CONSTITUENT ORDERING IN SPOKEN JAPANESE

by

Yumi Yamasaki
October 19, 2005

A dissertation submitted to the
Faculty of the Graduate School of
the State University of New York at Buffalo
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

Department of Linguistics

UMI Number: 3203907

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgments

First and foremost, my deepest appreciation and thanks go to my adviser and mentor of many years, Matthew S. Dryer for his tremendous guidance and enormous patience. His vast and technical knowledge of many different languages, which he shared with me, contributed to the development and completion of my project. He kept me in the right direction in every stage of the project. Without his insightful advice, comments and inspiration from the start of the project, throughout researching and writing the dissertation, I would not have been able to complete this thesis.

I also would like to thank the other members of my committee, Mitsuaki Shimojo for graciously sharing his knowledge on Japanese linguistics, and detailed comments and guiding discussions on Japanese data, and likewise, David A. Zubin for his invaluable comments and assistance on methodological issues, and his training in psycholinguistics classes. I also would like to say thank you to Karin E. Michelson, chair of the Linguistics graduate program, for her strong support on my finishing the dissertation, Jeri J. Jaeger for her support and encouragement throughout my time at UC Davis and SUNY Buffalo, Robert Van Valin, Jr. for his wonderful teaching in syntax at UC Davis and SUNY Buffalo, and Jean-Pierre Koenig for his support on my getting the degree. My very special thanks go to Carole Orsolits, assistant to Chair in Linguistic Department at SUNY Buffalo, for her constant support and encouragement on my achieving the goal.

Many thanks go to other professors at SUNY Buffalo: The late Paul Garvin, Wolfgang Wölck, Madeleine Mathiot, Leonard Talmy, David Shucard and Douglas Roland. Thanks also go to my professors at UC Davis who got me started on linguistics and lead me to complete a master thesis: David P. Wilkins, Janet S. Shibamoto Smith, Patrick Farrell, Nina Dronkers, Wilbur A. Benware, and Maria I. Manoliu. I am also indebted to the professors at UC Santa Barbara, Wallace Chafe, John W. DuBois, and Sandra A. Thompson, who gave me opportunities to audit their classes and to participate in linguistic discussions while I was a language instructor at UCSB.

I also like to thank you to my fellow students at UCD, UCSB and SUNY Buffalo for their inspiration, support and friendship, particularly to Mitsu Shimojo, Kean Kaufman, Ruth Shields, Brenda Redfern, Kanae Fujiwara, Wataru Nakamura, Isao Honda, Tomiko Kodama, Noriko Watanabe, Ki-Seong Park, David Houghton, Valerie Shafer, Toshi Nakayama, Kumiko Ishibashi, Tsuyoshi Ono, and Satoko Suzuki.

I am indebted to the World Languages Institutes at SUNY for their financial support throughout my graduate career in Buffalo. In particular, I wish to thank Mark A. Ashwill, Akiko Sugiyama, Kazue Masuyama, Akemi Sakamoto, and Sachiko Hiramatsu for their moral support and friendship. Thanks also to Clare Cucino who reviewed my draft while I resided in Beijing.

Finally my biggest thank you goes to my family, particularly to my husband Steve, who believed in my finishing the dissertation, for his patience, understanding and support in many ways. My parents, Kaoru and Hisako Yamasaki encouraged me to go to graduate school, and they made it possible. I owe this accomplishment to them.

Table of Contents

Abstract

In speaking, the speaker has a general idea about what they are going to talk about. They know how to form a grammatical sentence, and they know what grammatical roles they are going to assign to who and what. In a language like Japanese where the assignment of grammatical roles does not involve constituent ordering, the speaker has to decide in which constituent ordering they deliver the sentence. For example, when Japanese speakers want to say "Fred saw John", they first assign case particles, possibly the nominative case particle *ga* for "Fred" and the accusative case particle *o* for "John". The speaker then has to decide in which order, *Fred-ga John-o* or *John-o Fred-ga*, they will deliver the sentence.

I first investigate three factors that are pertinent to the choice of one constituent ordering over another: recency of previous mentions, cataphoric persistence and syntactic heaviness of the constituent. Recency is quantified by Referential Distance (RD), cataphoric persistence by Topic Persistence (TP), and syntactic heaviness by Syntactic Weight (SW). The relative measurement of RD, TP and SW between the two constituents in question indicates that the constituent with a lower RD, a higher TP or a higher SW tends to come before the constituent with a higher RD, a lower TP or a lower SW. Based on the patterns between factors and the earlier position, algorithms are formulated to examine if the algorithms can correctly predict the ordering choice between SOV and OSV. Applying the algorithms to the data shows not only that single predictors, RD, TP and SW, can make correct predictions, but also that the combination of three predictors can predict with a higher success rate than the single predictor or the combination of two

predictors. In addition, other factors, the focus structure, the use of the particle *wa* and demonstratives, the light verb construction, and scene-setting and topical elements, are explored. The algorithms based on the factors are further applied to clauses containing the subject, the direct object and other phrases to examine if the factors are relevant to predicting what comes first or what precedes what.

Notes on transcriptions

The transcription of Japanese words in this thesis follows the Kunrei-shiki romanization. Person's names and titles of literature are transcribed in the Hepburn system. The following examples demonstrate some differences between the two systems: *ti/chi, tu/tsu, hu/fu, sya/sha, zya/ja, si/shi, zi/ji, tya/cha* (the Kunrei system/the Hepburn system). In some cited Japanese examples, I change the original romanization to keep the transcription consistent. In some English cited examples, I change the original English gloss to keep the labeling consistent throughout the thesis. The question mark symbol [?] indicates that the sentence is odd for the average native speaker. The following abbreviations will be used throughout this thesis:

ABL	ablative
ACC	accusative
ADJ	adjective
ADV	adverb
AUX	auxiliary verb
CAU	causative
CC	complement copula
CJE	conjecture
CMPL	complementizer
COM	comitative
CONF	confirmation
CONJ	conjunction
COP	copula
DAT	dative
DEM	demonstrative
FP	final particle
GEN	genitive
HON	honorific
HUM	humble
IMP	imperative
INS	instrumental
LOC	locative
NEG	negative
NMLZ	nominalizer
NOM	nominative
NPST	non-past
PASS	passive
PL	plural
POL	polite
POT	potential
PPG	present progressive
PrNom	prenominal form
PRO	pronominal
PST	past/perfective
Q	question
QT	quotative
RES	resultative
Sg	singular
STAT	stative
TE	te-form of copula or verb
TEMP	temporal
To	topical element
TOP	topic marker
Vex	existential verb
Vi	intransitive verb
VOL	volition

Notes on pronominalization

I will use masculine pronominalization for ‘the speaker’ or ‘speakers’, whether male, female or generic. ‘The hearer/addressee’ or ‘hearers/addressees’ will receive feminine pronominalization, whether male, female or generic. When ‘the speaker’ or ‘the hearer’ denotes a group of speakers or hearers, I will use third plural pronominalization. When the sex of ‘the speaker’ or ‘the hearer’ is known, I will use masculine or feminine pronominalization accordingly.

Chapter 1

Introduction

One crucial attribute of spoken language is linear order in the temporal continuum. Spoken language, as spontaneous as it may sound to the ear, is the result of complex computing in the human brain. This complex computing is performed in a split second under linguistic and cognitive constraints, such as the limited capacity of short-term memory. What and how to compute is the ultimate question of how the human brain functions in order to form language, though there is no way of understanding the entire process. Previous research on this question has identified some linguistic properties which may correspond to certain cognitive functions like 'attention'. When a speaker wishes to communicate in the most economical way, there are a number of things that the speaker typically considers. In order to facilitate the hearer's comprehension, the speaker has to know what is in the hearer's consciousness, what the hearer knows about what the speaker is, or will be, talking about. This is a comprehension-based perspective. The speaker should also know what information he wants to convey the most by first sorting information in his consciousness. This is the speaker's intention about what the speaker is or will be talking about. Recent research reveals that not only the discourse status, but also the cognitive status of the referent of an entity that the speaker is talking about, plays an important role in determining a linguistic form in language production. Word order is a linguistic form that involves linear order in the temporal continuum. Therefore, studying the constituent ordering variation and its determining factors provides valuable information about language processing and production.

1.1. Objectives and outline

The Japanese language is known to be typical of an SOV (Subject-Direct Object-Verb) language (e.g. Greenberg 1963). Recent studies reveal that Japanese exhibits a certain degree of word order flexibility, including both scrambling[1] and postposing[2].

---

[1] “Scrambling” should be distinguished from “topicalization”. “Topicalization” in Japanese is a morpho-syntactic device of signaling a sentential topic by marking a phrase by the particle *wa*. The topicalized phrase is usually at the sentential initial position. The following is an example of a sentence with a topicalized phrase.

(a)	mottomo	ano	zyurassiku	paaku	tte	no	**wa**	watasi
	though	that	Jurassic	Park	QT	NMLZ	**TOP**	1sg
	hon	o	yomimasita		kedo	ne.		
	book	ACC	read.POL.PST		though	FP		

‘Talking about Jurassic Park, I read the book.’

There are three constituents in sentence (a). One is a topicalized phrase marked by *wa*, and the other two are subject and direct object that are subcategorized arguments of the verb *yomu* ‘to read’. i.e. the topicalized phrase is a peripheral. In this thesis, topicalized elements such as in (a) are categorized oblique noun phrases (i.e. other phrases than the subject and the direct object) in the ≥3NPs construction.

In some cases scrambling and topicalization occur in one grammatical process. This is the case when topicalization involves both fronting of a constituent and marking it by the particle *wa*. Sentence (b) involves the fronting of direct object noun phrase *kore o* ‘this-ACC’, and replacing the accusative *o* with the particle *wa*.

(b)	kore	wa	sakka	no	K-san	ga	nanika	de	kaita-n-desu	yo	ne
	this	TOP	writer	GEN	K-Mr	NOM	something	LOC	write.PST-NMLZ-COP	FP	FP

‘Talking about this, Mr. K wrote it for something.’

(c)	sakka	no	K-san	ga	kore	o	nanika	de	kaita-n-desu	yo	ne
	writer	GEN	Mr.K	NOM	this	ACC	something	LOC	write,PST-NMLZ-COP	FP	FP

‘Mr. K wrote this for something.’

The fact is that Sentences (b) and (c) are not allosentences. Sentence (b) has a different semantic reading from (c) in that, as Kuno (1973) pointed out, there is no exhaustive reading of the subject *K-san ga* ‘Mr. K-NOM’. In this study, Sentence (b) was alanyzed as a direct object-subject-locative-verb clause, and Sentence (c) as a subject-direct object-locative-verb clause.

[2] In Japanese postposing commomly refers to the placement of noun phrases, adverbial phrases and adjectival phrases after the final verb.

'Scrambling' in Japanese refers to the rearrangement of phrasal units[3] in a clause. The following sentences (1) are (2) are an example of scrambling.

(1)	taroo	ga	hanako	o	mita
	Taro	NOM	Hanako	ACC	saw
(2)	hanako	o	taroo	ga	mita.
	Hanako	ACC	Taro	NOM	saw

Sentences (1) and (2) both depict a scene where *Taro* is looking at *Hanako* (i.e. 'Taro saw Hanako.'). They are allosentences: The propositions which (1) and (2) represent are truth-conditionally identical. In (1) and (2), the grammatical relations are preserved, yet (1) and (2) have two different word orders. (1) has SOV (Subject-Direct Object-Verb) order (or Actor-Undergoer-verb order), and (2) has OSV (Direct Object-Subject-Verb) order (or Undergoer-Actor-verb order). Past research, as well as the data in this study, show that the SOV order such as in (1) is a dominant order[4], and the OSV order such as in (2) is rather a deviation. For example, Kuno (1973) states that the ratio of occurrence between the SOV order and the OSV order in Japanese is 17:1. My data shows the ratio of 3.7:1. The question naturally arises, then in what case does the speaker utter Sentence (1) and in what case does the same speaker utter Sentence (2). The native speaker's immediate

---

[3] A phrasal unit, which has been used in school grammar in Japan, is a way of grouping words together according to the function. One unit is, for example, an NP + a particle, a predicate, or an adverbial phrase. Therefore, the following sentence contains 3 phrasal units, i.e. hanako wa, hon o, and katta.

(d)	hanako	wa	hon	o	katta.
	Hanako	TOP	book	ACC	buy.PST
	'Hanako bought a book.'				

[4] In some literature the SOV order in Japanese is labelled as canonical order. My data indicates in every defined parameter the occurrence of the SOV order is more frequent than the OSV order. In that sense, the SOV order in Japanese should instead be called the default order. For example, the Japanese Language Engine (Kameyama 1995) does not posit any underlying constituent order from which scrambled ordering yields. My position, however, is not to deny the canonicity. The canonicity in word order is proved to exist by scientific studies of brain function, using event-related brain potential (ERP), magnetic resonance imaging (MRI), or positron emission tomography (PET), with respect to the filler-gap dependencies in sentence processing (e.g. Ueno and Kluender 2003).

response to the question is that in (1) the speaker is talking about *Taro* and in (2) about *Hanako*, or they say that the speaker's attention is on *Taro* in (1), and on *Hanako* in (2). The native speaker's intuition seems to appeal to pragmatic notions such as topic or focus. However, there have been few attempts to systematically integrate pragmatic notions as possible motivations into a theory of predicting a particular word order over another in Japanese.

This thesis will present an empirical study on the constituent ordering in spoken Japanese. It will explore those circumstances in which a particular constituent order is uttered over another. As a by-product, this study will also provide statistical data from spoken Japanese on constituent ordering of other grammatical roles as well as variations of case marking. In what follows, I will examine previous studies on word order in Japanese in different linguistic frameworks. Through the review of various frameworks, potential determining factors for the constituent ordering become visible. These potential factors are the starting point of this study.

1.2. Previous studies related to word order variation in Japanese

Linguists have paid attention to establishing parameters in order to categorize language. Since Greenberg (1966) suggested 45 universal statements based on a 30-language sample, the study of word order typology has made great progress. Research on exotic languages has lead to the discovery that some languages use word order to signal different discourse functions. Typological studies have come to classify languages into three types: languages in which the notion of grammatical relations is necessary in order to explain various linguistic phenomena; languages in which the notion of semantic roles is the most adequate to explain the phenomena; and finally pragmatically prominent languages that need pragmatic and discourse notions to explain the phenomena. Studies of the information structure of language have identified discourse notions, such as 'theme' vs. 'rheme' (e.g. Firbas 1966), 'given' vs. 'new' (e.g. Halliday 1967, Chafe 1976), or 'narrow focus' vs. 'broad focus' (e.g. Lambrecht 1994), and these notions have been used to describe the information status of discourse entities. Recent research reveals that both domains of the information structure of discourse and the information status of discourse entities are crucial to the study of constituent ordering.

Generative grammarians have also tried to find a rule that governs word order, and have posited a model of two layers of syntactic structure: surface level and deep structure. According to this model, there is an underlying word order in deep structure and syntactic rules move around the constituents to produce various surface syntactic structures. Many studies of word order in Japanese have tried to explain the ordering variation based on this model.

In the following sections, I will review some linguistic frameworks used to explain word order variation in Japanese. They are (1) the semantic approach, (2) the Prague school tradition, (3) traditional Japanese grammarians' approach, (4) the pragmatic/discourse approach, (5) psycholinguistics experimental approach, and (6) corpus analysis approach. As a supplement, I discuss briefly some studies on other SOV languages.

1.2.1. Semantic approach

One attempt to explain the constituent ordering variation is in terms of thematic role hierarchies. Thematic hierarchies are popular methodologies used to explain the parallelism between certain syntactic phenomena and semantics. The following sentences provide a basis for claiming that constituent ordering in Japanese is governed by thematic roles.

(3)	yamada-sensei	ni	wa	syakkin	ga	takusan	oarininaru.
	Yamada-teacher	DAT	TOP	debt	NOM	many	exist.HON

‘Teacher Yamada has a lot of debt.’

(4)	ude	ni	yori	o	kakeru
	arm	to-LOC	twist	ACC	hang

‘(Lit) To wring an arm. --> To get a hustle on.’

(5)	hyootan	kara	koma	ga	deta.
	a.gourd	from-LOC	a-top	NOM	came.out

‘(Lit) A top came out of a gourd. --> Unexpected things happened.’

Sentences (3), (4), and (5) represent three different sequences in terms of grammatical relations: in (3) the sequence is locative/dative - subject, in (4) locative - object, and in (5) locative - subject. The argument structure for (3), (4), and (5) are all locative - theme. Hence, Sentences (3), (4), and (5) can all be analyzed as having the sequence of locative - theme in terms of thematic roles.[5] This does not work, however, for the word order variation such as in (1) and (2), where the semantic roles are preserved. I repeat (1) and (2) below. Both sentences mean ‘Taro saw Hanako.’

(1)	taroo	ga	hanako	o	mita
	Taro	NOM	Hanako	ACC	saw

(2)	hanako	o	taroo	ga	mita.
	Hanako	ACC	Taro	NOM	saw

In both (1) and (2) the correlation of subject = actor, and object = theme is preserved. That is, the argument structure in (1) is actor – theme, whereas in (2) it is theme – actor.

---

[5] Dryer (personal communication) points out that some would claim that the pragmatic differences associated with semantic roles, rather than semantic roles per se, are a factor here. For example, at-locatives such as in (4) are in general scene-setting elements, and scene-setting elements are prone to be placed at the beginning of an utterance.

Thus, this approach by thematic roles does not completely answer the question regarding the speaker's choice of word order variation.

For existential sentences[6] with the verb *aru* 'to exist (inanimate)', Mikami (1969) points out that the different word order with respect to locative and subject encodes different semantic readings. For example, Sentences (6) and (7) can be both translated into English as 'There is a book on the desk', but (6) is concerned with the existence of the book and its location, while (7) is only concerned with the location of the book.

(6)	tukue	no	ue	ni	(wa)	hon	ga	aru
	desk	GEN	top	LOC	(TOP)	book	NOM	exist

(7)	hon	ga	tukue	no	ue	ni	aru
	book	NOM	desk	GEN	top	LOC	exist

Mikami provides a parallel example in Chinese,[7] explaining that in Chinese, in addition to the word order, different verbs, *yǒu* 'to exist/have' and *zài* 'to exist/be', are used to express the different readings.

(8)	zhuōzi	shàng	yǒu	shū
	table	top	exist	book
	'There is a book on the table'			

(9)	shū	zài	zhuōzi	shàng
	book	exist	table	top
	'The book is on the book.'			

---

[6] According to Kuno (1973), existential sentences only involve indefinite subject nouns, thus sentences like (b) are not, strictly speaking, existential sentences.

(a)	tukue	no	ue	ni	hon	ga	atta
	desk	GEN	top	LOC	book	NOM	exist.PST
	'There was a book on the desk.'						

(b)	sono	hon	wa	tukue	no	ue	ni	atta
	that	book	TOP	desk	GEN	top	LOC	exist.PST
	'That book was on the desk.'							

[7] Mikami gives examples (8) and (9) in *Hanzi* (Chinese characters). The transliteration in Pinyin is mine.

Similarly, Shibatani (1990) argues that scrambling or reordering of constituents has an effect on the semantic interpretation. The following examples (12) and (13) are from Shibatani (1990:261).

(12)	min'na	ga	dareka	o	aisiteiru
	all	NOM	someone	ACC	love.PPG
	'Everyone loves someone.'				

(13)	dareka	o	min'na	ga	aisiteiru
	someone	ACC	all	NOM	love.PPG
	'Everyone loves someone'.				

Sentences (12) and (13) have the same subject, *min'na ga* (everybody NOM), direct object, *dareka o* (someone ACC), and verb, *aisiteiru* (love.PPG) 'be loving', but in different constituent orders. Shibatani explains that (12) is ambiguous in the logical scope between "for all x, there is y, such that x loves y" (i.e. 'Everybody has the one who he loves,') and "there is y, for all x, such that x loves y" (i.e. 'There is one person who is loved by everyone.'), but (13) is only possible in the latter interpretation. Minami's argument and Shibatani's examples suggest that some constituent ordering is semantically motivated.

1.2.2. Prague School approach

Iijima (1997) attempts to explain the Complement-Subject-Intransitive Verb order[8] (hereafter CSVi) in Japanese in terms of Firbas's (1964) Communicative Dynamism. The dominant ordering of subject and complement in Japanese is Subject-Complement-Intransitive Verb (hereafter SCVi). The order CSVi is found in the last two

[8] The Subject-Complement-Intransitive Verb construction forms a closed class construction which occurs only with intransitive verbs *naru* 'to become' and *suru* 'to do'. Complement is an argument liscensed by the verb *naru* and *suru*, and marked by either the particle *ni* 'to' or *to* 'with'. The dominant order in this construction is SCVi, and CSVi is marked. See the section 2.3.1.4 for statistics in the present data.

lines of a famous poem by Miyazawa Kenji. This poem consists of 30 lines without any commas or periods. According to Iijima's analysis, this poem is divided into two parts, line 1 to 28 and the last two lines. The part in question is the last two lines, 'I want to be such a thing.', which contains the CSVi sequence.

