大きな差が観察されなかった。これは、他動詞の主語位置がトピックにとって望ましい位置と考える、Preferred Argument Structure によって説明ができる。また、OSV 語順の全体としての頻度が低かったのは、トピック要素が目的語位置で実現されているという点で有標性を帯びているからであると考えられる。