1	ame	ni	mo	makezu
	rain	by	also	be.defeated.NEG
	'Not being discouraged by rain'			
.				
.				
29	sooiu	mono	ni	
	such	thing	to	
	'to such a thing'			
30	watasi	wa	nari-tai	
	1sg	TOP	become-want	
	'I want to be'			

Firbas's Communicative Dynamism has three stages depending on the degree of force to push the development of communication forward. The minimum stage of Communicative Dynamism is called Theme (Th), the maximum stage is Rheme (Rh), and the stage in the middle is Transition (tr). Firbas assumes that the universal and common distribution of Communicative Dynamism through languages is from the minimum to the maximum, i.e. Theme-Transition-Rheme (Th-tr-Rh), and he calls it Natural Order. The flow in Natural Order, Th_1-tr_1-Rh_1 = Th_2-tr_2-Rh_2 ('=' indicates the sentence linkage, and subscripts 1 and 2 indicate sentence 1 and sentence 2, respectively), guarantees a smooth flow of communication because Rheme in sentence 1 (Rh_1) and Theme in sentence 2 (Th_2) are adjacent without any obstruction.

Communicative Dynamism is distributed across constituents of a sentence, and the interaction of these three stages in a sentence is called Functional Sentence Perspective (FSP). FSP realizes actual word orders, which may or may not be different from the basic word order (I assume 'basic' means 'dominant' in terms of frequency. The SOV order in SOV languages, for example, is basic order). FSP, in other words, is a motivating factor for word order variations.

Iijima claims that the normal distribution of Communicative Dynamism in Japanese is Th-Rh-tr, and the FSP places the maximum Communicative Dynamism (the most important newsworthy information, or the focus in Iijima's interpretation) on the immediate preverbal slot.[9] According to Iijima, this means that Rheme should be in the immediate preverbal position. Iijima supports his claim on the Th-Rh-tr flow by the frequent use of connectives, and the tendency to omit the sentence-final copula in Japanese. In the flow of Th_1-Rh_1-tr_1 = Th_2-Rh_2-tr_2 ('=' indicates the sentence linkage, and subscripts 1 and 2 indicate sentence 1 and sentence 2, respectively) such as in Japanese, the linkage of sentence 1 and sentence 2 is distracted by the intervention of tr_1 between Rh_1 and Th_2, resulting in a non-smooth flow of Rheme -Theme. Iijima reasons that the use of connectives and the ellipsis of the copula at the end of a sentence help make the linkage smoother.

Based on his claim, Iijima imposes the Th-Rh-tr flow on the possible word order variations for the last two lines of the poem. The three variations are shown in (14a), (14b), and (14c). The constituent ordering in (14a) is a dominant order SCVi. (14b) involves postposing of C and (14c) involves preposing of C.

---

[9] Iijima states that this pattern is observed in other SOV languages. He does not, however, specify the source of his claim.

(14) watasi wa sooiu mono ni nari-tai
1sg TOP such thing to become-want
'I want to become such a thing.'

a watasi wa sooiumono ni naritai (SCVi)
Th Rh tr

b watasi wa naritai sooiumono ni (SViC)
Th Rh tr

c sooiumono ni watasi wa naritai (CSVi)
Th Rh tr

Iijima assumes that the first part of the poem (lines 1-28) is an independent unit, and the unit as a whole is considered to be Rheme. He links the first part of the poem (lines 1-28) and the last two lines in Communicative Dynamism schema. In (15), Rh_1 stands for the lines 1-28 of the poem. Th_2 - Rh_2 - tr_2 in (15a), (15b), and (15c) correspond to (14a), (14b), and (14c) respectively. Underneath Th_2 - Rh_2 - tr_2, S stands for subject, C for complement, and V for verbal.

(15)

a. Rh_1 Th_2 - Rh_2 - tr_2
S C V
|__________|

b. Rh_1 Th_2 - Rh_2 - tr_2
S V C
|______________|

c. Rh_1 = Th_2 - Rh_2 - tr_2
C S V

From Iijima's claim, (15b) is ruled out because Rh_2 is not in the preverbal position. That gives us the choice between (15a) and (15b). He concludes that the poet chose the construction (15c) that creates the most efficient FSP, realizing the direct flow of information and the emphasis on the subject *watasi wa* 'I' in the immediate preverbal position.

In sum, Iijima's study was an attempt to account for the sentence initial complement *sooiu mono ni* 'to such a thing', preposed before the subject *watasi wa* 'I'. Given the fact that the noun phrases containing a demonstrative such as *sooiu* "such" in (14) tend to be preposed in Japanese, his account might find its value in explaining why noun phrases with demonstratives are placed at the beginning of a sentence.

Kuno (1973) provides a similar explanation for the constituent ordering with respect to the existential sentences. Following (10) and (11) are Kuno's examples of the locative and subject alternation in Russian.

(10)	Na	stole	kniga.
	on	table	book
	'There is a book on the table.'		

(11)	Kniga	na	stole.
	book	on	table
	'The book is on the table.'		

He claims that in many languages the preferred order for existential sentences is indefinite things (subject) – location (locative), and that it is so because it forms an uninterrupted chain of information, new information (indefinite things) --> old information (specific locations), in discourse, as diagramed in (12), wherein the character "S" stands for "sentence".

(12) [Old New]$_{S1}$ [Old New]$_{S2}$ [Old New]$_{S3}$

The new information introduced at the end of the first sentence in discourse, S1, is carried on to the next sentence, S2, as old information at the beginning of this sentence, and at the end of the second sentence, S2, another piece of new information is introduced, and so on.

Whether a language has the Theme - Rheme pattern or the Rheme - Theme pattern has been argued in the cases of different languages. Some studies support the Prague school Theme - Rheme tradition (e.g. Halliday 1967), and others indicate the Rheme - Theme pattern among flexible word languages (e.g. Tomlin and Rhodes 1979, Givón 1983). It seems that both patterns have cognitive explanations.

1.2.3. Traditional Japanese grammar

The most intensive study on word order in Japanese comes from the traditional Japanese grammarians (e.g. Sakakura 1966; Saeki 1960, 1975, 1976, 1998; Miyajima 1963, 1964; Kitahara 1970). In traditional Japanese grammar, word order has been captured in terms of the *Kakari-Uke* relationship. *Kakari-Uke* is a relative relation where *Kakari* is 'a throw', and *Uke* is 'a catch'. At the clause level, *Uke* is basically the predicate, and *Kakari* is a clausal constituent other than the predicate. At the sentence level, the subordinate clause is *Kakari*, and the matrix clause is *Uke*. The dominant order is defined as the one that has a flow from *Kakari* to *Uke* on the assumption that the *Kakari-Uke* flow approximates the consciousness of Japanese speakers. What follows is an example of the *Kakari-Uke* schema. The double underline indicates *Kakari*, and the single underline indicates *Uke*. The relation between *Kakari* and *Uke* are not absolute, but relative.

(16) kyonen no aki watasi wa amerika kara nihon e modotte-kimasita.
last-year GEN autumn 1sg TOP America from Japan to return-came
"Last fall I came back to Japan from America."

kyonen no aki watasi wa amerika kara nihon e modotte-kimasita.
watasi wa amerika kara nihon e modotte-kimasita.
amerika kara nihon e modotte-kimashita.
nihon e modotte-kimashita.

Saeki is one of the earliest linguists who paid attention to the order of preverbal constituents in Japanese. In the above sentence, for example, preverbal constituents are *kyonen no* 'last fall', *watasi wa* 'I', *amerika kara* 'from America', and *nihon e* 'to Japan', which are preverbal constituents. His analysis of constituent order variation identifies motivations for certain constituent ordering. Saeki (1975) retrieved preverbal constituents marked by *ga* (nominative), *o* (accusative and locative), *ni* (locative and temporal), *e* (locative), *to* (comitative), *kara* (allative/locative and temporal), and *de* (locative and instrumental) from a total of 67 pages of novels by four writers. First, he investigated the order of noun phrases that have the same length of phrasal units[10] and found the following five tendencies.

Tendency 1: Temporal and locational noun phrases < other noun phrases.
Tendency 2: Temporal *ni* and *kara* < locative *ni*, *de*, *kara*, and instrumental *de*.
Tendency 3: *ga*-marked Subjects < Noun phrases other than locative and temporal.
Tendency 4: Dative *ni* < *o*-marked Object.
Tendency 5: Allative/locative *kara* < locative *ni* and *de* indicating the destination.

---

[10] The following (a) is an example sentence in which the constituents have the same length.

(a) syuzin ga saikun o furo ni yatta ato
husband NOM wife ACC bath LOC sent after
'After the husband let his wife go to bathe herself,'

Three noun phrases in question, *syuzin ga* 'husband' (subject), *saikun o* 'wife' (direct object) and *furo ni* ' to the bathroom' (locative) are all one phrasal unit length.

He proposed two parameters to account for the above tendencies. One is what he calls the Depth of *Kakari* elements, and the other is what he calls the Scope of *Kakari* elements.

The Depth of *Kakari* elements concerns the distance of an entity (*Kakari*) from its predicate (*Uke*). The *Kakari* element that has a longer distance from its corresponding *Uke* comes earlier in the sentence than the *Kakari* element that has a shorter distance from its *Uke*:

(17)

tookyoo	wa	kinben-na	hito	ga	takusan	ite	issyookenmei
Tokyo	TOP	diligent	people	NOM	many	be-and	very-hard

hataraite-iru	no	da	ga,	mati	mo	kitanai	si,
work-PPG	NMLZ	COP	but	town	also	be.dirty	and,

miti	mo	kitanai.
streets	also	be.dirty.

‘As for Tokyo there are many diligent people and they are working hard, but the town is dirty and the streets are dirty too.’

In (17) *tookyoo wa* ‘As for Tokyo’ (*Kakari* A) corresponds to the last part of sentence *mati mo kitanai si, miti mo kitanai* ‘the town is dirty and the streets are dirty too’ (*Uke* A), and *kinben-na hito ga* ‘diligent people’ (*Kakari* B) corresponds to its predicate *ite* ‘be-and’ and *hataraiteiru* ‘working’ (*Uke* B). That is, the distance between *Kakari* A and *Uke* A is farther than the distance between *Kakari* B and *Uke* B. In Saeki’s term, *tookyoo wa* (*Kakari* A) is ‘deeper’ than *kinben-na hito ga* (*Kakari* B), and thus *Kakari* A comes before

Kakari B. Saeki reports a number of order patterns that show a reversed order[11] but treats them as statistically dismissible deviations.

Miyajima (1963) employed a quantitative way of determining the distance between *Kakari* and *Uke*. He introduced a formula (18) to calculate the value of each *Kakari* element in terms of the distance from *Uke*:

(18) $T = 2r - (n + 1)$

T is a score of the r-th *Kakari* element from *Uke* in a sentence where the number of *Kakari* elements is n. In (19), *taroo wa* 'Taro' is the 4th element from *Uke*, *at-ta* 'met' in the sentence where there are 4 *Kakari* elements. The value of r is 4, and n is 4. Thus, T (taroo wa) = 2x4 – (4+1) = +3.

(19)

taroo	wa (+3)	kinoo (+1)	eki	de (-1)	hanako	ni (-3)	at-ta
Taro	TOP	yesterday	station	at	Hanako	DAT	meet-PST

'Taro met Hanako at the station yesterday.'

The bigger the score is, the more likely that the *Kakari* element comes earlier in a sentence. Miyajima, using this formula, examined extract passages. His result showed the general tendency of the order, which conforms with Saeki's claim:

---

[11] The sequence of (a) is subject-locative (from) and in (b) locative (from)-subject. He points out the difference is that in (a) the subject is human, and in (b) the subject is an abstract concept.

(a)

yuunouna	sinsingaka	ga	aituide	furansu	kara	kichoshite,
capable	rookie-painters	NOM	successively	France	from	come-home,

fobisumu	no	hana	o	nigiyakani	sakaseteiru
Fauvism	GEN	flower	ACC	cheerfully	making-bloom

'Capable rookie painters successively came home from France, and are cheerfully blooming the flower of fauvism.

(b)

me	kara	wa	keibetsu	ga	kiete,
eye	from	TOP	despite	NOM	disappear,

magaimono	no	sinkensa	ga	nozoiteita
fake	GEN	seriousness	NOM	peeping.

'From eyes despite disappears, and fake seriousness is showing.'

A --> B (A comes before B)

Time	-->	Place-*de*
Place-*ni*	-->	Subject-*ga*
Subject-*ga*	-->	Descriptive-*to*
Subject-*ga*	-->	Object-*o*
Object-*o*	-->	Complement Copula

The Scope of *Kakari* elements is based on Miyajima's study (1964), and it concerns the semantic link between an argument and its predicate. Miyajima examined the co-occurrence of the locational particle *o* and its predicate. The locational particle *o* marks the location as in *miti o magaru* 'to turn the street', and it occurs with a certain group of verbs. Miyajima discovered in one study that 21 of 49 locational *o* occur with the verb *deru* 'to leave', as in *ie o deru*, 'to leave the house'; that 4 out of 49 locational *o* occur with the verb *tatu* 'to stand/depart', as in *Tokyo o tatu* 'to leave Tokyo'; that 3 out of 49 locational *o* co-occur with the verb *saru* 'to leave', as in *Nihon o saru* 'to leave Japan'; and that 3 out of 49 locational *o* co-occur with the verb *oriru* 'to descend', as in *yama o oriru* 'to descend the mountain'. He also examined the occurrence of the locative particle *e* and its predicates. He found that each of the 25 instances of locational *e* occurs with a different verb. From this observation he quantified the degree of a specific *Kakari* element (such as a noun phrase marked by locational *o*) occurring with a specific predicate, and noted a tendency that *Kakari* elements with a higher concentration on a specific predicate are placed closer to the predicate than the *Kakari* elements that do not have a special connection to a specific predicate. For example, in the above instance of the *o* and locative *e*, noun phrases marked by locational *o* tend to be placed closer to their predicate while the locative *e* does not have such a tendency.

In addition, Saeki (1975) looked at the noun phrases that are not included in his initial study of 67 pages of novels by four different writers. These noun phrases are either marked by the particles *wa* (topic marker) or *mo* (particle denoting a sense of addition 'also') or contain a demonstrative. He ignored the difference in phrasal length and compared the noun phrases whose lengths were different. Saeki drew the following five tendencies from this study.

Tendency 6:	Long noun phrases < short noun phrases[12]
Tendency 7:	Noun phrases with demonstrative[13] < Noun phrases without demonstrative
Tendency 8:	When Noun phrase 1 (NP1) affects the meaning of Noun phrase 2 (NP2), the order is NP1 < NP2
Tendency 9:	When the cohesion of a noun phrase and the verb exists, the noun phrase with cohesion < the noun phrase without cohesion
Tendency 10:	*wa*-marked noun phrases < other noun phrases

For Tendency 6, he argues that introducing long noun phrases later in a clause confuses the identification of the *Kakari* element and its corresponding *Uke*, and that writers try to prevent this confusion by preposing the longer noun phrases. He also reasons that longer noun phrases are preposed because they exhibit the characteristics of connectives, which usually come in the initial position. This reasoning requires more explanation. According to Saeki, a noun phrase consisting of a noun and a case particle carries its most important information in the noun, not in the case particle. Similarly, a connective consists of a verb and a connective particle, and the verb carries the central information of the connective.[14] Example (20) is from Saeki (1975:116):

---

[12] According to Saeki, Mitsuya (1908) mentioned the tendency of long subordinate complements being placed before the subject and verb.

[13] The term 'demonstrative' here includes demonstrative pronouns such as *sore* 'it' and the nominals containing a demonstrative adjective such as *sono* 'that' as in *sono hon* 'that book' and a demonstrative modifier such as *son'na* 'such that' as in *son'na hon* 'a look like that'.

[14] It seems that Saeki uses the term 'connective' in a broad sense including independent connectives and connective phrases.

(20)

umi	ni	nozonda ……	gake	no	tottan	de
sea	Loc	faced ……….	cliff	GEN	edge	LOC

suimin'yaku	o	nonde	taoreteiru	tokoro	o
sleeping.pill	ACC	drank	be.collapsed	TOKORO (place/state/time)	ACC

kaizyoo	no	gyosen	ni	hakkensareta	no	da	to	iu
ocean	GEN	fishing.boat	by	be.discovered	NOM	COP	QT	say

'The state of him/her being collapsed on the edge of the cliff after taking sleeping pills was discovered by a fishing boat on the sea.'

The important information in (20) is carried by the underlined verb phrase 'being collapsed after taking sleeping pills', but not by the head noun *tokoro* 'place/state/time'. The underlined verb phrase is to *tokoro* 'place/state/time' in long noun phrases as the verb is to the connective particle in connectives. Thus, the noun phrase *Suimin'yaku o nonde taoreteiru tokoro* 'the state of being collapsed after taking sleeping pills' functions more as a connective than as a noun phrase. The way I understand Saeki's claim is that a longer noun phrase tends to be a nominalized noun phrase, and the nominalizer, which is analogous to a connective particle, projects a property of connectives (i.e. being placed at the beginning of a clause) over the long noun phrase, resulting in the more connective-like behavior of the long noun phrase.

Saeki posited communicative ease as a motivation for the earlier introduction of pronouns and demonstratives without providing further explanation for other tendencies. His 'communicative ease' can be interpreted in two ways. One is to ease the hearer's comprehension, and another is to reduce the burden on the speaker's production. These two issues are one of the central concerns of this thesis.

The study of word order by traditional Grammarians provides information on the syntactic, semantic, morphosyntactic, and discourse-pragmatic properties of initial constituents in a sentence. The traditional Japanese Grammarians' study is in a sense a pursuit of marked and unmarked constituent orders[15]. Such studies have value in that they provide the information on the environment where the unmarked word order occurs. No mention to the importance of study on the marked order, studying the properties of the unmarked order sheds light on the mapping of cognitive functions to linguistic realization. There should be a reason why the constituents in a clause with the unmarked word order are introduced in the way they are introduced, as well as the marked order.

1.2.4. Psycholinguistic experimental evidence

1.2.4.1. Long-before-short

Yamasita and Chang (2001) introduced experimental evidence that Japanese speakers prefer 'long before short' in production. In their experiments, the participants were asked to produce a sentence from a set of sentence components, including a verb, a subject and a direct object, through visual prompting. Two arguments were presented under three conditions: (a) both are short, (b) the subject is long, and the object is short, or (c) the direct object is long and the subject is short. Concrete examples are shown in Table (1).

---

[15]Dryer (1995) argues that if one word order is characterized as the one that is used elsewhere while another order is characterized with certain features, the order occurring elsewhere is pragmatically unmarked, and that the pragmatically unmarked order is not necessarily the most frequently used order in a language. In this paper, I will describe word orders in my data in terms of frequency, i.e. the default order or non-default order. When the terms, 'unmarked' or 'marked' are used in the citation, I will leave the author's terminology as it is.

Table (1) Phrase Length Condition in Yamashita and Chang's experiment (2001:B47)

(a) All-Short

keezi-ga	han'nin-o	oikaketa
Detective-nom	suspect-acc	chased.

'The detective chased the suspect.'

(b) Long-Subject

we-ga	takakute	gassiri-sita	keezi-ga	han'nin-o	oikaketa.
Height-nom	tall-and	big-boned	detective-nom	suspect-acc	chased.

'The detective who is tall and big-boned chased the suspect.'

(c) Long-Direct Object

keezi-ga	se-ga	takakute	gassiri-sita	han'nin-o	oikaketa
Detective-nom	height-nom	tall-and	big-boned	suspect-acc	chased.

'The detective chased the suspect who is tall and big-boned.'

Yamashita and Chang's study shows a clear favor of fronting long objects to the initial position. They further used the same experiment format and tested the effect of length in the ditransitive sentences containing a subject (S), a direct object (DO), an indirect object (IO), and a verb. Their interest is in the occurrences of 'IO-shift', 'DO-shift' and 'internal-shift' from the 'canonical order', S-IO-DO-V. The 'IO-shift' produces IO-S-DO-V order; the 'DO-shift' produces DO-S-IO-V order; and the 'internal-shift' procudes S-DO-IO-V order. The three arguments were presented under the three conditions: (a) all three are short, (b) the direct object is long and the other two are short, and (c) the indirect object is long and the other two are short. Participants were asked to produce ditransitive sentences through visual prompting. They report the results that the occurrence of IO-S-DO order is higher when the IO is long (i.e. under the condidtion (c)) than when the IO is short (i.e. under the condition (a) and (b)), and the occurrence of DO-S-IO is more frequent when the DO is longer (i.e. under the condition (b)) than when the

DO is short (i.e. under the condition (a) and (c)). They also report that the occurrence of S-DO-IO order is more frequent than DO-S-IO order when the DO is longer than when it is short. There is no explanation for the last result in Yamasita and Chang. There are a few possible motivations that I can propose. The subject-initial construction is dominant in the 3NPs construction involving the subject, the direct object, and the indirect object as well as in the 2NPs construction containing the subject and the direct object. Their use of *wa* for the subject marking in the ditransitive experiment also contribute to the occurrence of subject-initial ordering. Unfortunately in their experiment, the subject is always short, and there is no report on the effect of length in S-IO-DO-V clauses, which they treat as 'canonical order'. Thus, it is difficult to see the systematic pattern of the effect of length. Nevertheless, their experiment indicates a relation between the initial entity and syntactic heaviness.

Yamashita and Chang argue that the tendencies of 'short before long' in English and 'long before short' in Japanese are not contradictory[16] if they are working at two different levels during incremental production, the conceptual level and the lexical level. At the conceptual level, conceptually more accessible entities are more salient. At the lexical level, more semantically and pragmatically loaded entities are more informational,

---

[16] Dryer (1980) provides evidences for sentential noun phrases' positional tendencies for final over initial, final over internal, and initial over internal. Sentential noun phrases are "subordinate clauses which function as subject or object of their sentence" (Dryer:124). Thus, sentential noun phrases are by definition syntactically heavy entities. The following (i) and (ii) are from Dryer (1980:124). (i) contains a sentential subject and (ii) a sentential object. The sentential noun phrases are italicized.

(i) *That John is tall* is obvious.
(ii) Bill know *that John is tall*.

The sentential noun phrases' positional tendency for final over internal is relevant to short-before-long preference in English, and that for initial over internal is relevant to long-before-short preference in Japanese.

and thus more salient. They suggest that the salience at different levels competes for an earlier position, and which level wins which position will vary from language to language.

There are two contradictory word order tendencies in Japanese. One is the tendency that longer noun phrases come before shorter noun phrases and another is for demonstratives, which in general are short, to appear at the beginning of the sentence. Previous studies (e.g. Miyajima 1964, Saeki 1975, 1998) suggest that the reason for fronting heavier noun phrases is to facilitate the processing of the sentence structure (i.e. *Kakari-Uke* flow), and the occurrence of demonstratives at the clause-initial position has been discussed in terms of the referential dependency of a demonstrative on an entity in the anaphoric discourse. If Yamashita and Chang's suggestion is valid, syntactic weight and referentiality are both salient at two different levels of sentence production and compete for earlier positions.

1.2.4.2. Given-before-new

Ferreira and Yoshita (2003) conducted a long-term-recall-task experiment to explore the effect of given-new ordering in spoken Japanese and its implication as to two stages of processing in sentence production suggested in psycholinguistic literature.[17] The two processing stages include the first stage, function assignment, and the second stage, constituent integration. The first stage retrieves the lexical-syntactic forms (lemma) corresponding to entities in a conceptual representation, assigns the grammatical functions to the entities, and connects the lexical-syntactic forms (lemma) and the grammatical functions. The second stage determines the sequential positioning of entities

---

[17] E.g. Bock 1995, Bock & levelt 1994, Garrett 1975, Levelt 1989 in Ferreira and Yoshita's citation.

whose lexical-syntactic forms and grammatical forms are bound at the first stage. The premise behind these stages is that if the given-before-new preference exists in Japanese (i.e. if the givenness is a motivation for constituent ordering), given-before-new ordering is considered to arise at the constituent integration stage since constituent ordering in Japanese does not involve the grammatical role assignment. This is distinguished from English in which the word order signals the grammatical role assignment. Ferreira and Yoshita claim, then, that if we find the given-before-new effect in Japanese, we can say that we do not place an entity to the first position because of its givenness per se, but we do so because the information is more available. Speaking more available information first buys time for speakers to retrieve less available information or to form new information.

In Ferreira and Yoshita's experiment, sixteen Japanese speakers were first instructed to listen to twenty-four sets of eliciting sequences, and then 24 dative target ditransitive sentences. They were instructed to remember the target ditransitive sentences and to determine which eliciting sequence each target sentence belongs to. Then, the same eliciting sequences were presented again. This time, the subjects were instructed to produce a target ditransitive sentence after listening to each eliciting sequence. Each subject produced ninety-six dative target sentences in four sessions.

Table (2) Examples of eliciting sequences (a-b) and dative targets (c) in Ferreira and Yoshita's experiment.

---

(1) The same form reference

a. okusan-ga <u>otetudaisan-ni</u> kansyasiteita
housewife-NOM housekeeper-DAT was grateful
'The housewife was grateful to the housekeeper.'

b.	sorekara	doosita?		
	What happened next?			
	'What happened next?'			
c.	okusan-ga	otetudaisan-ni	purezento-o	okutta
	housewife-NOM	housekeeper-DAT	present-ACC	gave
	'The housewife gave the housekeeper a present.'			

(2) The different form reference

a.	okusan-ga	okurimono-o	katta	
	housewife-NOM	gift-ACC	bought	
	'The housewife bought a gift.'			
b.	sorekara	doosita?		
	What happened next?			
	'What happened next?'			
c.	okusan-ga	purezento-o	otetudaisan-ni	okutta
	housewife-NOM	present-ACC	housekeeper-DAT	gave
	'The housewife gave a present to the housekeeper.'			

The given argument in dative targets and its previous mention was manipulated in order to test the effect of the same form reference and the different form reference. For example, in (1) in Table (2), the given argument in the dative target (c), *otetudaisan* 'housekeeper', has the same lexical form as its previous mention, i.e. the corresponding argument in the eliciting sequence in (a), *otetudaisan* 'housekeeper', while in (2) in Table (2), the given argument in the dative target, *okurimono* 'gift', is lexically different from its corresponding argument in the eliciting sequence, *purezento* 'present'. The target sentences were also given in SDOV or SODV orders.

The result of the experiment indicated two factors for the production of dative target sentences that are relevant to this study. First, the subjects produced the sequence of subject-indirect object-direct object (SDOV) more than that of subject-direct object-indirect object (SODV). Ferreira and Yoshita term the former sequence as 'canonical' word order and the latter as scrambled word order.[18] Second, the subjects produced given-before-new ordering more often when the dative targets were presented with new-before-given ordering than when they were presented with given-before-new ordering, and importantly given-before-new ordering was produced equally as much in 'canonical' and scrambled word order sentences. The effect of this given-before-new preference was also present in both same or different lexical form conditions, but it was strongly observed when the given argument in dative targets and its previous mention have the same lexical forms.

Ferreira and Yoshita argue that Japanese speakers' given-before-new preference is less frequent under the condition of the different form reference because the information on the given argument in dative targets is only conceptually available whereas in case of the same form reference, the information is lexically and conceptually available, and thus "given-new ordering seems to emerge as a consequence of availability-sensitive processing at the level of constituent integration", but not to "communicate givenness per se" (2003:689).

Ferreira and Yoshita's work is significant for the present study in that it provides experimental evidence to suggest that givenness is a possible motivation for constituent ordering in spoken Japanese, and that constituent ordering, subject-indirect object-direct object, is the unmarked ordering for ditransitive construction in Japanese.

1.2.5. Corpus analysis approach

Yamashita (2002) proposes that scrambling in Japanese serves the purpose of facilitating the speaker's production rather than the hearer/reader's comprehension.

---

[18] I prefer calling the former sequence the default order and the latter a non-default order since the former is produced more frequently than the latter.

She looked at three types of "scrambled" sentences in terms of syntactic 'heaviness' and discourse 'referentiality'.[19] Three types of scrambling are (22) "short-distance" scrambling, (23) "internal" scrambling, and (24) "long-distance" scrambling. Yamashita's "short-distance scrambling" is fronting one of the arguments to the sentence-initial position; her "internal scrambling" is placing the direct object in front of the indirect object in a clause; and her "long-distance scrambling" is the topicalization of an argument by fronting it from the embedded clause to the sentence-initial position (i.e. the initial position of the matrix clause). Examples (22) to (24) illustrate each of these types. Example (21) is 'canonical' order. In (22) to (24), the 'moved entities'[20] are underlined. All examples are from Yamashita (2002:601-602), except (24a) which I supplemented to show the sentence before the topicalization.

(21) "Canonical" order of ditransitive clauses: 'John gave Mary an apple.'

John-ga	Mary-ni	ringo-o	ageta.
John-NOM	Mary-DAT	apple-ACC	gave

(22) Short-distance scrambling: 'John gave Mary an apple.'

a.	<u>Mary-ni</u>	John-ga	<u>e</u>	ringo-o	ageta.
	Mary-DAT	John-NOM		apple-ACC	gave
b.	<u>ringo-o</u>	John-ga	Mary-ni	<u>e</u>	ageta.
	apple-ACC	John-NOM	Mary-DAT		gave

[19] Yamashita's 'referentiality' includes reference by explicit mention and by inference in anaphoric discourse.

[20] The empty category 'e' in Examples (22)-(24) reflects Yamashita's assumption that *Mary-ni* (Mary-DAT) in (22a), *ringo-o* (apple-ACC) in (22b) and (23), and *ookina-ie-o* (big-house-ACC) in (24b) are all preposed from their 'canonical' positions. This thesis does not have such an assumption. For example, moving 'John' from initial position to position after 'Mary' is as true as 'Mary' moving after 'John' to initial position. We may rather say that scrambling two constituents 'John' and 'Mary' in (21) reverses the order of the two constituents and results in (22a) without one being moved and the other staying where it is.

(23) Internal scrambling: 'John gave Mary an apple.'

John-ga	ringo-o	Mary-ni	e	ageta.
John-NOM	apple-ACC	Mary-DAT		gave

(24) Long-distance scrambling: 'John thinks that Mary bought a big house.'

a.

John-wa	Mary-ga	ookina-ie-o	katta-to	omotte-iru.
John-TOP	Mary-NOM	big-house-ACC	bought-COM	think-PPG

b.

ookina-ie-o	John-wa	Mary-ga	e	katta-to	omotte-iru.
Big-house-ACC	John-TOP	Mary-NOM		bought-COM	think-PPG

'John thinks that Mary bought a big house.'

Yamashita's data shows that 73.7% of the preposed constituents have the property of syntactic heaviness and 26% of the proposed constituents make some reference to the previous utterances. Most importantly, she indicates that those with referentiality are syntactically light, and thus they are in almost complementary distribution with those syntactically heavy preposed constituents. In sum, 95 % of the preposed constituents in Yamashita's data have either the syntactic property of heaviness or the discourse property of referentiality. With respect to the heavy shift, she argues that the shift of heavy constituents either to the sentence-initial position or to the sentence-final position is based on production-based motivations, and reasons that the preposing or postposing of heavy constituents lightens the working memory of speakers. Yamashita also looked at her data in terms of the discourse principle given-before-new. The result was inconclusive; she found some preposed constituents were given, and some were new. Yamashita, following up previous research suggesting that sentences with a 'marked' order are used to cue a topic change, further examined preposed constituents to see whether they signal a change of topic. The function of signaling a change of topic is a comprehension-based motivation for the ordering. In her data, 89 % of sentences containing preposed constituents continued on the current topic. The remaining sentences containing preposed constituents did not function as the main signal of topic change. Given this result, Yamashita dismisses the signal of a topic change as a comprehension-based motivation for preposing. This result, however, indicates a higher percentage of preposed constituents are persistent in the following discourse. The accountability of preposed constituents by anaphoric and cataphoric referentiality, and syntactic weight in Yamashita's data is worth pursuing.

1.2.6. Discourse pragmatic approach on postposing

In the discourse pragmatic approach much of the work has been done for postposing, which involves postverbal elements, rather than for scrambling involving preverbal constituents. In this section, I will discuss two studies of postposing which have used a similar approach and data to those of this study.

Utilizing Givón's referential distance measurement, Fujii (1991) examined the eligibility for ellipsis of postverbal elements in postposing constructions in Japanese based on the concept of active, semi-active, and inactive information (Chafe 1987). She classified the postposing sentences into two groups: one with postverbal elements eligible for ellipsis and another with postverbal elements not eligible for ellipsis. She found two seemingly contradictory results. Among the postposing sentences with postposed elements not eligible for ellipsis, there are those where the postverbal elements are more focused than the preceding elements, and there are those where the preverbal elements are more focused than the postposed elements. Her interpretation of the former is that the postverbal elements should be in the canonical position, but are mistakenly dropped, and are thus postposed in order to compensate the mistakenly dropped information.[21] She

[21] It is 'recoverable' information in terms of Kuno (1978)

interprets the latter to be an example of 'important information first' in terms of Payne's "pragmatic markedness" (1985).[22] Fujii proposed that in the latter case the preceding elements are more focused simply because they contain characteristics of "pragmatic markedness". She identifies one type of construction, among the postposing constructions with the postverbal elements eligible for ellipsis, where the postverbal elements are highly topical, in terms of Topic Continuity, and thus highly activated. She analyzes the function of those postverbal elements as confirming their discourse status to the hearer when they do not have to be mentioned.[23] Although she recognized the pragmatic notions such as importance, topicality, or activation, her explanation for why the postposing construction is used instead of canonical order is not clear. Nevertheless, her findings that preverbal elements are more focused than the postposed entities, and that the postposed entities are more topic continuous and activated than the preverbal elements, are informative in that they indicate the ordering of focus (i.e. nonactive) < nonfocus (i.e. active) in a clause.

Shimojo (1995) approaches the postposing from a different angle than Fujii, though he also uses Givón's referential distance methodology. He investigated the postposing constructions in terms of activation suggested by Dryer (1996). He first examined the postposed construction in question-answer exchanges, and determined the acceptability of postposing elements in terms of Lambrecht focus structures. His analysis revealed that the postposing in the predicate-focus structures is more acceptable than that

---

[22] Payne (1985) conducted an empirical study and wrote a detailed analysis of the types of clauses containing preverbal nominals, and established the characteristics of preverbal elements: single focus contrast, double focus contrast, restatement, added detailed restatement, (wh-) questions, answer to (wh-) question, counter expectation, negation, threats, and unexplained as "pragmatic markedness".

[23] It is 'deducible' information in Shimojo (2005).

in sentence-focus structures, which in turn is more acceptable than that in argument-focus structures, and that acceptability can be accounted for by the relative degree of activation between the postverbal elements and the preceding elements. Based on this observation as well as further examination of the referential distance of postverbal elements and their preceding elements in the 140-minute conversational text, Shimojo proposed the following Acceptability Hierarchy in terms of activation. In the Hierarchy below, "A" is the proposition expressed by the preverbal elements, including the predicate, and "B" refers to the elements in the postverbal position.

Acceptability Hierarchy (Shimojo 1995:143)

Most acceptable	$A \geq B$[24]	B is more active than A, or A and B are equally active
|	$A = B$	A and B are equally nonactive
Least acceptable	$A < B$	A is more active than B

This Acceptability Hierarchy successfully rules out the least acceptable postposing construction. The most acceptable postposing construction is the case of $A \geq B$, and as Shimojo states, "Postposing construction manifests the focus structure of the sentence by its postverbal unit as opposed to the preceding unit; the postverbal unit represents nonfocus as its primary function[25], or it may represent focus secondary to the focus involved in the preceding unit" (Shimojo 1995:148). Shimojo's finding, as well as Fujii's, indicates the ordering of focus (i.e. nonactive) < nonfocus (i.e. active).

Shimojo's work also provides insight for this thesis in that he refers to the interplay between the postnominal particles *wa/ga* and postposing. His study examined

---

[24] $A \geq B$ means the RD of A is greater than or equal to B, so that A is not more activated than B.

[25] In recent study of argument endcoding, Shimojo (2005) suggests that postposing is best defined in terms of cataphoric nonpersistence than anaphoric activation. He calls postposing a "defocusing device" (2005:132) which defocusses previously activated referents.

the token distribution of *wa/ga*-marked postverbal elements in postposing constructions, and revealed that *wa*-marked postverbal elements and *ga*-marked postverbal elements are complementarily distributed in the Accessibility Hierarchy of postposing.[26] He likened postverbal elements to *wa*-marked elements in that both are nonfocus, and are to "serve as a basis for a new proposition by making the new proposition relevant at that point of discourse" (Shimojo 1995:148).

1.2.7. Word order in other SOV languages

Kim (1988) claims, based on the crosslinguistic examination of the syntactic position of the subject WH-words, that the focus position in rigid verb-final languages is immediately preverbal. His claim, then, predicts (25a), which is OSV order, and penalizes (25b), which exhibits the default order, SOV, in Japanese.

(25) 'Who will go to Osaka?' (Kim 1988:159)

a.	Osaka	wa	dare	ga	iku-ndai.
	Osaka	TOP	who	NOM	go-Q
b.	Dare	ga	Osaka	ni	iku-ndai.
	Who	NOM	Osaka	to	go-Q

He argues that the subject WH-word, *dare-ga* (who-NOM) in (25b) is not in the immediately preverbal position because *Osaka-ni* in (25b) forms a semantic unit with the verb (i.e. "phrasal topic" in Kim's terminology), and the verb-finality of Japanese prevents the "phrasal topic" from being preposed. In other words, semantics and the verb-

[26] Shimojo's new data shows that among *wa*, *ga* and zero-particle marking on the postverbal elements, the zero-particle is the most frequent, *wa* is the second most frequent, and *ga* the least frequent (2005:202). He analyzes that the zero-particle is associated with the cataphoric defocusing, and *wa* and *ga* with cataphoric focusing. See Shimojo (2005) for further development on argument encoding in Japanese.

finality do not allow *dare-ga* to occur between *Osaka-ni* and the verb.[27] An alternative explanation would be a different contextual reading that Kim himself pointed out. Kim states that (25a) and (25b) occur in different contexts. For example, in the context such as "Suppose a company in Tokyo is considering establishing a new branch in Osaka and accordingly selecting personnel for the new post", " a company staff member's concern may be expressed most suitably by" (25a), and in the context such as "If the company has several branches in major cities and the branch managers are routinely rotates", (25b) would be more appropriate (Kim 1988:159). In short, the difference in constituent ordering between (25a) and (25b) is factored by a different semantic reading.

Herring and Paolillo (1995) present a study on presentational focus and WH-focus in Tamil and Sinhala.[28] Their presentational focus correlates with the "newness of mention"; thus, the distribution of new mentions portrays the position of presentational focus in their term. Their analysis indicates that among the constituents in the initial position, 5.3 % are new mentions; among preverbal constituents, 21.4% are new mentions; and among the post-verbal constituents that are not afterthoughts and emphatics, 90.9% are new mentions. Based on the ratio of new mentions, they claim that the post-verbal position is specialized for the presentational focus in Sinhala. Their analysis of new mentions in Tamil shows the same tendency that the sentence-final position is for the presentational focus, except that in Tamil, the sentence-final position

---

[27] Kim states that the topicalization of "phrasal topic" is only realized by pseudo-cleft construction as in (a).

(a)	Osaka	ni	iku	no	wa	dare-dai	(Kim 1988:159)
	Osaka	to	go	NMLZ	TOP	who-Q	

'Who is the person assigned to go to Osaka?'

[28] Tamil and Sinhala are SOV languages with flexible verb finality and with postpositional marking. Tamil is a Dravidian language spoken in India and Sri Lanka, and Sinhala is an Indo-European language spoken in Sri Lanka.

involves the verbless construction instead of post-verbal constituents. With respect to WH-focus, Herring and Paolillo looked at the position of WH-question words. Their close examination of WH-words shows that information-seeking WH-words more often occur in initial position in Sinhala.[29] Their data analysis in Tamil was only conclusive for information-seeking WH-adverbs (i.e. 'where', when', and 'how') that prefer the preverbal position.[30]

Testelec (1998) introduces word order variation in Archi, a Daghestanian language.[31] Example (26a) is SOV, (26b) is OVS, and (26d) is OSV. According to Testelec, in Archi the preverbal position is normally reserved for the focused NP. In SOV, as in (26a), however, being the unmarked order, it is not the case that the preverbal object is a focal NP. In order to put the object in focus, the OVS order, as in (26b), is used. For the subject focus, OSV, as in (26d), is used instead of SVO order, as in (26c), which is inappropriate for the subject focus reading.

(26) a. Boxlotu-mu xams a-b-c'u.
hunter-ERG bear kill<3CL>.AOR
'The hunter killed a bear.'

b. xams a-b-c'u Boxlotu-mu.
bear kill<3CL>.AOR hunter-ERG
'The hunter killed a BEAR.'

---

[29] Herring and Paolillo (1995) examined WH-words in three categories, classical rhetorical questions, thematicizing rhetorical questions (i.e. WH-question which is followed by the answer), and information-seeking WH-questions.

[30] Their data show that in Tamil the positional preference of subject WH-words is the initial position and that of object WH-words is preverbal, which coincide with subject and object positions in the 'unmarked' word order in Tamil. Because of the possible grammatical conditioning, they suggest the inconclusive result for WH-words. WH-adverbs, however, are not in this case since "adverbs are not assigned a fixed position in Tamil grammar" (Herring and Paolillo, 1995:189).

[31] The Daghestanian languages, which are spoken in the Caucasian region, have SOV order. Unlike Japanese, they are ergative languages, and manifest agreement systems in class and number in verbs and adjectives.

c.# Boxlotu-mu a-b-c'u xams.
hunter-ERG kill<3CL>.AOR bear

d. xams Boxlotu-mu a-b-c'u.
bear hunter-ERG kill<3CL>.AOR
'The HUNTER killed a bear.'

Vilkuna (1998) argues that in Tundra Nenets constituents carrying new information are placed immediately before the verb.[32] Example (27a) is SOV, and Example (27b) is OSV. Thus, in (27a) the subject 'the woman' is old information, and the object 'the man' is new information, whereas in (27b) the object 'the man' is old information, and the subject 'the woman' is new information. Vilkuna's analysis suggests that in Tundra Nenets the constituent ordering is given-before-new.

(27) a. Nye xasawam ladø^0
woman man:ACC hit:SUBJ-3SG
(What happened?) 'The woman hit the man.'

b. Xasawam nye ladø^{0}da
man:ACC woman hit:OBJ-SG.SUBJ2-SG
'(It was) the woman (who) hit the man.'

1.2.8. Summary

The attempt to explain the word order in terms of thematic roles failed when the correlation between the grammatical relation and the semantic roles of constituents is preserved. Iijima, based on Communicative Dynamism and Functional Sentence Perspective in the Prague school tradition, claims that the information flow in Japanese is 'topic' (= 'theme') to 'focus' (= 'rheme'), and that the entity in the immediate preverbal position is 'focus'. Iijima's claim is supported by studies of other SOV languages

[32] Tundra Nenets is an Uralic language spoken in Northern Russia. It is a rigid verb-final language. Unlike Japanese, Tundra Nenets has subject agreement in number and person, and finite subordinate clauses are practically nonexistent.

suggesting the preverbal position for focal entities. Nevertheless, the direction of informational flow from 'topic' to 'focus' is still debatable among linguists. Fujii and Shimojo's work, although their focus on postposing is beyond the scope of this thesis, demonstrated that the notions of 'importance', 'topicality' and 'activation' are relevant to the discourse and psychological status of a linguistic entity. Traditional Japanese grammarians provided ample data on the default and nondefault constituent orders in Japanese. The *Kakari-Uke* schema (i.e. argument-predicate linkage) was measured in terms of distance between an argument and its predicate, as well as in terms of semantic tightness between them. The distance and semantic tightness of argument-predicate linkage are the two parameters used to account for the constituent ordering in Japanese grammar. Finally, data from psycholinguistic experiments showed the Japanese speaker's preference, long-before-short, and given-before-new.

1.3. Possible motivations underlying word order variation

In this section, I will discuss three frameworks postulating possible factors that underlie constituent ordering. First, the Givónian tradition of discourse measurements is introduced. Givón's (1983, 1988) discourse measurements of referential distance, Possible Interference, and topic persistence measure the properties of referents, such as 'unpredictability', 'accessibility', and 'importance'. Second, the notion of 'activation' proposed by Dryer (1996) is presented. Recent cross-disciplinary study in experimental psychology and cognitive linguistics has suggested that the 'activation' status of referents plays an important role in ordering constituents. Third, Hawkins' parsing theory is discussed. Hawkins (1992, 1994, 2004) argues that syntactic weight is a ruling factor for fronting noun phrases in left-branching languages.

The attempt to integrate discourse, cognitive and syntactic factors for constituent ordering gave rise to a debate over the production-based vs. comprehension-based motivations. I will briefly touch the issue. In addition, I will discuss possible relations of word order with postnominal particles and with reference forms.

1.3.1. Givónian tradition

The Givónian tradition is an implementation of a quantitative methodology by which certain discourse properties can be measured. Givón (1983) introduced three discourse measurements: referential distance, topic persistence and potential interference. Referential distance (RD) is the distance from the most recent mention of an entity in anaphoric discourse. In Givón, it measures the discourse property of 'predictability' of a referent to the hearer --whether the hearer can identify the referent that the speaker refers to. The more predictable the referent is, the easier it is for the hearer to identify. Givón's topic persistence (TP) measures how long a referent remains in the following discourse, and Givón correlates the measurement to the discourse property of 'importance'. The referent with a higher value of TP is translated as a more informationally important entity. Givón's 'importance' is interpreted as the speaker's assessment of an entity with regard to its behavior in the following discourse.

Crosslinguistic text-based studies of preverbal and postverbal constituents using Givón's measurements indicate the correlation between less predictable referent and preverbal constituents, and between more important referent and preverbal constituents. Based on this finding, Givón proposes two parameters, 'unpredictable' before 'predictable' and 'important' before 'unimportant'. A more 'unpredictable' referent and more 'important' information are introduced earlier to facilitate the hearer's overall comprehension. Givón views, as other functionalists, syntactic structures as the realization of cognitive functions, and proposes that the motivation behind the factors ‘predictability’ and ‘importance’ is the “communicative task urgency” as he states “A communicative task is more urgent when the information to be communicated is either less predictable or more important” (Givón 1988: 275). He further argues that the reason behind the urgency is ‘attention’. Entities in the initial position of a string attract the hearer’s attention, and the attended information can be stored and retrieved more efficiently. He indicates that ‘task importance’ and ‘informational predictability’ are the most important pragmatic parameters grounded in more general human cognition, and the two parameters might coincide, or may be independent of each other, in predicting a specific constituent order, Where the two parameters conflict, Givón observes, ‘importance’ overrides ‘predictability’. In Givón (1993), RD is used as an index of activation. For example, zero anaphora and unstressed PRO in English, which has a mean RD of 1, are devices to signal the continued activation, and other devices such as stressed PRO, Y-movement, definite nouns, which has a mean RD of greater than 1, are devices to signal the de-activation of the currently active referent. If RD indicates activation, we can see TP as an indicator of the decay of activation for the currently active referents.

Another discourse measurement, Potential Interference (PI) concerns the presence of semantically compatible (‘most commonly in terms of animacy, humanity, agentivity or semantic plausibility as object or subject’: Givón 1983) referents within the preceding three clauses, which possibly causes difficulty in a hearer’s identifying the referent of an entity in question. Thus, the entity with ‘potential interference’ is introduced earlier in a clause to avoid the confusion in the hearer’s mind. Based on previous research, Givón assumes that semantically compatible referents beyond the preceding three clauses do not cause much interference in a hearer's ability to identify the referent that the speaker refers to. In Givón, the two anaphoric measurements, RD and PI, are understood as a comprehension-based factor.

1.3.2. Dryer's activation theory

1.3.2.1. Activation theory

Recent crossdisciplinary studies in experimental psychology and cognitive linguistics center around the notion of 'attention'. The concept of 'attention' in cognitive processes parallels the linguistic concept of 'activation'. Many functional linguists agree on the view that language is an integrated part of human cognition, and thus linguistic processes are instantiations of more general cognitive processes. In other words, linguistic structures are within the scope of cognitive structures.

Dryer's activation theory is one framework that attempts to map certain linguistic forms to certain cognitive processes. He argues that activation is the status of cognitive entities, and its relevance to linguistics "lies in the fact that certain linguistic choices (like pronouns vs. noun, or active vs. passive, or position of focal accent) may be systematically related to if not determined by, the activation status of cognitive entities", and "in some cases may be the activation status of entities in the mind of the speaker that is relevant, while in other cases it may be the speaker's assumptions about the activation status of corresponding cognitive entities in the mind of the hearer" (1996:482). The status of 'activation' of an entity is measured by examining recency, frequency, or inferability of its previous mention in anaphoric discourse. The more often an entity is mentioned in the previous discourse, the more prominently the referent of the entity is registered in the speaker's and the hearer's mind, and thus more activated. The more recent the previous mention is, the fresher its referent is in the speaker's and the hearer's memory, and thus more activated.

One of the linguistic issues central to this thesis is a speaker's choice of one particular word order over another. Following Dryer's view, the speaker's choice is related to the activation status of cognitive entities that correspond to linguistic entities. In what follows, I will sketch the theory of activation proposed by Dryer (1996).

Dryer (1996) argues that nonfocus in simple focus sentences involves activation. That is, nonfocus corresponds to an activated entity.

(28) A: Who saw John? (Dryer 1996)
B: MARY saw John.

Nonfocus in the simple focus sentence (28B) is *saw John*. That part of the utterance has no focal accent and the proposition 'X saw John' is more activated, being mentioned in the immediately preceding context (28A), than MARY is. MARY in (28B) is not activated because its corresponding cognitive entity has not been mentioned before in the above exchange. Dryer examined single focus sentences, negated clauses, questions, and conditional clauses, and demonstrated how successfully the degree of activation predicts nonfocus elements.

Activation is not a binary value of active and nonactive, but rather a continuous notion. The activation status of entities is dynamic; it changes through time. The first mention of an entity causes it to become activated in the mind of the hearer. Activated entities may become deactivated and eventually fade away, or may remain activated. No mention of an entity causes its gradual decay. Since activation is a continuum, somewhere in the middle of the continuum there are entities that are less activated than other entities, but still more activated than entities that are not activated or decayed. Dryer labeled those intermediate entities ‘semi-deactivated’ entities. The notion that activation is a continuum becomes important in predicting focus and nonfocus distribution in the exchange, as in (29).

(29) A: Did MARY kiss john or did SALLY kiss him? (Dryer 1996:496)
B: MARY kissed John.

The propositions that someone kissed John (X kissed John), that Mary kissed John (Mary kissed John), and that Mary kissed someone (Mary kissed X) are all activated by the first part of the question in (29A). In addition, the propositions that someone kissed John (X kissed John), that Sally kissed John (Sally kissed John), and that Sally kissed someone (Sally kissed X) are all activated by the second part of (29A). At the time B utters (29B), these six propositions are all activated in B’s mind. However, the proposition that someone kissed John (X kissed John) is more highly activated than the other five propositions, occurring in both clauses in (29A).

Dryer identifies another type of entity in intermediate status, called ‘accessible’ entities. Accessible entities occur where the activation of one entity causes another entity related to the activated entity to be accessible to activation by inference or association. Accessibility is also a continuous notion. Among fully activated entities, some entities may be more highly activated than other entities and become what Dryer calls ‘focus of attention’, which is based on “the cognitive notion of focal attention” (1996:). Focal attention is a cognitive notion that is well grounded in the psychology literature (Tomlin 1995:518). The following diagram (30) summarizes the four activation levels in the continuum that Dryer discusses. The symbol > denotes ‘more activated’.

(30) focus of attention >
activated but not focus of attention >
recently activated but now semi-deactivated/accessible to activation >
nonactivated

One important thing to remember is that the activation status of cognitive entities corresponds not only to noun phrases in an utterance but also to the proposition represented by the utterance.[33] When speaker A says *Who saw John*?, the proposition that someone saw John (X saw John) must be activated in the speaker's mind. At the same time, the same proposition is activated in the hearer's mind at the moment of utterance. When speaker B says *MARY saw John* as a reply to A, the cognitive entity corresponding to the noun phrase *MARY* is activated, as well as the proposition that Mary saw John.

1.3.2.2. Predicting narrow focus in Japanese

We have seen that activation can be a useful tool for predicting the distribution of focus and nonfocus. That is, by examining the activation status of cognitive entities corresponding to linguistic entities in an utterance, we can identify the distribution of the active and nonactive, and thus the distribution of focus and nonfocus in the utterance.

Let us practice using two sentences that involve scrambling in Japanese. Sentences (1) and (2), used earlier, are repeated here as (31) and (32) respectively. (31a), (31b), and (31c) are propositions of the sentence (31a), and (32a), (32b), and (32c), and (32d) are propositions of the sentence (32a).

(31) a. Taro ga Hanako o mita.
b. (Taro saw Hanako)
c. (X saw Hanako)
d. (Taro saw X)

---

[33] A similar concept can be seen in Givón (1983). He states that reduntant semantic information, which facilitates the hearer's referent-identification, "comes primarily from the predicate of the clause".

(32) a. Hanako o Taro ga mita.
b. (Taro saw Hanako)
c. (X saw Hanako)
d. (Taro saw X)

If *Taro ga* in (31a) is the focus in a particular discourse context, and *Hanako o mita* is therefore nonfocus, then Dryer's theory claims that the proposition that X saw Hanako should be more activated than the cognitive entity corresponding to *Taro*. By the same token, if *Hanako o* in (32a)[34] is focus, and thus *Taro ga mita* is nonfocus, his theory claims that the proposition Taro saw X is more activated than the cognitive entity corresponding to *Hanako*. Given the theoretical background, we can say it is feasible to examine the word order variation such as (31a) and (32a) in terms of activation.

1.3.2.3. Pragmatic presupposition

Dryer (1996) distinguishes two kinds of pragmatic presupposition, basic presuppositions and metapresuppositions. Metapresuppositions connote basic presuppositions, but the reverse is not valid. Dryer states, "In general, we can say that when an utterance pragmatically presupposes (in the basic sense) a proposition *p*, it also metapresupposes the proposition that the hearer believes the proposition *p*." (1996:502). Basic presuppositions in Japanese are triggered by the use of the adverbial particle *mo* that conveys the meaning of addition.[35] For example, the sentences in (34) and (35) presuppose a proposition (Someone went to the park) in the speaker's mind. On the other

---

[34] There is a native speaker's testimony that when a speaker utters (32a) instead of (31a), the primary focal accent falls on *o,* and the secondary focal accent on *ga*, as formulated in *Hanako WÓ Taro gà mita.*

[35] In traditional Japanese grammar the function of the particle *mo* contrasts with that of the particle *wa*. The function of the particle *mo* is defined in terms of 'addition', as opposed to 'elimination/separation' which is construed as the function of the particle *wa*. While *wa* picks up one entity in a group of entities by eliminating others, *mo* adds an entity to the group. Some studies on Japanese particles (e.g. Aoyagi 1998) treat both *wa* and *mo* as a focus inducer.

hand, when the speaker utters the sentence in (33), there is no such presupposition involved.

(33) watasi kooen ni ikimasita
1sg park LOC go.POL.PST
'I went to the park.'

(34) watasi **mo** kooen ni ikimasita
1sg also park LOC go.POL.PST
'I went to the park.'

(35) kooen ni watasi **mo** ikimasita
part LOC 1sg also go.POL.PST
'I went to the park.'

The use of the particle *mo* indicates that an entity parallel to the *mo*-marked entity in terms of thematic and grammatical roles, as well as an open proposition with the *mo*-marked entity as a variable is present in the speaker's mind at the time of utterance.[36] In (33), the open proposition involving the subject as variable is (X went to the park). When X is marked by particle *mo*, as in (34) and (35), the proposition (Someone other than X went to the park) is presupposed in the speaker's mind at the time of utterance. The utterance in (34) and (35), then, involves the metapresupposition that the hearer believes the proposition (Someone went to the park). Therefore, in the context where both the speaker and the hearer believe the proposition (someone went to the park), the speaker introduces a new entity X in the position of (someone). Is the new entity X not focus of the sentence? Can we call it presuppositional focus?

Dryer argues that both activation and metapresuppositions "involve beliefs by the speaker about the hearer's cognitive state." (1996:501), but that the important notion

[36] There is a case where the implication of a proposition in the previous discourse is the presupposition of *mo*-marked entity.

insofar as what plays an influential role on the position of the focal accent (in English) is what the speaker believes is activated in the hearer's mind (activation), but not what the speaker believes is in the hearer's beliefs (metapresuppositions)'.

We discussed above that pragmatic presupposition is signaled by the use of a particular particle in Japanese. In the preceding section, it was demonstrated that narrow focus in Japanese could be predicted by the activation status of constituents in Japanese. If we examine the positional preferences of *mo*-marked constituents and constituents in narrow focus, we can compare the role of activation and metapresupposition, if any, on influencing constituent ordering in Japanese. If there is any positional preference of *mo*-marked constituents, metapresuppositions may be argued to be influential for the ordering choice. If we find a strong positional preference of constituents in narrow focus, activation may be considered an effective factor of constituent ordering.

1.3.3. Hawkins' performance theory

1.3.3.1. Syntactic weight

Hawkins (1992, 1994, 2004) claims that syntactic weight is a primary motivation for word order variation, and that pragmatic factors such as predictability or importance are secondary.[37] Hawkins' claim is based on his processing theory. The underlying notion of his processing theory is the 'Constituent Recognition Domain' (CRD). The CRD is the number of words that we have to parse in order to recognize syntactic groupings. The words that Hawkins counts are all 'Immediate Constituents' (ICs) of a phrasal mother

[37] Systematic heavy noun phrase shifts either to clause-initial or to clause-final throughout languages was originally noted by Dryer (1980) in terms of sentential noun phrases' positional preference.

node (M). The core principle of Hawkins' claim is that the shorter CRD makes the processing faster, therefore easier to parse.

The example in (36) illustrates Heavy NP Shift in English (Hawkins 1992:199). Both (36a) and (36b) have three ICs in VP (verb phrase). They are a V (verb), a PP (prepositional phrase) and a NP (noun phrase). The length of the CRD for (36a) is eleven words. It means that the hearer of (36a) has to go through eleven words until she recognizes the sentence structure of (36a). When the hearer hears 'to' in the PP [to Mary] in (36a), she recognizes the sentence structure Subject-Verb-Direct Object-Indirect Object. The length of the CRD for (36b) is four because the hearer goes through four words until she recognizes the sentence structure of (36b), which is Subject-Verb-Indirect Object-Direct Object. Hawkins claims that Sentence (36b) is easier to parse since the CRD is shorter.

(36)

a) I $_{VP}$[introduced $_{NP}$[some friends that John had brought to the party] $_{PP}$[to Mary]]
1 2 3 4 5 6 7 8 9 10 11

b) I $_{VP}$[introduced $_{PP}$[to Mary] $_{NP}$[some friends that John had brought to the party]]
1 2 3 4

Moreover, Hawkins formulated Early Immediate Constituents (EIC) in order to substantiate his claim. Hawkins states, "The human parser prefers to maximize the left-to-right IC-to-word ratios of the phrasal nodes that it constructs" (1992:200). The EIC is the IC-to-word ratio. The IC-to-word ratio is calculated by dividing the IC total by the word total at a point. Let us illustrate his calculation in the same Heavy NP Shift example.

(37)

a) I $_{VP}$[introduced $_{NP}$[some friends that John had brought to the party] $_{PP}$[to Mary]]

1/1 2/2 2/3 2/4 2/5 2/6 2/7 2/8 2/9 2/10 3/11

100% 100% 67% 30% 40% 33% 29% 25% 22% 20% 27%

= 47% aggregate ratio

b) I $_{VP}$[introduced $_{PP}$[to Mary] $_{NP}$[some friends that John had brought to the party]]

1/1 2/2 2/3 3/4

100% 100% 67% 75%

= 86% aggregate ratio

The IC total in (37a) is 1 for the Verb 'introduced', 2 for the NP, and 3 for the PP. The word total at the point of 'John', for example, is 5. Therefore, the IC-to-word ratio at the point of the constituent 'John' is 2 divided by 5 = 40%. In order for a hearer to recognize the sentence structure of (37a), the hearer has to hear 'to' in the PP. When the hearer hears 'to' in the PP, the immediate constituents of the clause have been recognized. Only the IC-to-word ratio of the constituents up to 'to' is necessary to calculate the average IC-to-word ratio for (37a). The average ratio is expressed as an aggregate ratio by Hawkins. The aggregate IC-to-word ratio of (37a) is 47%. For the sentence (37b), a hearer has to hear 'some' in the NP in order to recognize the sentence structure of (37b). Therefore, the IC-to-word ratio of constituents up to 'some' in the NP is included to calculate the aggregate ratio of IC-to-word ratio for (37b). The aggregate ratio of (37b) is 86%.

The constituency information becomes more informative as the ratio increases. The higher the aggregated IC-to-word ratio is, the more optimal it is for processing. In the above Heavy NP Shift example in English, the word order of Indirect Object-Direct Object of (36b) is more optimal for processing than the order of Direct Object-Indirect Object of (36a).

Hawkins (1994) argues that the principle of the shorter CRD making the processing faster is valid for a left-branching (or head-final language) like Japanese. According to Hawkins, left-branching languages that have right-peripheral complementizers will shorten the CRDs by preposing the sentential noun phrase marked by the complementizer to the left. It means that it is easier to process a sentence when a long nominalized clause appears at the beginning of the sentence than when it appears after other constituents. He explains how the CRD is shortened by preposing the center-embedded clause. Examples (38) and (39) are from Hawkins (1994:8, 1994:66). I underline the center-embedded clause in (38).

(38)	Mary-ga	kinoo	John-ga	kekkonsi-ta	to	it-ta
	Mary	yesterday	John	married	that	said
	'Mary said that John got married yesterday.'					

(39) a. $_{S1}$[$_{NP}$[Mary-ga] $_{VP}$[$_{S'}$[$_{S2}$[kinoo john-ga kekkonsi-ta] to] it-ta]]

b. $_{S2}$[$_{S'}$[$_{S1}$[kinoo john-ga kekkonsi-ta] to] $_{NP}$[Mary-ga] $_{VP}$it-ta]]

In (39a) the parser starts constructing the S when the *ga* is reached. In (39b) it starts constructing the S when *to* 'that' is reached. In both cases, the parser completes constructing the S when it reaches *itta* 'said'. Therefore, the CRD for (39b) is much shorter than for (39a). Hawkins compared IC-to-word ratios for four logical possible constituent structures[38] of the construction [NP1-NP2-V] and their reverse order versions in two hypothetical conditions, (1) where NP1 is shorter than NP2 by two ICs, and (2) where NP1 is shorter than NP2 by four ICs. Four variations are shown in (40).

---

[38] Hawkins discusses further which constituent structure is optimal in terms of a better EIC ratio. Since it is not the concern to this thesis, I will not introduce his discussion here. The important point is that he claims that long-before-short for preverbal elements involves an optimal EIC ratio.

(40) a. $_{S}$[NP1 NP2 $_{VP}$[V]] versus $_{S}$[NP2 NP1 $_{VP}$[V]]
b. $_{VP}$[NP1 NP2 V] versus $_{VP}$[NP2 NP1 V]
c. $_{S}$[NP1 $_{VP}$[NP2 V]] versus $_{S}$[NP2 $_{VP}$[NP1 V]]
d. $_{S}$[$_{VP}$[NP1] NP2 [V]] versus $_{S}$[$_{VP}$[NP2] NP1 [V]]

His calculation indicated that IC-to-word ratios was higher for all four variations when the longer NP was placed before the shorter NP, and IC-to-word ratios improved as the difference in IC increased.

Based on the EIC, Hawkins claims that when the difference of word counts in two constituents is greater, the heavier constituent tends to come later in a clause in right-branching languages like English, and come earlier in a clause in left-branching languages like Japanese. Hawkins' claim predicts the tendency of fronting heavy noun phrases in Japanese written texts that was observed by Japanese traditional grammarians.

Hawkins points out that it is necessary to assume that a certain ordering in languages has been grammaticalized. For example, he argues that the ordering of a noun phrase (NP) before a prepositional phrase (PP) in English, such as (36a), has been grammaticalized because an NP is typically shorter than a PP in English, and the ordering of NP-PP in English has a better EIC ratio than the reverse ordering.[39] For Japanese, Hawkins argues that the grammaticalization of [NP1-ga NP2-o] can explain the 45 % tolerance for short-before-long order in [NP1 NP2 V] constructions, where NP is marked *ga* (nominative), *o* (accusative), or *ni* (dative), and when the difference in IC between NP1 and NP2 is one to two.

---

[39] Hawkins also argues that the ordering of a subject noun phrase before the verb phrase in German has been grammaticalized because an NP is typically shorter than a VP in German. And so the orderings of NP-VP in German provides better EIC ratios than reverse orderings.

1.3.3.2. Syntactic weight and pragmatic factor

Hawkins' claim that the shorter constituent tends to come earlier in a clause contradicts Givón's apparent prediction that the shorter constituents tend to be postposed. Givón claims that referents of constituents with shorter length are more accessible and more predictable, and that such constituents are likely to be placed later in the clause. Similarly, Givón's discourse principle of 'unpredictable' entity before 'predictable' entity predicts postposing of shorter length constituents. It is important to note that Givón's prediction is based on the data involving preposing (i.e. right-dislocation) and postposing (i.e. left-dislocation), as SV or OV versus VS or VO, while what Hawkins tries to predict is the ordering of scrambled elements such as preposing heavier NPs in Japanese and the dative shift in English.

Having said that, we continue with Hawkins' argument against Givón's claim. Hawkins examined his interpretation of Givón's predictions in the construction [NP V [PP_1 PP_2}] in English, such as in *The raven slept on a perch behind the back door* versus *The raven slept behind the back door on a perch*. The two locative phrases, *on a perch* (L1) and *behind the back door* (L2), can be in the order of L1 < L2 or L2 < L1. The following (41) summarizes Hawkins' interpretation of Givón's prediction by Task Urgency (henceforth abbreviated as TU) which involves new-before-given and important-before-unimportant principles (1994:218). 'P' stands for the value of 'unpredictability' and 'I' stands for that of 'importance'. The subscription $_1$ indicates the constituent that comes earlier, and the subscription $_2$ indicates the constituent that comes later in the sequence of AB or BA. Thus, $P_1 > P_2$ means that P_1 (the constituent introduced earlier in the sentence) is more unpredictable than P_2 (the constituent introduced later in the sentence). Similarly, $I_1 > I_2$

indicates that I_1 (the constituent introduced earlier in the sentence) is more important than I_2 (the sentence introduced later in the sentence).

(41)	a. Single Orders Predicted: AB	either	$P_1 > P_2$ (where $I_1 = I_2$)
		or	$I_1 > I_2$ (where $P_1 = P_2$)
		or	$P_1 > P_2$ & $I_1 > I_2$
	b. Both Orders Predicted AB or BA	either	$P_1 = P_2$ & $I_1 = I_2$ (all values can be Ø)
		or	$P_1 > P_2$ & $I_1 < I_2$
		or	$I_1 > I_2$ & $P_1 < P_2$
	c. Counterexamples: *BA	either	$P_1 < P_2$ (where $I_1 = I_2$)
		or	$I_1 < I_2$ (where $P_1 = P_2$)
		or	$P_1 < P_2$ & $I_1 < I_2$

The case of (41a) is the correct prediction when either RD or TP make predictions, or when both RD and TP make predictions. (41b) is the case when neither RD nor TP makes a prediction, or when RD and TP conflict (i.e. RD predicts AB whereas TP predicts BA, or vice versa). The case of (41c) occurs when both RD and TP make incorrect predictions, or when one makes an incorrect prediction while the other makes no prediction. Hawkins' data indicated that 13/62 was the ratio for (41a), 44/60 for (41b), and 12/62 for (41c). In other words, TU made the correct predictions for 21% of the data, predicted both orders with 71%, and made incorrect predictions for 19.4% of the data. The same data showed that the syntactic weight made the correct predictions (i.e. in case of English, the heavier constituent comes after the shorter constituent) in 38/69 (i.e. 55.1%), made no prediction in 21/69 (i.e. 30.4%), and made incorrect predictions in 10/69 (i.e. 14.5%). Hawkins' analysis indicated a higher success rate in prediction made by the syntactic weight. Hawkins observed an inverse correlation between EIC and TU, in that (41a) and (41b) fall into the case where the EIC does not predict the constituent order. Givón's incorrect predictions in (41c), on the other hand, represent a case where the EIC strongly predicts the ordering.

Hawkins, then, focused on one pragmatic factor, (un)predictability in Givón's terms, and examined the interaction between EIC and Givón's new-before-given, and between EIC and the Prague school's given-before-new in English, Hungarian, German and Japanese. He found a positive correlation between EIC and given-before-new in English, Hungarian, and German, where the category node C in the constituent structure is recognized on its left periphery (such as complementizer *that* in English), and between EIC and new-before-given in Japanese where the C is recognized on its right periphery (such as complementizer *to* in (38)). By correlation he means that new-before given is correctly predicted whenever EIC is correctly predicted in English, Hungarian and German, and given-before-new is correctly predicted whenever EIC is correctly predicted in Japanese. Based on the general tendency that new information is more linguistically loaded, and predictable entities are more likely pronouns or simple noun phrases, he proposes a solution for two contradictory correlations in two different language types: "when EIC predicts short before long ICs, there is a positive correlation with given-before-new; when it predicts long before short, there is a positive correlation with new-before-given." (1994:238). With respect to which one, EIC (or syntactic weight) or pragmatic factor (either new-before-given or given-before-new), is the primary factor for the ordering choice, he argues that EIC is because the pragmatic factor cannot explain the language-specific correlation between the syntactic weight and pragmatic factors but EIC can.

Hawkins' raw data in English and Japanese are consulted in order to compare the three ordering factors, syntactic weight, given-before-new, and new-before-given. The Japanese data are shown in Table (3), and English in Table (4).

Table (3) Hawkin's analysis of the interaction of syntactic weight and pragmatic factor

Japanese	Correct		No Prediction		Incorrect		Total	
[{PP1 PP2] V]	#	%	#	%	#	%	#	%
Syntactic Weight	19	39.6	22	45.8	7	14.6	48	100
Given-before-New	15	31.3	22	45.8	11	22.9	48	100
New-before-Given	11	22.9	22	45.8	15	31.3	48	100

Table (4) Hawkin's analysis of the interaction of syntactic weight and pragmatic factor

English	Correct		No Prediction		Incorrect		Total	
[NP V [PP1 PP2}]	#	%	#	%	#	%	#	%
Syntactic Weight	38	55.1	21	30.4	10	14.5	69	100
Given-before-New	28	40.6	21	30.4	20	29.0	69	100
New-before-Given	20	29.0	21	30.4	28	40.6	69	100

As Hawkins claimed for EIC, the success rate (ratios of correct to incorrect predictions) of the predictions by the syntactic weight is the highest among the three factors. Tables (3) and (4) show that both in English and Japanese, the heavier and given entity comes earlier.

1.3.4. Production-based motivations

As discussed in the above section, syntactic weight has been presented as one of the governing factors of constituent ordering, and researchers, as represented by Hawkins, have explained the choice of a particular constituent ordering within the framework of parsing theory. The concept of efficient parsing, thus, is concerned with the hearer's comprehension. Researchers who have focused on the discourse status of constituents, as represented by Givón, have tried to map the discourse status of constituents within a broader picture of cognitive functions such as 'task urgency', and viewed the constituent ordering as the ordering of 'attention'.

Arnold, Wasow, Losongco, and Ginstrom (hereafter Arnold et al.) (2000) argue that 'grammatical complexity (heaviness) and discourse status (newness)'[40] influence constituent ordering, and those two factors work in a way to facilitate the speaker's production. Their corpus analysis and an elicitation experiment indicate that speakers tend to postpose new and heavy noun phrases in English, thus introducing the constituent ordering of given-before-new and light-before-heavy. This constituent ordering pattern can be interpreted in two ways. Given and light entities are more accessible and simpler, thus it is easier for a hearer to process (comprehension-based motivation). New and heavy entities are more complex and linguistically more loaded, thus more difficult to produce. By postposing entities that are difficult to produce, speakers can earn themselves extra time to plan and formulate their utterance (production-based motivation). Extra time is important when producing an utterance is under production constraints. Speakers have to go through complex stages, "conceptualizing a message, formulating the grammatical characteristics of the message, and articulating it" (Levelt 1989, cited in Arnold et al. 2000:46), and they have to do it in a short period of time. Arnold et al. found that when speakers have to choose between facilitating the hearer's comprehension and ease of production, they choose the task that benefits themselves. They also claim that 'disfluencies' at the beginning of an utterance are evidence of utterance planning. They support their claim by introducing experimental evidence that when speakers are disfluent at the beginning of an utterance, they use more constructions with 'goal' than with 'theme' in the utterance because constructions with 'theme' are more complex than

[40] Their 'heaviness' indicates syntactic weight in terms of number of words, and 'newness', as opposed to 'givenness', of an entity reflects its referent's previous mention in the preceding discourse. If the referent of an entity has never been mentioned before, the entity is 'new', and if the referent has been mentioned in the previous discourse, the entity is 'given'.

the constructions with 'goal'.[41]

Arnold et al. also indicate that when the length of two constituents are closer, the discourse status of constituents is a more influential factor for the ordering choice, and when the discourse status of constituents are both given, heaviness has more effect on the constituent ordering. This result conforms to Hawkins' observation that his parsing theory strongly predicted the ordering where the discourse motivation fails to predict.

Further, they point out that "experimental evidence has shown that speakers tend to put NPs early in the sentence when they refer to referents that are conceptually accessible (Bock & Irwin 1980, Bock & Warren 1985, Tomlin 1998) or lexically primed (Bock 1986, bock & Irwin 1980)" and "word-order variation is sensitive to the accessibility of word representations at the phonological level" (2000:47).

1.4. Factors to be taken into account

There are two factors that should be taken into account when studying constituent order in Japanese. One is reference form and another is case marking. Givón's study on topic continuity revealed that different reference forms reflect a certain discourse status of referents. Case marking particles and adverbial particles in Japanese demonstrate discourse properties traditionally described in terms of topicality or new/old information.

---

[41] Arnold et al. list examples of goal-first and theme-first constructions, which are used by participants in their dative alternation experiment. Table (i) is an extract from Arnold et al. (2000:43).

Table (i)

Construction	Example
Passivized goal	The white rabbit should be given the carrot.
Double object, goal topicalized	The white rabbit, give the carrot.
Prepositional, goal topicalized	To the white rabbit, give the carrot.
Passivized theme	The carrot should be given to the white rabbit.
Prepositional, theme topicalized	The carrot, give to the white rabbit.

1.4.1. Particle and word order

The Japanese language uses postnominal particles to signal grammatical relations. They are case marking particles such as the nominative *ga*, accusative *o* or dative *ni*. In addition to these case particles, Japanese has adverbial particles such as *wa* and *mo*. The particle *wa* is commonly understood as a topic marker, marking a topic or contrastive topic. The particle *mo* denotes a sense of addition, and the use of particle *mo* suggests the presence of presupposition. It is known that monomoraic particles in Japanese are polysemous, or multi-functional. The adverbial particle *wa* may be contrastive or thematic.[42] The case particle *ni* can be dative, temporal, or locational. Another case particle *o* functions most of the time as a direct object marker of transitive verbs, but also functions as a locational marker for a closed group of intransitive motion verbs. Since the adverbial particle *wa* calls for ellipsis of the case particles *ga* and *o* when it attaches to a subject or direct object, the sentences in (42) and (43) are produced.

(42)	taroo	ga	hanako	o	mita.
	Taro	NOM	Hanako	ACC	saw
	‘Taro saw Hanako.’				

(43)	taroo	wa	hanako	o	mita
	Taro	TOP	Hanako	ACC	saw
	‘Taro saw Hanako.’				

The two sentences, (42) and (43), are allosentences because they depict the same event and the grammatical relations of constituents are the same. The difference in reading lies in the semantic or pragmatic property that particle *wa* denotes. This parallelism has invited intensive studies (e.g. Kuroda 1965; Kuno 1972, 1973; Miyagawa 1987, Shimojo 1995, and many others) of the case particle *ga* and the adverbial particle *wa*. Many

---

[42] Some linguists propose contrastiveness as a single function of *wa* (e.g. Shibatani 1990, Shimojo 2005).

studies appeal to pragmatic notions for describing the properties of the two particles. For example, some view the *ga*-marked noun phrases as carrying 'new' information, or *wa*-marked noun phrases as carrying 'given' information in the previous discourse.

On the one hand, there is a claim that Japanese particles signal discourse properties such as 'new' information and 'given' information, and on the other hand, there is a claim that 'given information tends to be introduced earlier in a sentence in Japanese (Ferreira and Yoshita 2003). A question I raise is, "Can a discourse function be signaled by two different linguistic devices, word order and particles?" If this is possible, how do they interact?

Hakuta (1982) studied the interaction between case marking particles and word order in first language acquisition in Japanese. He conducted experiments on the two possible strategies that Japanese children could use to interpret a sentence. One is the word order strategy, and the other is the particle strategy. Using the word order strategy, children are expected to impose an Agent-Patient-Action sentence pattern on a series of constituents XYZ, interpreting X as agent, Y as patient, and Z as action. Using the particle strategy, children are expected to interpret each constituent according to the information conveyed by the postpositional particle attached to it. Remember that Japanese uses postpositional case particles to signal grammatical roles.

In his experiments, children between the ages 2 and 6 were tested on comprehension of reversible and nonreversible sentences.[43] The examples in (44) to (47)

---

43 Reversible sentences are ones like 'John hit Bill.' The Agent-Patient relation can not be understood from the semantics of constituents alone. Only grammatical roles assign Agent-Patient relation to constituents. In English, for example, word order signals grammatical relations. Nonreversible sentences are one like 'John ate apples.' The Agent-Patient relation can be understood from the semantics of the words 'John' and 'apples'. Commonly humans eat fruit, but not the reverse. Fruit does not eat humans. Therefore, without information on grammatical roles of constituents, John is usually interpreted as agent, and apples as patient.

are reversible sentences.

(44) Active/SOV

AGENT-ga	PATIENT-o	'bit'-active

Taro	ga	Jiro	o	butta.
Taro	NOM	Jiro	ACC	hit.PST

"Taro hit Jiro.'

(45) Active/OSV

PATIENT-o	AGENT-ga	'bit'-active

Jiro	o	Taro	ga	butta.
Jiro	ACC	Taro	NOM	hit.PST

"Taro hit Jiro.'

(46) Passive/SOV

PATIENT-ga	AGENT-ni	'bit'-passive

Jiro	ga	Taro	ni	butareta.
Jiro	NOM	Taro	DAT	hit.PASS.PST

"Jiro was hit by Taro.'

(47) Passive/OSV

AGENT-ni	PATIENT-ga	'bit'-passive

Taro	ni	Jiro	ga	butareta.
Taro	DAT	Jiro	NOM	hit.PASS.PST

"Jiro was hit by Taro.'

The result showed that children comprehend poorly Passive/OSV and Active/OSV. This result suggested, furthermore, that children did not use either the word order strategy or particle strategy. Regarding Passive/SOV, children systematically interpreted *Jiro* as Agent and *Taro* as Patient. This systematic reversal indicates the tendency of children to impose the interpretation of the first noun marked by *ga* as Agent.

In another experiment, children were tested on sentences with Noun-Noun-Verb (NNV), Noun-Verb-Noun (NVN) and Verb-Noun-Noun (VNN) sequences without any noun marking particles in order to assess the role of word order. The result suggested that

Japanese children do not have a particularly strong tendency to impose Agent-Patient-Action interpretation on NNV sequence.

The production test revealed that children acquire control over the particle *ga* before *o*. When they were asked to repeat sentences, they performed poorly on Active/OSV sentences because they impose the particle *ga* on the initial noun while retaining the order of the sentence, Object-Subject-Verb.

Hakuta concludes that Japanese children require the agreement between the particle *ga* and its position within the sentence. This study shows that word order and particles are intertwined closely in Japanese language acquisition.

One strategy to investigate the issue of interaction of particle and word order is to categorize sentences according to their particle markings. When particle marking variation is taken into consideration, a simple sentence like 'X bought Y' (*Katta* 'bought') can yield eight possible variations.

S-O-V					O-S-V				
X	ga	Y	o	katta.	Y	o	X	ga	katta.
X	wa	Y	o	katta.	Y	o	Z	wa	katta.
X	ga	Y	wa	katta.	Y	wa	X	ga	katta.
X	wa	Y	wa	katta.	Y	wa	X	wa	katta.

In this study, all eight possible constituent order variations with different particle markings will be investigated.

1.4.2. Reference form

When a speaker talks about an entity, the speaker has the choice of selecting a linguistic form to describe the entity. This linguistic form is termed the 'reference form'. Givón (1983), seeing topicality as a continuum, introduced the scale of reference forms where zero-anaphors (ellipsis of noun phrases) are more continuous/accessible topics and full noun phrases are more discontinuous/inaccessible topics.

(48) more continuous/accessible topics

zero anaphors
unstressed/bound pronouns ('agreement')
stressed/independent pronouns
full NPs

more discontinuous/inaccessible topics

The underlying principle is that the more difficult it is to identify the referent, the more linguistically loaded the anaphor is. Hinds (1983) examined referential distance and topic persistence of spoken Japanese, and revealed that different reference forms have the property of different continuity in anaphoric and cataphoric discourse. Before I go into the details of his work, I will briefly describe a pattern of typical Japanese discourse.

The Japanese language is well known for its frequent use of ellipsis. Ellipsis is an omission of noun phrases by a speaker when the speaker assumes that the hearer can easily identify the referent. The following illustrates the prototypical behavior of noun phrases in Japanese. EM stands for explicit mention and Ø for zero-anaphor.

```
          EM                    Ø                   EM
------------|----------------------------|------------------------|-------------------->
        Mary went to school.
                              Ø met Sally.
                                                  Mary bought a book.
```

After an entity has been introduced by an explicit mention, the entity is usually referred to with zero-anaphors when the speaker believes that the entity has been registered in the hearer's mind. As the discourse proceeds, the entity gradually decays either with no mention or low frequency of mention. Then, the entity is eventually referred to with explicit mention again when the speaker believes that the hearer might not identify the referent that the speaker is referring to. If there is no other interfering semantically and syntactically comparable entity, the entity retains the high-activated status, being continuously referred to with zero-anaphora. In some cases, a comparable entity is introduced in the recent discourse. Then, the speaker has to explicitly mention the entity in order to avoid confusion in the hearer's mind. In sum, zero-anaphora has the property of low referential distance, and explicit mention has higher referential distance. The use of explicit mention and zero-anaphora in Japanese roughly corresponds to the use of full noun phrases and third person pronouns in English, respectively. It has been suggested that zero-anaphora has a particular pragmatic property. For example, native speakers of Japanese find cases where the recovered subject noun phrases cannot be applied with either the adverbial particle *wa*, case particle *ga*, or zero-particle.[44]

Hinds (1983) acknowledged discourse properties of particles and reference forms in Japanese. He measured referential distance and topic persistence of twelve "grammatical forms" in a narrative, in a semi-structured conversation, and in a natural conversation. The twelve grammatical forms are as follows:

Subject NP-ga	Direct Object NP-o/ni/ga
Subject NP-wa/mo	Direct Object NP-ø
Subject NP-ø	Direct Object Pronoun-o/ni/ga
Subject Pronoun-ga	Direct Object Pronouns-ø
Subject Pronoun-wa/mo	Direct Object Ellipsis
Subject Pronoun-ø	
Subject Ellipsis	

The choice of reference forms is: full noun phrases (NP), pronouns (PRO), and ellipsis (i.e. zero-anaphor). The choice of particles are: *ga, wa, ø, mo, and ni*. The measurements

[44] Zero-particle and particle ellipsis are two distinctive linguistic forms with distinctive functions. The function of zero-particle will be discussed in Section 2.1.in Chapter 2. See Shimojo (2005) for further discussion on zero particle encoding.

are defined as follows (Hinds 1983:58):

(a) Distance – The distance from the present mention of a noun phrase by a particular device and the last clause where the same referent was a semantic argument of that clause, in numbers of clauses.

(b) Decay – The number of clauses to the right from the locus of study that the same referent remains an argument of the predication.

In all three corpora, a narrative, a semi-structured conversation and a natural conversation, the ellipsis showed the property of high discourse continuity. In the semi-structured conversation corpus, there was a big difference in the value of referential distance between (subject ellipsis) and (subject pronouns and subject noun phrases). Subject ellipsis showed high continuity whereas (subject pronouns and subject noun phrases) equally exhibited higher discontinuity. In the natural conversation corpus, the difference in referential distance was gradual from the highest continuity of subject ellipsis to subject pronouns, and to the highest discontinuity of subject noun phrases.

Tables (5) to (8) are from Hinds (1983). In both the semi-structured conversation corpus and the natural conversation corpus, the anaphoric continuity scale was ellipsis > pronouns (PRO) > noun phrases (NP) with ellipsis more continuous.

Table (5) Topic/Subject in semi-structured conversation (Hinds 1983)

category	N(umber)	Average (referential) distance
NP	11	13.1
PRO	3	12.3
Ellipsis	83	1.6

Table (6) Topic/Subject (animate) in natural conversation (Hinds 1983)

category	N(umber)	Average (referential) distance
NP	24	8.6
PRO	20	4.8
Ellipsis	83	2.7

As for the referential distance of inanimate direct objects, the demarcation was between noun phrases on the one hand, and pronouns and ellipsis on the other. Pronouns and ellipsis had relatively higher continuity, and noun phrases showed much lower continuity.

Table (7) Direct Object (inaminate) (Hinds 1983)

category	N(umber)	Average (referential) distance
NP	42	12.5
PRO	7	3.7
Ellipsis	16	2.4

Hinds's data also showed that ellipsis had the property of higher continuity in the succeeding discourse compared to noun phrases and pronouns.

Table (8) Decay of Subject/Topic (inaminate) (Hinds 1983)

category	N(umber)	Average Decay
NP	81	1.1
PRO	13	1.7
Ellipsis	28	2.7

Hinds's study conformed to Givón's reference form scale in that ellipsis is more continuous, that full noun phrases are least continuous, and that pronouns had intermediate continuity status in between ellipsis and full noun phrases.

Given the fact that different reference forms exhibit different discourse properties, it is important to consider the factor of reference forms in examining the function of word ordering. The reference forms relevant to this study are independent pronouns and demonstrative noun phrases. Examples (49) and (50) are extracts from my database. They illustrate a choice of reference form by one speaker, a noun phrase with a demonstrative[45] modifying a noun and an independent pronoun.

(49)

1A:	R-tte	kata	mo	nagaiki-dattatte.			
	R-QT	person	also	long life-COP.PST.TE			
	'Mr. R also lived long.'						
1B:	ee,	<u>ano</u>	<u>kata</u>	91,2	de	kakeoti	yarimasite.
	Yes,	that	person	91,2	at	runaway	did.POL.TE
	'Indeed, he eloped at the age of 91 or 92.'						

[45] Some Japanese grammar books use the term 'attributive' for *kono* 'this~', *ano* 'that~', and *sono* 'its'.

(50)

2A:

H-tte	kata	mo	kanari	de	nihon	ni	irassyaimasita	yo	ne.
H-QT	person	also	fairly	at	Japan	to	come.HNO.POL.	FP	FP

'Mr. B at fairly old age came to Japan, didn't he?'

2B:

so	desu	ne,
so	COP.NPST	FP

'well,'

demo	<u>kare</u>	<u>wa</u>	pianisuto	tosite	wa	son'nani	tyomei	dewanakatta.
but	he	TOP	pianist	as	TOP	such	long.life	COP.PST.NEG

'But he didn't live that long as pianist (i.e. He didn't live that long compared to other pianists).'

Speaker B above chose the demonstrative noun phrase *ano hito* 'that person' to refer to Mr. R in the immediately preceding discourse in (49), whereas the same speaker chose an independent pronoun *kare* 'he' to refer to Mr. B that is also introduced in the immediately preceding discourse in (50). Saeki (1975, 1976, 1998) observed, in his study of positional tendencies of Japanese pronouns and demonstratives, that they tend to come earlier in the sentence, and the sentence-initial demonstratives have no competitive element to confuse the hearer to identify the referent.

1.5. Starting point

Based on the study of R-dislocation and L-dislocation, Givón suggests discourse principles, 'unpredictable' before 'predictable', and more 'important' before 'unimportant', and "Task Urgency" as reasoning behind the principles. Arnold et al. showed that English speakers tend to produce the constituent ordering of given-before-new, and light-before-heavy. Ferreira and Yoshita provide experimental results that Japanese speakers prefer given-before-new. Whichever order preference, new-before-given or given-before-new, a language has, Givón's discourse measurements, referential distance and topic persistence, provide important information on the discourse pragmatic nature of constituents with respect to the choice of one word order over another.

Dryer's activation theory introduces another powerful tool to investigate governing factors for word order choice, namely the status of 'activation' of corresponding cognitive entities of constituents and their focus status. The concept of 'activation' supplements a shortcoming of discourse measurement which only counts the linguistic form of referents (i.e. antecedent). Non-linguistic events, such as visual or conceptual prompts, also activate the cognitive entity corresponding to a linguistic entity.

Hawkins' parsing theory shows that syntactic weight can predict constituent ordering to a certain degree. Yamashita and Chang demonstrate that Japanese speakers' preference is long-before-short. As Hawkins indicated, the EIC accounts for the ordering where discourse principles failed to predict. Yamashita's finding that neither syntactic weight nor referentiality accounts for the ordering of her data as a sole factor, but when they are combined, 95% of her data was accounted for, conforms to Hawkins' statement.

Arnold et al. argue that the earlier introduction of given and light constituents in English is more beneficial to the speaker under the "constraints on planning and production" (2000:47) than to the hearer's comprehension. Givón's referential distance pertains to comprehension-based perspectives, whereas topic persistence does to a productive-based perspective. Arnold et al.'s claim links constituent ordering directly to a conceptual level. Yamashita and Chang suggest the bi-level production system, conceptual level (i.e. discourse factors) and form level (i.e. syntactic weight factor). Ferreira and Yoshita discuss an availability-based framework.

Givón's task urgency, Hawkins' EIC, Japanese scholars' *Kakari-Uke* flow, production-based motivation suggested by Arnold et al., and Dryer's activation all provide theoretical reasoning for the choice of constituent ordering. As Arnold et al. point out, there will probably be many factors that influence the choice of a particular ordering, and the question to ask would be when one factor is effective, and when another is more influential. The issue is not 'which claim is right or wrong' anymore, but to organize various factors and seek a reasoning or a system that can account for all factors.

Chapter 2

Database

2.1. Corpus

The data investigated in this study are spontaneous speech samples from conversations on a TV talk show. The talk show is led by a female host, and the conversation is carried between the female host and a guest of the day. The talk show format involves the female host asking questions to the guest of the day based on roughly prepared topics. With some guests, though not all, the viewer can sense that the female host studied the guest of the day prior to the talk show, and the guest of the day has been informed of what he or she might be asked. However, there are no predetermined conversational turns between the conversation partners, and the speech in this talk show reflects the spontaneous flow of the speaker's thought. The characteristics of natural conversation, such as the overlapping of the speech by the conversation partners, the repetition of the speech, disfluencies and the disruption of the speech, are all present in this data. Therefore, the data in this study demonstrates the nature of natural spoken language. Each talk show lasts about 35-45 minutes. A total of 24 talk show episodes[1]

[1] The data contains 24 episodes, of which four episodes (two female and two male guests) were used in Shimojo (1995).

with 12 female guests and 12 male guests were reviewed for this study. The host and the 24 guests are all native speakers of Japanese. The transcription of 24 episodes contains about 16 hours of conversation.

Each token in the database of this study contains a lexical subject and a lexical direct object[2], and possibly one or more other noun phrases. In Japanese, grammatical relations are marked by postnominal particles. Verb semantics specifies the number of noun phrases in the argument structure. For example, the verb *kawu* 'to buy', as in X *ga* Y *o kawu* 'X buys Y', has two arguments in its semantics, the buyer (Actor) and the things that the buyer buys (Undergoer). Each argument (i.e. noun phrase) is given an appropriate case marking. In case of the verb *kawu* 'to buy', for example, nominative is assigned to Actor, and accusative is assigned to Undergoer. In the same vein, the verb *miseru* 'to show', as in X *ga* Z *ni* Y *o miseru* 'X shows Y to Z', has three arguments, Actor, Undergoer and Recipient which is a person to whom X shows Y. The case marking for the verb *miseru* 'to show' is nominative for Actor, accusative for Undergoer, and dative for Recipient. It is rare that native speakers make mistakes on particle assignment, although there are cases which appear to be an obvious slip of the tongue. Those sentences exhibit the mismatch of postnominal marking according to verb semantics. Another issue surrounding postnominal marking in Japanese is that it is common for case markers to be dropped, and in many cases, it is not certain which case marker is dropped. For example, the sentence (1) lacks both subject and object marking;

[2] Henceforth, "object" stands for "direct object" or "O". "Indirect object" will be noted as "indirect object" or "IO".

yet, the hearer understands ‘your wife’ as subject and ‘pocket money’ as object from the verb semantics and the meanings of noun phrases.

(1)	okusama-ø	okozukai-ø	son’nani	kudasaranai-n-desu	tte
	your.wife	pocket.money	that.much	give.HON.NEG-NOM-COP	COM
	‘I heard that your wife doesn’t give you much pocket money.’				

Researchers used to think that zero-marked noun phrases had no significant linguistic functions. Kuno (1976) claims that the instances of noun phrases without particles (i.e. zero-marked noun phrases) are ellipted *wa*, and Hinds (1976) thought of zero-marked noun phrases as casual speech. Recent work has revealed that zero-particle is a distinct paradigmatic choice with unique functions, which should be distinguished from particle ellipsis. Suzuki (1995) argues that zero-marked phrases have two functions that are distinct from those of *wa*-marked phrases and zero anaphora: a discourse boundary and backgrounding the entities in order to foreground *wa*-marked entities or zero anaphora. Shimojo (2005) reports that zero-particle is typically used with referents whose reference the speaker terminates in the immediate following discourse (defocusing in his term). Fujii and Ono (2000) observe that objects receive zero-particle when their referents are already activated or when the speaker judges that the referents do not need to be activated, and see the choice of object marking, with either accusative o or the ellipsis of accusative *o*, as a device to signal the activation status of the referent to the listener. Given the claims, it is necessary to treat exactly in the same fashion two linguistic forms, zero-marked noun phrases and noun phrases marked by a particle. In this study, any clause which contains two or more noun phrases, regardless of the presence of particles, or the kind of particle, was included in the database.

The tokens in this study for analysis are nonsubordinate clauses. Therefore, adverbial clauses that are non-embedded clauses marked by subordinate conjunctions such as *node* 'because', *kara* 'because/since', *tara* 'if/when', *no ni* 'despite', *te mo* 'even if', *ba* 'if', and *totan* 'as soon as' are not included in the database. For example, Example (2) is a token in this study, but Example (3) is not since it is an adverbial clause marked by the subordinate conjunction *node* 'because'.

(2)	oyazi	ga	sakan'gyo	yatte-ite
	my.old.man	NOM	carpenter	do.TE-PPG.TE
	'My old man was carpenter, and …'			

(3)	oyazi	ga	sakan'gyo	yatte-masita	node
	my.old.man	NOM	carpenter	do.TE-POL.PST	because
	'because my oldman was a carpenter, …'				

The constituents in a token can be simple nouns or noun phrases. Noun phrases can be coordinate phrases, or complex phrases with embedded subordinate clauses such as noun modifying clause (i.e. relative clauses), nominalized clauses and subordinate complement clauses that constitute a part of noun phrases. Nominalized clauses are headed by *no* (nominalizer) and nouns such as *koto* 'fact'. Subordinate complement clauses are marked by *to* (quotation marker). For example, the subject in (4) *kaisya no syatyoo* ' the president of the company' is a coordinate noun phrase, and that in (5) is a complex phrase with a relative clause *hon o yonda* 'read the book' that is a part of the subject noun phrase. The objects in (4) *piano* 'piano' and (5) *tegami* 'letter' are simple nouns. The subject *geizyutu gakka to iu no* 'so-called the art department' in (6) is a nominalized clause headed by the nominalizer *no*.

(4)

kaisya	no	syatyoo	ga	piano	o	hii-ta.
company	GEN	president	NOM	piano	ACC	play-PST

‘The president of the company played piano.’

(5)

hon	o	yonda	hito	ga	tegami	o	kurete
book	ACC	read.PST	person	NOM	letter	ACC	give.TE

‘The person who read the book wrote to me, and …’

(6)

geizyutu	gakka	to	iu	no	ga	sore	o	yatte-iru
art	department	QT	say	PRO	NOM	it	ACC	do.TE-PPG

‘So-called the Art Department is doing it.’

2.2. Measurements

Givón established a quantitative methodology by which some discourse properties of a referent can be measured. In this thesis, Givón’s two discourse measurements, referential distance and topic persistence, are adopted to measure the recency of previous mentions and cataphoric persistence of the preverbal noun phrases in various constituent orders. In addition, the measurement of syntactic weight is employed in order to assess the syntactic heaviness. It should be noted that referential distance, topic persistence and syntactic weight are textual factors, and their measurements are a sort of crude method to approximate some psychological phenomena such as activation.

In what follows, I will discuss what a specific measurement measures, and how it is measured.

2.2.1. Referential distance (RD)

The count of referential distance (RD) is the number of clauses between the entity in question and its most recent coreferential expression in the preceding entity clauses. A coreferential expression could be an overt noun phrase, pronoun, demonstrative, or zero-anaphora. The coding procedure is as follows: if the coreferential expression is found in the second previous clause, it is coded as RD = 2. No mention in the previous 20 clauses involves two cases; one is the case where there is no previous mention at all in the preceding text. They are coded as FM (first mentions). Another case is when there is no coreferential expression in the preceding 20 clauses but it has been mentioned before somewhere in the previous text. They are coded as RD = 20+. Entities with lower referential distance indicate that their referents are mentioned recently in discourse. In addition to counting the RD of an entity in question, the same method is applied to count the RD of the corresponding entity of a predicational proposition.

According to Givón, the discourse properties to be measured by these two measurements are the 'predictability' and 'importance' of a referent. Givón claims that there is a correlation between the measurement of RD and constituent ordering; a higher RD corresponds to noun phrases occurring earlier in the clause.

2.2.2. Topic persistence (TP)

Topic persistence measures how long an entity remains in cataphoric discourse. The count of topic persistence (TP) is the number of mentions of the entity in question in the succeeding 10 clauses. There are two ways of counting topic persistence. One is to count consecutive following clauses that contain the referent of an entity in question. For example, when the corresponding cognitive entity of a constituent in question is mentioned in the immediately succeeding clause, it is coded as TP = 1. As it is mentioned successively in the immediately following clauses, the count goes up as TP = 2,3…10. When the mention is interrupted, the count stops. When there is no repeated mention of the corresponding cognitive entity of an entity in question in the immediately succeeding clause, it is coded as TP = 0. Therefore, even the corresponding cognitive entity of an entity in question is found in the second clause from the utterance, it is also coded as TP = 0, since it is not mentioned in the immediately following clause. Another way of counting TP is to count how many of the following ten clauses contain the referent of an entity in question. If four clauses out of ten contain the corresponding cognitive entity of the entity in question, the value 4 (TP = 4) is assigned. The most persistent entity has, thus, the value 10 (TP = 10), and the least persistent entity has the value zero (TP = 0).

Let us compare the count by these two methods. Say, an entity in question is referred to in the second, fourth, sixth and eighth clauses after the clause in question. It has a TP of 0 by the former method of counting, but of 4 by the latter method. This entity in question is very different from the entity which is not referred to at all in the subsequent discourse. The former method, however, treat both entities the same by assigning both TP = 0. For this reason, I employ the latter method of counting TP in this study.

The lower TP indicates a rapid decay of cognitive status of an entity in question after it is uttered. The higher count of TP indicates the greater cataphoric persistence. In Givón, the higher TP corresponds to constituents introduced earlier in a clause.

2.2.3. Syntactic weight (SW)

A phrasal unit, which has been used in school grammar in Japan, is a way of grouping words together, and is a common way of measuring syntactic heaviness among Traditional Japanese grammarians. One phrasal unit is, for example, a noun with a particle, a predicate, or an adverbial phrase. For example, Sentence (7) contains 5 phrasal units: *omise de* 'in the shop', *ten'in-san ga* 'a sales person', *boosi o* 'a hat', *kabutte* 'wearing', and *sekkyakusiteiru*[3] 'taking care of customers'.

(7)	omise	de	ten'in-san	ga	boosi	o	kabut-te
	shop	LOC	sales-person	SUB	hat	OBJ	wear-TE

sekkyakusiteiru
take.care.of.customer.PPG

"A sales person, wearing a hat in the shop, is taking care of customers."

Hawkins (1994), on the other hand, counts a word, a tense morpheme and a grammatical particles, such as conjunction particles and the sentence-final question particle -*ka*, as a single syntactic unit. According to Hawkins' method, *watasi-wa* (1sg TOP) or *tabe-ta* (eat-PST), for example, is counted as two syntactic units.

Example (8) demonstrates how traditional Japanese grammarians' phrasal unit count and Hawkins' syntactic unit count[4] differ in number. The phrasal unit is marked by the separator (|), and the syntactic unit of Hawkins is separated by a space in between. In (8), the phrasal unit count is 9, and Hawkins' count yields 21.

---

[3] *Sekkykusiteiru* is decomposed as *sekkyaku-site-iru* (take.care.of.customers-do.PPG) 'taking care of customers'.

[4] Since Hawkins does not provide an exhaustive list of morphemes and particles that are counted as one syntactic unit, I will follow my interpretation of Hawkins' way of counting syntactic units with discretion to be consistent as to what to count for one syntactic unit throughout the data.

(8) koo iu | huu ni | bunsyoo de | hyoogen suru | koto ga |
this say way DAT sentence INS expression do fact NOM

dekiru no dat te | iu no o |
can GEN COP TE say GEN ACC

atasi wa | omot ta |
1sg TOP think PST

‘That (we) can express in this way, I thought.’

In this thesis, I will use Hawkins’ method. Hawkins’ method reflects the presence or absence of the postnominal marking while the phrasal unit count does not. For example, both *watasi-ø* (1sg-zero marking) and *watasi-wa* (1sg-TOP) are counted as one in the phrasal unit count, but in Hawkins’ method the former is one syntactic unit and the latter has two syntactic units. Since my data include zero-marked noun phrases, it would be more accurate to use Hawkins’ method than the phrasal unit count.

In Hawkins’ parsing theory, syntactically heavier entities come before syntactically lighter ones in left-branching languages such as Japanese.

2.2.4. The definition of ‘clause’ for the purpose of counts

One important concept for discourse measurements in the Givónian tradition is that a clause is an informational unit in language processing, and a clause is defined as a syntactic unit which conveys a proposition. There are different definitions of the notion of clause, such as the view that defines a clause as a construction with a finite verb (Lehman 1988), or the generative Grammarian view that defines a clause as a string that is dominated by the node S (e.g. Kuno 1973). Moreover, not all languages necessarily have the same type of clauses.

In principle, I count a verb (or a predicate) for a clause in this study. In what follows, I will explain how the one-verb-for-one-clause principle works by referring to examples from two studies that provide insight on identifying a clausal unit in Japanese. One is the classification of interclausal relational structure by Ohori (1992), and another is a study of TE-linkage (cf. 'gerund' in Myhill 1992) by Hasegawa (1992). Ohori (1992) and Hasegawa (1992) investigated Japanese clause linkage in the framework of Role and Reference Grammar (hereafter RRG, cf. Foley and Van Valin 1985, Van Valin 1984, 1987, 1993, Van Valin and LaPolla 1997), which assumes a layered structure in the clause. The innermost layer is the nucleus, which is the predicate, then the core which consists of the nucleus and the arguments of the predicates, and the clause which contains at least one core, with or without periphery (i.e. non-argument of the predicate), and the core-external precore slot (e.g. to accommodate focused elements). RRG posits two parameters for clausal linkage. One is Juncture and another is Nexus. Juncture has three layers of grammatical levels which correspond to three layers of the clause, nuclear level, core level, and clausal level. Nexus is the trichotomy of Coordination [-dependent, -embedded], Cosubordination [+dependent, -embedded], and Subordination [+dependent, +embedded], based upon the nature of the dependency between the linked units.

2.2.4.1. Clausal coordination

Clausal coordination is linkage between two coordinate clauses. Clausal coordination linkage is linkage by conjunctives such as *ga* 'and' as in (9) and 'verbal head linkage' (Ohori 1992) as in (10)-(12). 'Verbal head linkage' includes SI-linkage as in (10) and TE-linkage as in (11) or (12). Example (11) is TE-linkage with disjoint subjects[5] and Example (12) is TE-linkage with overtly present subjects. Examples (9) and (11)-(12) are from Hasegawa (1992) and Example (10) is from Ohori (1992). Each sentence in (9)-(12) consists of two clauses, and each clause contains one verb.

(9)	keeki	ga	ar-imasu	**ga**,	tabe-masen	ka
	cake	NOM	exist-POL	**CONJ**,	eat-POL.NEG	Q

‘There is cake (= (I) have cake). Would you like to have some?

(10)	ame	ga	hut-ta-**si**	kaze	mo	hidoku	hui-ta.
	rain	NOM	fall-PST-**SI**	wind	also	terribly	blow-PST

‘It rained, and the wind also blew terribly.’

(11)	zyoon	ga	gitaa	o	hi-ite	hiro	ga	utau
	Joan	NOM	guitar	ACC	play-TE	Hiro	NOM	sing-NPST

‘Joan will play the guitar, and Hiro will sing.’

---

5 Hasegawa (1992) points out that with TE-linkage, disjoint reference is only permitted in case when the subjects of both linked units are overtly present as (11) or (i). If either or both of the subjects are missing, disjoint reference is prohibited as in (ii), (iii), and (iv). If the switch reference is signaled by morpho-lexical devices, such as an honorific predicate *go*, as in (v), or psych-predicate, disjoint reference is permitted. Examples (i)-(vi) are from Hasegawa (1992).

(i)	zyoon	ga	setumee-site	hiro	wa	nattoku-simasita.
	Joan	NOM	explanation-do.TE	Hiro	TOP	compliance-do.POL.PST

‘Joan explained, and Hiro understood (it).’

(ii)#	zyoon	ga	setumee-site	Ø	nattoku-simasita
	Joan	NOM	explanation-do.TE	Ø	compliance-do.POL.PST

‘Joan explained, and (he) understood (it).’ [intended]

(ii)#	Ø	setumee-site	hiro	wa	nattoku-simasita
	Ø	explanation-do.TE	Hiro	TOP	compliance-do.POL.PST

‘(She) explained, and Hiro understood (it).’ [intended]

(iv)#	Ø	setumee-site	Ø	nattoku-simasita
	Ø	explanation-do.TE	Ø	compliance-do.POL.PST

‘(She) explained, and (he) understood (it).’ [intended]

(v)	zyoon	ga	go-setumee-site	Ø	nattoku-nasaimashita.
	Joan	NOM	HON-explanation-do.TE	Ø	compliance-do.HON.PST

‘Joan explained, and (he) understood (it).’

(12) zyoon ga gitaa o hi-ite zyoon ga utau.
Joan NOM guitar ACC play.TE Joan NOM sing.NPST
'Joan will play the guitar, and Joan will sing.'

2.2.4.2. Clausal subordination

Clausal subordination is linkage between two clauses, one of which is subordinate to another. There are two types of subordination at the clausal level. One type of subordination functions as arguments and another as adjuncts.[6] The former includes noun-modifying clauses (i.e. relative clauses) as in (13), nominalized clauses as in (14), and subordinate complement clauses as in (15). The example in the latter case is adverbial phrases as in (16). Examples (13)-(15) are from my database, and Example (16) is from Ohori (1992).

(13) Noun-modifying clause

anata o sizi-site-**kudasatta** sensei wa
2sg ACC support-do.TE-**give.PST** teacher TOP

anata no go-katuyaku o o-yorokobide
2sg GEN HON-activity ACC POL-be.glad.TE

'The teacher who supported you is happy about your being around as an actor.'

(14) Nominalized clause

o-mise de ten'in-san ga boosi o kabutte-i-ru
POL-shop LOC sales.person NOM hat ACC wear.TE-STAT-NPST

no wa hen-da
NMLZ TOP strange-COP

'It is strange that a sales person wears a hat in the shop.'

---

[6] Ohori (1992) calls the former type of subordination clause as 'nominal head linkage', and the latter as 'verbal head linkage'.

(15) Subordinate complement

H	wa	boku	ni	dake	okur-e	**tte**	it-ta.	(4-14)
H	TOP	1sg	DAT	only	send-IMP	**QT**	say-PST	

'H said to send (it) only to me.'

(16) Adverbial

ame	ga	hut-ta	**node**	taroo	wa	gakkoo	ni	ik-anakatta
rain	NOM	fall-PST	**because**	Taro	TOP	school	DAT	go-NEG.PST

'Because it rained, Taro didn't go to school.'

The clause marked by *tte* (colloquial form of comitative *to*) in (15), i.e. *boku ni dake ok-re* 'sent (it) only to me', is a complement of the matrix verb *i-tta* 'said'. The noun-modifying clause in (13) and the *no*-marked clause in (14) are sentential arguments of the matrix predicates *o-yorokobi-de* 'is happy' and *hen-da* 'is strange', respectively. When the subordinate linkage unit occupies an argument position of the matrix predicate, it is not considered as a clausal unit. In other words, nominalized clauses headed by nominalizers *no* or *koto*, noun-modifying clauses (i.e. relative clauses), and subordinate complement clauses marked by quotative *to, te*, and *tte* are a part of a sentential noun phrase, and thus they are not treated as independent clauses. Therefore, in this study, the verb in the linked units in (13)-(15) is not counted for the clause.

The adverbial clause in (16), *ame ga hutta node* 'because it rained', which is signaled by a subordinate conjunction node 'because', functions as an adverbial. The adverbial clause is subordinate to the matrix clause *taroo wa gakko ni ikanakatta* 'Taro didn't go to school.', but not a part of a sentential argument. Thus, the adverbial clause and the matrix clause are treated as independent clauses. Therefore, the sentence (16) consists of two clauses, each of which contains a verb.

What we have to distinguish is two types of noun modifying clauses. One is noun modifying clauses such as in (13), and another is noun modifying clauses whose head noun is grammaticalized as in (17).[7] Example (17) is from my database.

(17) Grammaticalization of the head noun

min'na	obentoo	toka	tabete-i-ru	**toki**	ni,
everybody	lunch-box	etc.	eat-STAT-NPST	**time**	TEMP

'When everybody is eating (their) lunch box,'

watasi	dake	tabe-nai-de	kon'na	yat-te-ru	to,
1sg.	only	eat-NEG-TE	this.manner	do-TE-NPST	if/when

'if I am doing like this,'

Ohori (1992) points out that when the head noun in a noun-modifying clause expresses some relational concept such as spatial (e.g. *tokoro* 'place'), temporal (e.g. *toki* 'time'), or qualitative (e.g. *kuse* 'bad habit') concepts, they are an adjunct in the matrix predicate and structurally indistinguishable from an adverbial subordinate clause such as (16). Ohori (1992) uses, as a diagnostic tool, the omissibility of the oblique case *ni* after the head noun, which indicates grammaticalization. Moreover, the semantic shift in some of the head nouns (e.g. *kuse* 'bad habit'--> 'in spite of the fact that/although') in this construction signals grammaticalization. While a noun-modifying clause such as (14) is not considered as an independent clausal unit, the linkage unit lead by grammaticalized head nouns such as (17) is counted as an independent clause.

---

[7] See Ohori (1992) for a list of grammaticized nouns in nominal head linkage.

2.2.4.3. Clausal cosubordination

In section 2.3.5.1., we saw TE-linkage in clausal coordination. Example (18) demonstrates TE-linkage in clausal cosubordination. Clausal cosubordination is linkage between two clauses, one of which is dependent on another. Example (18) is from Hasegawa (1992).

(18)	hayaku	sigoto	o	**sumasete**	uti	ni	kaeri-nasai.
	quickly	work	ACC	**finish-TE**	home	LOC	return-IMP

‘Finish (your) work quickly, and go home!’

The first unit *sumasete* ‘finish-TE’ in (18) exhibits dependency on the second unit with respect to illocutionary force[8], i.e. imperative *nasai*. Despite the dependency, there are two separate events: ‘your finishing work quickly’ and ‘your going home’. Therefore, (18) consists of two clauses, each of which contains a verb.

Hasegawa (1992) presents another example of clausal cosubordination, V1-*ta-ri* V2-*ta-ri s*- Construction ‘do V1 and V2 among other things’ as in (19)

(19)	kinoo	wa	[tomodati	ga	**ki-ta-ri**	
	yesterday	TOP	friends	NOM	**come-PST-CS(connective suffix)**	
	conpyutaa	ga	**koware-ta-ri**	**si-ta**]	node	
	computer	NOM	**break-PST-CS**	**do-PST**	because	
	nanimo	deki-nakat-ta.				
	anything	can.do-NEG-PST				

‘Yesterday, [a friend came and the computer broke down], so I couldn’t accomplish anything.’

8 Illocutionary force is a clausal Operator in RRG. Operator consists of morphemes which are the realization of grammatical categories such as aspect, tense, modal, negation, and illocutionary force.

In (19) *si-ta* 'do-PST' in the second linked unit appears to be the only element indicating tense, and it scopes over the first linked unit. Although the two units share the same tense marker, each linked unit has disjoint referents and disjoint events.

Ba-marked clauses are also dependent on the matrix clause with regard to tense. Example (20) is from Ohori (1992) and (21) is from my database.

(20) ame-ga hure-**ba** siai wa tyuusi-dat-ta
rain-NOM fall-**BA** game TOP suspended-COP-PST
'If it had rained, the game would have been suspended.'

(21) ik-**eba** wakar-imasu yo.[9]
go-**BA** understand-POL.NPST FP
'if (you) go, (you) will find (it) out, certainly.'

The linked units, however, involve two independent events; in (20) they are 'to rain' and 'the game being suspended', and in (21) they are 'your going' and 'your finding out'.

2.2.4.4. Core coordination

When the linked units share the same argument (i.e. the subject in most of the cases) and each linked unit exhibits disjoint events, the linkage is categorized as core coordination in RRG. An example of core coordination is *zu-ni* construction 'without

---

[9] The subjects of the linked units may be coreferential or noncoreferential.

(vii) i-**eba** wakaru hazu-desu.
say-**BA** understand should-COP.NPST
'if (you) say (it), (he/she/they) should understand (it).'

The utterance (vii) has no overt subject NPs, as seen in the gloss. The semantics of the linked clauses construes the missing subject of *i-eba* 'if say' as 'I', and that of *wakaru hazu-desu* ' as a third person/persons 'he or she or they'. The linked units *i- eba* 'if say' and *wakaru hazu-desu* 'should understand' have disjoint reference. (18) and (21) are also cases of subject ellipsis. In (18) and (21) the subjects, construed by the semantics of predicate and illocutionary force of linked units, are coreferential.

doing ~', as in (22), and another is TE-linkage with the second subject ellipsis as in (23). Both (22) and (23) are from Hasegawa (1992).

(22)

zyoon	wa	repooto	o	**yoma-zu**	**ni**	kaigi	ni	ki-ta.
Joan	TOP	report	ACC	**read-NEG**	**CMPL**	meeting	LOC	come-PST

'Joan came to the meeting without reading the reports,'

(23)

zyoon	ga	okane	o	**tame-te**	kuruma	o	kawa-nakatta	koto[10]
Joan	NOM	money	ACC	**save-TE**	car	ACC	bought-NEG.PST	fact

'the fact that Joan saved money and didn't buy a car.' [11]

Core coordination involves two clauses, each of which contains a predicate.

2.2.4.5. Core subordination and core cosubordination

Core subordination involves the subordination of one verbal element to another verbal element in a predicate, as in (24). An example of core cosubordination is purpose *ni* construction, as in (25). Both examples are from Hasegawa (1992).

(24)

anata	wa	moo	kaette	i-	i
you	TOP	already	go-home-TE	be-permitted	NPST

'You may go home now.'

---

[10] Hasegawa prefers to formulate the example with the nominative *ga*, and presented this example in the form of a nominalized clause with *koto* 'fact/(abstract) thing', where the nominative *ga* is mandatory.

[11] Hasegawa (1992) points out that clauses like (23) are structurally ambiguous. Example (23) may be interpreted as (viii).

(viii)

zyoon	ga	okane	o	**tame-te**	kuruma	o	kawa-nakatta	koto
Joan	NOM	money	ACC	**save-TE**	car	ACC	bought-NEG.PST	fact

'the fact that Joan bought a car without saving money.'

The interpretation in (viii) indicates that the negation *nakatta* 'NEG.PST' scopes both *tamete* 'save-TE' and *kawa-* 'bought'. The intonation might fall after the first linked unit *tame-te* 'save-TE' in (23). Nevertheless, it is subtle, and it cannot be denied that such an intonation boundary might be present when (viii) is uttered. In this thesis, regardless of structural differences, TE-linkage clauses such as (23) and (viii) are counted as two clausal units.

(25) zyoon wa hon o kari ni tosyokan ni ika-nakat-ta
Joan TOP book ACC borrow CMPL library LOC go-NEG-PST
'Joan didn't go to the library to borrow some books.'

In this study, core subordinate construction such as TE-ii (permission) in (24) is treated as an extended predicate, thus, (24) has a subject and a verb, SV construction. The purpose phrase like *kari-ni* 'in order to borrow' as in (25) is treated as one noun phrase, thus (25) is 4NPs construction, Subject-Direct Object-Purpose NP-Locative NP-Verb construction.

2.3. Constructions in Japanese

The study of word order in Japanese has been overlooked, probably because Japanese is labeled as a rigid verb final language. The data in this study, however, demonstrate notable variations in constituent order, including postverbal constituents[12]. Tables (1), (2) and (3) below present constructions and their constituent order variations.

Table (1) Constituent order variation of constructions with verb

Construction with Verb	Constituent Order Variation				
SOV	SOV	OSV	SVO	OVS	VSO
SQV	SQV	QSV	SQV	QVS	
SDV	SDV	DSV			VSD
SCVi	SCVi			CViS	
SLVm	SLVm	LSVm	SVmL		
SLVi	SLVi	LSVi		LViS	
SLVex	SLVex	LSVex	SVexL	LVexS	
SOVi	SOVi	OSVi		OViS	

S=subject, O=direct object, V=verb, Q=quotative, D=dative, C=complement, Vi=intransitive verb, L=locative, Vm=motion verb, and Vex=existential verb

[12] In this study, when there is no intervening pause between the verb and the postverbal element, the postverbal element is identified as a postposed element. When there is a pause, the postverbal element is treated as an independent phrase categorized as afterthought.

Table (2) Constituent order variation of constructions with adjective

Construction with Adjective	Constituent Order Variation	
Subject-Topic-Adjective		T-S-Adj
Experiencer-Stimulus-Adjective	Ex-Sti-Adj	Sti-Ex-Adj

S=subject, T=topic, Adj=adjective, Ex=experiencer, and Sti=stimulus

Table (3) Constituent order variation of constructions with copula

Construction with Copula	Constituent Order Variation			
Experiencer-Stimulus-C	Ex-Sti-C	Sti-Ex-C	Sti-C-Ex	C-Ex-Sti

Ex=experiencer, Sti=stimulus, and C=copula

All constructions of verb, except the one with complements (i.e. the SCVi construction), had the verb-final construction: SXV and XSV. In the verb-medial construction, most of the constructions exhibited subject-final XVS order instead of subject-initial SVX order; only constructions with motion verb and existential verb preferred subject-initial SVX order, where X is locative. Only two constructions with verb had the verb-initial construction, and the constituent order was subject-medial VSX order instead of subject-final VXS order. The construction with complements (i.e. the SCVi construction) did not show much flexibility in constituent variation. Interestingly, the constructions with adjective and copula exhibited the same pattern in constituent order variation as the constructions with verb. In the next section, I will give a brief description of each construction and an example for each variation of constituent order. I will also show the number of tokens in each constituent order variation in both main clauses and subordinate clauses.

2.3.1. The construction with verb

2.3.1.1. SOV (subject-object-transitive verb) construction

The SOV construction contains a subject, object and transitive verb. Mainly the subject is the Actor and the object is the Undergoer or Theme. The object is prototypically marked by the accusative *o*.[13] Sentence (26) is an example of the SOV construction.

(26)	SOV			
4-14	boku	wa	sore o	kotowatta-n-desu
	1sg.male	TOP	that ACC	refuse.PST-NMLZ-COP
	'I turned it down.'			

The order variations found in the data are SOV, OSV, SVO, OVS[14] and VSO. All word order variations, except VSO, are found in both the main and subordinate clauses.

Table (4) The number of tokens in each constituent order variation of the SOV construction

SOV	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SOV		OSV		SVO		OVS		VSO				
2NPs	130	99	37	17	13	2	10	2	1	0	191	120	311
≥3NPs	92	42	22	6	6	0	7	0	0	0	127	48	175
Total	222	141	59	23	19	2	17	2	1	0	318	168	486

---

[13] Occasionally the locational *o*, as in *miti o aruku* (street ACC walk) 'to walk on the street', is found. The particle *o* is also found in an utterance like (a). Ohori (1992) provides structural analysis for this type of *o* usage in terms of "internally-headed relative clauses".

(a)	o-isogasii	tokoro	o	o-zyma-site	sumimasen
	POL-busy	place	ACC	POL-bother-do.TE	sorry
	'I am sorry to bother you when you are busy.'				

[14] According to Shimojo's data, among 119 postposed elements, 37 are subjects while only 7 are objects (1995:112). Recent research on postposing in Japanese has elucidated that postposed elements may represent new information but the information is relatively unimportant compared to the pre-predicative elements (e.g. Simon 1989, Fujii 1991). More precisely, they function as defocusing (e.g. Clancy 1982, Shimojo 2005), i.e. "the process of deactivating a referent in one's cognitive focus of attention. Defocusing of a referent occurs if there is no longer focusing of the referent". (Shimojo 2005:34)

(27) OSV

12-2	sore	o	watasi	wa	haiken-simasita
	it	ACC	1sg	TOP	see.HUM-do.PST.POL

‘I took a look at it.’

(28) SVO

11-12	atasi	mi-masita	toroya-no-on’na-tati
	1sg	see-PST.POL	Troy-GEN-woman-PL

‘I saw *The Women of Troy*.’

(29) OVS

23-51	sono	doitu	no	eiga	toka	mi-masita	yo	atakusi
	its	Germany	GEN	movie	etc.	see-PST.POL	FP	1sg.POL

‘I saw that German movie and others.’

(30) VSO

22-10	mata	yoku	oboete-irassyaru-n-desu
	also	well	remember.TE-HON.PPG-NMLZ-COP

kono	kata	ga	iron’na	koto	no	nen’goo	toka
this	person	NOM	various	matter	GEN	year	etc.

‘Also this person remembers well the year of an event and such.’

2.3.1.2. SQV (subject-quotation-transitive verb) construction

The SQV construction involves a subject, quotation and transitive or intransitive verb. The verb in the SQV construction is predominantly *iu* ‘to say’ and *omou* ‘to think’. Sentence (31) is an example of the SQV construction.

(31) SQV

14-3b	anata	ga	manzyusyage	to	ossyata	no	yo
	2sg	NOM	cluster-amaryllis	QT	say.HON.PST	NMLZ	FP

‘You said “(I want) Cluster-amaryllis”, indeed.’

Four word order variations are found: SQV, QSV, SVQ and QVS. Only SQV and QSV are found in both main and subordinate clauses. The number of SQV clauses is the greatest number of tokens in the data in this study.

Table (5) The number of tokens in each constituent order variation of the SQV construction

SQV	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SQV		QSV		SVQ		QVS						
2NPs	147	44	33	13	2	0	10	2			192	59	251
≥3NPs	29	8	5	0	1	0	0	0			35	8	43
Total	176	52	38	13	3	0	10	2			227	67	294

(32) QSV

11-15 yakamasii to syuzin wa omotteru-kamosiremasen
noisy QT my.husband TOP think.PPG-may
'My husband may be thinking it is noisy.'

(33) SVQ

5-22 watasi siranakatta-n-desu yo
1sg know.NEG.PST-NMLZ-COP FP

AD-san ga taihen'na sigoto da tte
AD-Mr. NOM tough job COP QT

'I didn't know that AD (Assistant Director) is a tough job.'

(34) QVS

13-14 kuru to omo-imasita yo watasi wa
come QT think-PST.POL FP 1sg TOP
'I thought (she) would come.'

2.3.1.3. SDV (subject-dative-intransitive verb) construction

The SDV construction contains a subject, dative and intransitive verb. The dative is subcategorized by the verb and takes *ni* as a particle, which denotes the meaning of 'involvements'. The verbs in this construction are, for example, ~ *ni au* 'to meet someone', ~ *ni niteiru* 'to look like someone/something'. Sentence (35) is an example of the SDV construction.

(35) SDV
1-3 T-san ga anata ni ai-ni-iku
T- Ms NOM 2sg DAT meet-DAT (purpose)-go
'Ms. T goes to see you.'

Table (6) The number of tokens in each constituent order variation of the SDV construction

SDV	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SDV		DSV						VSD				
2NPs	15	12	1	3					1	0	17	15	32
≥3NPs	9	2	1	0					0	0	10	2	12
Total	24	14	2	3					1	0	27	17	44

Three constituent order variations are found: SDV, DSV and VSD. SDV and DSV are found in both main and subordinate clauses.

(36) DSV
5-14 I-san ni kao ga niteru tte iw-areru
I-Ms DAT face NOM resemble QT say-PASS
'People say that my face looks like Ms. I.'

(37) VSD
13-21 ie-nai-n-desu ne boku wa keisatu ni
say.POT-NEG-NMLZ-COP FP 1sg.male TOP police DAT
'I cannot tell the Police.'

2.3.1.4. SCVi (subject-complement-intransitive verb) construction

The SCVi construction contains a subject, complement and intransitive verb, such as *naru* 'to become' or *suru* 'to do'. The complement takes the particles *to* or *ni*, which denotes the meaning of 'to become' or 'to change to'. *~ ni naru or ~ to naru* denotes a natural occurrence of something becomes something, and *~ ni suru or ~ to suru* involves an intention of making something become something. The SCVi construction was the least flexible construction in terms of order variation. There are only two orders found, SCVi and CViS. The word order of SCVi order is found in both main and subordinate clauses.

Table (7) The number of tokens in each constituent order variation of the SCVi construction

SCVi	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SCVi						CViS						
2NPs	48	32					3	0			51	32	83
≥3NPs	8	2					0	0			8	2	10
Total	56	34					3	0			59	34	93

Sentence (38) is an example of SCVi order tokens, and (39) is of CViS order tokens.

(38) SCVi
16-19 sore ga kasuregoe ni natte
it NOM hoarse.voice to become.TE
'It became a hoarse voice.'

(39) CViS
2-8 noirooze ni nat-tyatta-n-desu
neurosis to become-PST-NMLZ-COP
'I have had a nervous breakdown.'

2.3.1.5. SLVm (subject-locative-motion verb) construction

The construction SLVm contains a subject, locative and motion verb, such as *iku* 'to go', *kuru* 'to come', or *kaeru* 'to return'. The locative is marked with the particles *ni*, *kara*, *made* or zero marking. The particle *ni* 'to' denotes a goal, *kara* 'from' a starting point, and *made* 'up to' a destination. Sentence (40) is an example of the SLVm construction.

(40) SLVm
17-3 watasi wa pari e iku no ne
1sg TOP Paris LOC go FP FP
'I will go to Paris.'

Table (8) The number of tokens in each constituent order variation of the SLVm construction

	SLVm	Verb-Final				Verb-Medial		Total		Total
		main	sub	main	sub	main	sub	main	sub	
Locative ni/he 'to'	order	SLVm		LSVm		SVmL				
	2NPs	51	35	14	7	4	2	69	44	113
	≥3NPs	16	4	1	0	1	0	18	4	22
	Total	67	39	15	7	5	2	87	48	135
Locative kara 'from'	order	SLVm		LSVm						
	2NPs	1	5	4	1			5	6	11
	≥3NPs	3	2	3	0			6	2	8
	Total	4	7	7	1			11	8	19
Locative made 'upto'	order	SLVm								
	2NPs	2	0					2	0	2
	≥3NPs	0	0					0	0	0
	Total	2	0					2	0	2
	Total	73	46	22	8	5	2	100	56	156

There are three order variations found, SLVm, LSVm and SVmL. The verb-final order, SLVi and LSVi, are found in both main and subordinate clauses.

(41) LSVm
14-18 byooin ni kyuukyuusya ga yattekita
hospital LOC ambulance NOM come.PST
'An ambulance came to the hospital.'

(42) SVmL
9-9 hiiru nomerikomu wake zyanai desu ka undoozyoo ni
heel plunge reason COP.NEG COP Q playground LOC
'The heel naturally plunges into the playground, doesn't it?'

2.3.1.6. SLVi (subject-locative-stative verb) construction

The construction SLVi is similar to the SLVm construction, but different in that SLVi involves stative verbs instead of motion verbs. Stative verbs are intransitive verbs, such as *nokoru* 'to remain', compound verbs in the form of *te-iru* (TE-be), denoting a stative state, such as *aite-iru* (open.TE-be) 'is open', or compound verbs in the form of *te-aru* (TE-exist),[15] denoting a resultative state, such as *oite-aru* (place.TE-exist) 'it is placed (as a result of someone placing it)'. The postnominal marking of the locative is *ni*. Sentence (43) is an example of SLVi order.

(43) SLVi

23-24	sore	ga	zutto	kioku	ni	nokor-imasu	ne
	it	NOM	for.a.long.time	memory	LOC	remain-POL.NPST	FP

'It stays in the memory for a long time.'

Table (9) The number of tokens in each constituent order variation of the SLV iconstruction

SLVi	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SLVi		LSVi				LViS						
2NPs	10	6	5	7			1	0			16	13	29
≥3NPs	1	1	1	0			0	0			2	1	3
Total	11	7	6	7			1	0			18	14	32

There are three variations, SLVi, LSVi and LViS. SLVi and LSVi are found in both main and subordinate clauses.

(44) LSVi

11-11	waki	no	semai	tokoro	ni	watasi-domo	wa	sundorimasu
	side	GEN	small	place	LOC	1sg-PL.HUM	TOP	live.HUM.POL

'We live in a small place on a byway .'

(45) LViS

9-16	buranko	ni	notteru-n-desu	haizi	ga	ne
	swing	LOC	ride.PPG-NMLZ-COP	Haidi	NOM	FP

'Haidi was on a swing.'

---

[15] The TE-verbs in the compound structure of *te-aru* are transitive verbs, such as *akeru* 'to open something' or *simeru* 'to close something', and the use of transitive verbs implies the actor's intention about the action.

2.3.1.7. SLVex (subject-locative-existential verb *aru* 'exist' and *iru* 'be') construction

The construction SLVex contains a subject, locative and existential verb, *aru* 'exist' or *iru* 'be'. The verb *aru* 'exist' co-occurs with inanimate subjects, and the verb *iru* 'be' with animate subjects. In the present data, the occurrence of the verb *aru* is much greater than that of the verb *iru*. Sentence (46) is an example of the SLVex construction with the verb *aru*, and (47) is with the verb *iru*.

(46)
22-11 meron ga mukoo no hoo ni
melon NOM over.there GEN direction LOC

atta-n-desu yo ne
exist.PST-NMLZ-COP FP FP

'There was a melon over there.'

(47)
18-21 U-san ga usiro ni iru
U-Mr. NOM behind LOC be.NPST
'Mr. U is behind.'

Table (8) The number of tokens in each constituent order variation of the SLVex construction

SLVex	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SLVex (iru)		LSVex (iru)		SVexL (iru)		LVexS (iru)						
2NPs	7	5	24	6	1	0	1	0			33	11	44
≥3NPs	3	1	0	0	0	0	0	0			3	1	4
Total	10	6	24	6	1	0	1	0			36	12	48
order	SLVex (aru)		LSVex (aru)		SVexL (aru)								
2NPs	9	6	78	32	10	0					97	38	135
≥3NPs	1	1	15	3	0	0					16	4	20
Total	10	7	93	35	10	0					113	42	155

Two variations are found with the verb *aru* 'exist', SLVex and LSVex, and three variations with the verb *iru* 'be', SLVex, LSVex, and SVex L. Sentence (48) is an example of LSVex order with the verb *aru*, (49) is an example of SVexL order with the verb *aru*, and (50) is an example of LVexS order with the verb *iru*.

(48)	LSVex					
10-20	mukoo	ni	sooiu	mono	ga	aru
	over.there	LOC	such	thing	NOM	exist.NPST
	'There is that kind of thing over there.'					

(49)	SVexL				
16-19	denwa	an-no[16]	soko	no	uti
	telephone	exist-Q	that	GEN	house
	'Is there any phone in that house?'				

(50)	LVexS									
11-11	asoko	ni	Kanda-gawa	ni	orimasu	kedomo	ne	koi	ga
	there	LOC	Kanda-river	LOC	be.HUM.POL.NPST	though	FP	carp	NOM
	'The carp is there, in the Kanda river.'								

Some locatives have a possessive reading[17]. The locative in (51) is a second person singular pronoun and the possessor of the entity denoted by the subject, *kodomo* 'children'.

(51)				
20-6	anata	kodomo	san'nin	iru
	2sg	children	3.people	be
	'You have three children.'			

The possessor reading of the locative is more frequent with the verb *aru* 'exist' (41.2%) than *iru* 'be' (6.8%). The token distribution in terms of order variations is not much different for possessor reading locatives and non-possessor reading locatives. However, the postnominal marking of locatives makes a difference in the frequency of a possessive

[16] *an-no* is the contraction form of *aru-no*.

[17] Shibatani (1990) differentiates the possessive verb from the existential verb in that with the possessive verb the *ni*-marked nominal is a subject while it is a locative, with the existential verb.

reading or a non-possessive reading of locatives. Tables (11) and (12) show the postnominal marking of locatives and the number of tokens with possessive reading locatives and non-possessive reading locatives.

Table (11) Postnominal marking of the locative in constructions with the verb *iru* 'be'

Marking on locative	Order Variations												Total # of tokens	
	SLVex (iru)				LSVex (iru)				SVexL (iru)		LVexS (iru)			
	Main		Sub		Main		Sub		Main		Main			
	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs
ni	6	3	5	1	16	0	4	0	1	0	1	0	33	4
ni-wa	0	0	0	0	1	0	1	0	0	0	0	0	2	0
ni-mo	0	0	0	0	1	0	0	0	0	0	0	0	1	0
de	0	0	0	0	0	0	1	0	0	0	0	0	1	0
wa	0	0	0	0	1	0	0	0	0	0	0	0	1	0
mo	0	0	0	0	2	0	0	0	0	0	0	0	2	0
ø	1	0	0	0	3	0	0	0	0	0	0	0	4	0
Total	7	3	5	1	24	0	6	0	1	0	1	0	44	4
Total	10		6		24		6		1		1		48	

Table (12) Postnominal marking of the locative in construcctions with the verb *aru* 'exist'.

marking on locative	Order Variations										Total # of tokens	
	SLVex (aru)				LSVex (aru)				SVexL (aru)			
	Main		Sub		Main		Sub		Main			
	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs	2NPs	≥3NPs
ni	5	0	4	1	15	2	6	2	3	0	33	5
ni-wa	1	0	0	0	2	2	3	0	2	0	8	2
ni-mo	2	0	1	0	0	0	1	0	0	0	4	0
de	0	0	0	0	2	0	0	0	0	0	2	0
ga	0	0	0	0	3	0	2	1	0	0	5	1
wa	1	0	0	0	23	5	12	0	2	0	38	5
mo	0	0	0	0	7	1	5	0	0	0	12	1
ø	0	1	1	0	15	4	3	0	2	0	21	5
tte	0	0	0	0	7	1	0	0	0	0	7	1
de-wa	0	0	0	0	0	0	0	0	1	0	1	0
de-mo	0	0	0	0	2	0	0	0	0	0	2	0
tositewa	0	0	0	0	2	0	0	0	0	0	2	0
Total	9	1	6	1	78	15	32	3	10	0	135	20
Total	10		7		93		35		10		155	

Tables (11) and (12) indicate a strong tendency that locatives marked by non-locative particles such as *wa*, *ga*, *mo*, or zero-particle appear in the initial position, while locatives marked by locative particles such as *ni* and *de* appear in all positions, initial, medial and final in a clause.

2.3.1.8. SOVi (NP_1-NP_2-intransitive verb) construction

The SOVi construction is commonly known as a dative-subject or *ga*-marked object construction[18] with cognitive verbs, such as *dekiru* 'can do', *wakaru* 'understand', *ki-ni-naru* 'be worried'; and verbs with potential inflection such as *yurus-e-nai* (forgive-potential-negation) 'cannot forgive'. It might be more appropriate to describe the SOVi construction in terms of thematic roles, as the experiencer-stimulus-intransitive verb construction. Sentence (52) is an example of the SOVi construction with the verb *dekiru* 'can'.

(52)	SOVi						
14-6	tinpanzii	wa	hatuon	ga	dekinai-n-desu	yo	ne
	chimpanzee	TOP	pronunciation	NOM	can.NEG-NMLZ-COP	FP	FP
	'Chimpanzees cannot pronounce.'						

Table (13) The number of tokens in each constituent order variation of the SOVi construction

SOVi	Verb-Final				Verb-Medial				Verb-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	main	sub	
order	SOVi		OSVi				OViS						
2NPs	10	8	4	5			3	0			17	13	30
≥3NPs	1	0	1	0			0	0			2	0	2
Total	11	8	5	5			3	0			19	13	32

---

18 One of the arguments of treating the second noun phrase as the object is the agreement of the honorific verb to the first noun phrase. In (ix), the speaker uses the honorific verb form of *dekiru* 'can' to show respect to the first noun phrase, J, but not to the second noun phrase, English.

(ix)	J	wa	eigo	go	o-deki-ninaru
	J	TOP	English	NOM	HON-can-HON
	'J can speak English.'				

Three variations are found, SOVi, OSVi and OViS. SOVi and OSVi are found in both main and subordinate clauses.

(53) OSVi

16-12	sono	hun'iki	ga	zibun	ni	wa	yoku	wakar-anakatta
	that	atmosphere	NOM	self	DAT	TOP	well	understand-NEG.PST

'I didn't quite understand that mood/atmosphere.'

(54) OViS

17-21	eigo	dekiru	no	sono	hito	wa
	English	can	Q	that	person	TOP

'Can that person speak English?'

Despite the label "*ga*-marked object construction", the actually postnominal marking was more versatile. Table (14) shows the postnominal marking of subjects and objects in the SOVi construction.

Table (14) Postnominal marking of Subject and Object in SOVi construction

S	O	SOVi		O	S	OSVi				OViS	Total	
		Main	Sub			Main		Sub		Main	# of tokens	
		2NPs	2NPs			2NPs	≥3NPs	2NPs	≥3NPs	2NPs	2NPs	≥3NPs
wa	ga	3	2	ga	wa			1			6	0
ø	ga		1	ga	ø	1		2			0	0
mo	ga		1								1	0
demo	ga		1								1	0
				ga	niwa	1					0	0
				wa	ø			1			0	0
				wa	mo					1	0	0
wa	mo	1		mo	wa	1					2	0
ga	ø		2	ø	ga						2	0
wa	ø	2		ø	wa					1	0	0
ø	ø		1	ø	ø	1				1	0	0
				ø	ni			1			0	0
				ø	dewa		1				0	1
ga	tte	1									1	0
ø	tte	2									2	0
wa	wo	1									1	0
Total		10	8	Total		4	1	5	0	3	30	1
Total		10	8	Total		5		5		3	31	

Table (14) shows that the subject is mostly marked with *wa* or zero-particle, while the object is dominantly marked with *ga* and zero marking. Note that in the present data there is no token with double *ga* marking, such as in *hanako ga eigo ga dekiru* (Hanako-NOM English-NOM can) 'Hanako can speak English', which is often discussed as a grammatically possible sentence in the literature of Japanese syntax.

2.3.2. The construction with adjectives

In Japanese, adjectives form a class of the predicate. In the present data there were a number of tokens containing two noun phrases, NP_1 and NP_2, and an adjective, i.e. the NP_1-NP_2-adjective construction. The NP_1-NP_2-adjective construction with the subject occupying the NP_2 position is widely understood as the topic-subject construction as first illustrated by Mikami (1960), as in (55).

(55)	zoo	wa	hana	ga	nagai
	elephant	TOP	nose	NOM	long.
	'Talking about the elephant, his nose is long.'				

The topic in the topic-subject-adjective construction is a left-dislocated topicalized noun phrase, and the subject is the syntactic subject of the adjective. [19] In 54% of tokens of the topic-subject-adjective construction in the data, the syntactic unit of the subject and the adjective are cohesive denoting an idiomatic expression such as *se-ga-*

[19] The topic-subject-adjective construction is not unique to the adjective class of predicates. There are parallel constructions with verbs, such as in (a), which is the topic-subject-intransitive verb construction.

(a)

24-1	watasi	mo	sorosoro	pasupooto	ga	kireru-n-de
	1sg	also	soon	passport	NOM	expire-NMLZ-COP.TE
	'As for me, the passport will soon expire. --> My passport will soon expire too.'					

takai (stature-NOM-high) 'to be tall (in height)'. In addition, 70.3% of topics in the topic-subject-adjective construction were human. The postnominal marking of the subject is dominantly by *ga*, followed by zero-particle and *mo*, while the topic was dominantly marked with *wa*, followed by zero-particle, *ga*[20], *tte* and *mo*.

When the subject is animate and experiencer, the NP_1-NP_2-adjective construction is better described in terms of experiencer and stimulus, i.e. the experiencer (Ex)-stimulus (Sti)-adjective construction. Two order variations were found in the current data. Sentence (56) is an example of Ex-Sti-adjective order where the NP_1 is the subject (=experiencer), and Sentence (57) is an example of Sti-Ex-adjective order where the NP_2 is the subject (=experience).

(56)

7-20	uma	wa	on'na-no-hito	ni	yasasii
	horse	TOP	women	to	kind

'Horses are nice to women (while they are not nice to men).'

---

[20] The nominative *ga* may be used for non-subject NPs in the double nominal construction. One is the experiencer-stimulus-adjective construction, discussed in this section, such as in Example (a), another is the existential construction, discussed in 2.3.1.7., such as in Example (b), and the other is with intransitive cognitive verbs, discussed in 2.3.1.8., such as in Example (c). See Shibatani (1990) for the exhaustive study on the relation between the *ga*-marked nominal and the syntactic subject.

(a)	taroo	wa	hanako	ga	suki-da
	Taro	TOP	Hanako	NOM	like-IMP

'Taro likes Hanako.'

(b)	taroo	ni	musume	ga	iru/aru
	Taro	DAT	daughter	NOM	exist

'Taro has a daughter.'

(c)	taroo	ni	eigo	ga	dekiru
	Taro	DAT	English	NOM	can/be able to

'Taro can speak English.'

(57)
4-13 sore ga boku uresikatta
it NOM 1sg.male glad.PST
'I was happy about it.'

There were tokens where the NP_1 is the subject and stimulus, and the NP2 is the experiencer that is often marked by the particle *ni* (i.e. locative/dative). Sentence (58) is an example of this kind. I categorized as these tokens as the Ex-Sti-adjective construction with Sti-Ex-adjective order. There were no tokens where the NP_2 is the subject and stimulus.

(58)
5-12 kono koe o dasu koto ga watasi ni wa muzukasii
this voice TOP produce thing NOM 1sg LOC TOP difficult.NPST
'To produce this voice is difficult to me.'

Table (15) summarizes the NP_1-NP_2-adjective construction and the frequency of order variation.

Table (15) Order variations in the construction with adjective and the number of tokens

Adjective	Adjective-Final				Total		Total
	main	sub	main	sub	main	sub	
Order			T-S-Adj				
2NPs			23	6			29
≥3NPs			6	0			6
Total			29	6			35
Order	Ex-Sti-Adj		Sti-Ex-Adj				
2NPs	7	6	5	4	12	10	22
≥3NPs	0	0	0	0	0	0	0
Total	7	6	5	4	12	10	22
Total	7	6	34	10	12	10	57

2.3.3. The construction with complement copula

The complement copula is a syntactic structure in the form of a nominal adjective followed by a copula, as in *suki-da* (nominal adjective *suki* 'to like' and copula *da* 'to be') 'to like'. In the present data, there are a number of constructions involving two noun phrases and the complement copula, i.e. the NP_1-NP_2-complement copula construction. The relationship between NP_1 and NP_2 is best described as experiencer-stimulus. Sentence (60) has the experiencer-stimulus-complement copula construction, in which *watasi wa* is NP_1 = experiencer, *doobutu ga* is NP_2 = stimulus, and *suki-de* is complement copula.

(60) NP_1 (experiencer)-NP_2 (stimulus)-CC
21-18 watasi sugoku doobutu ga suki-de
1sg very animal NOM like-COP.TE
'I like animals very much.'

The complement copula found in the present data, other than *suki-da* 'to like', are *kirai-da* 'to dislike', *dame-da* 'to be no good', *hituyoo-da* 'to be necessary', and so forth. Table (16) displays the order variation and the frequency of each order.

Table (16) The number of tokens in each constituent order variation of the construction with complement copula

CC	CC-Final				CC-medial		CC-Initial		Total		Total
	main	sub	main	sub	main	sub	main	sub	main	sub	
Order	Ex-Sti-CC		Sti-Ex-CC		Sti-CC-Ex		CC-Ex-Sti				
2NPs	26	13	6	3	1	0	1	0	34	16	50
≥3NPs	3	0	0	0	0	0	0	0	3	0	3
Total	29	13	6	3	1	0	1	0	37	16	53

CC=complement copula, EX=experiencer, and Sti=stimulus

Four order variations were found, Ex-Sti-CC, Sti-Ex-CC, Sti-CC-Ex and CC-Ex-Sti.

(61) Sti-Ex-CC
15-17 H-san hizyooni boku wa suki-de
H-Mr. extremely 1sg.male TOP like-COP.TE
'I am extremely fond of Mr. H.'

(62) Sti-CC-Ex
10-6 zitensya yappa suki ne otoko-no-ko wa
bicycle as.I.said like FP male-GEN-child TOP
'Boys, as I said, like bicycles.'

(63) CC-Ex-Sti

1-30	taihen-datta	watasi	yomu	no
	tough-COP.PST	1sg	read	NMLZ

‘It was tough for me to read.’

2.4. Statistics

In this section, I will show various statistics of subject-object order (hereafter SO order) and object-subject order (hereafter OS order) in 2NPs, 3NPs, 4NPs and 5NPs nonsubordinate clauses. Tokens with 2NPs clauses contain a subject and an object (hereafter 2NPs construction), those with 3NPs clauses contain a subject, an object and another noun phrase (hereafter the 3NPs construction), those with 4NPs clauses contain a subject, an object and two noun phrases (hereafter 4NPs construction), and those with 5NPs clauses contain a subject, an object and three noun phrases (hereafter 5NPs construction). Some clauses in the ≥3NPs construction contain more than one NP of the same grammatical role which may or may not be coreferential. In my data, each NP in a clause is counted as one NP. Clauses (64) and (65) are examples of the 2NPs construction, (66) and (67) are of the 3NPs construction, (68) and (69) are of the 4NPs construction, and (70) and (71) are of the 5NPs construction.

(64) The 2NPs construction in SO order: SOV

7-14	anata	wa	o-mise	o	hira-ita
	2sg	TOP	POL-shop	ACC	open-PST

‘You opened up a shop.’

(65) The 2NPs construction in OS order: OSV

8-19	anata	no	koto	o	Y-sensei	ga	homete-iru
	2sg	GEN	matter	ACC	Y-teacher	NOM	applause.TE-be

‘Mr. Y applauds you.’

(66) The 3NPs construction in SO order: SLOV

19-1	sore	ga	noo	ni	sigeki	o	okutte
	it	NOM	brain	LOC	stimulus	ACC	send.TE

‘It sends stimulus to the brain.’

(67) The 3NPs construction in OS order: OSLV

22-13	sore	o	N	ga	gakuya	de	kiite-ita
	it	ACC	N	NOM	backstage	LOC	listen.TE-PPG.PST

‘N was listening to that backstage.’

(68) The 4NPs construction in SO order: LSLOV

12-6	sono	gakkoo	de	sensei	ga	soko	no	tokoro	ni
	its	school	LOC	teacher	NOM	there	GEN	place	LOC

sooiu	panhuretto	o	oite
like.that	pamphlet	ACC	place.TE

‘At that school a teacher placed such a pamphlet at that place.’

(69) The 4NPs construction in SO order: OXOSV

21-7a	kore	dooiu-no-de	kore	min’na	de	utusita	no
	this	what.kind-GEN-INS	this	everybody	INS	take	Q

‘On what occasion did you all take this picture?’

(70) The 5NPs construction in SO order

18-19b	boku	wa	soko	de	zyazu	konsaato	de
	1sg.male	TOP	there	LOC	jazz	concert	LOC

syooappu-suru	no	ni	isu	motte-iku
show.up-do	NMLZ	Purpose	chair	take-go

‘There, in the jazz concert, I bring a chair with me for a show.’

(71) The 5NPs construction in OS order: TOSOQV

6-10	saki-hodo	hon	no	daimei	o	watasi	Hituzi-no-uta
	a.while.ago	book	GEN	titile	ACC	1sg	Sheep-GEN-song

no	koto	o	Hituzi-no-mure	tte	itte-ta
GEN	matter	ACC	Sheep-GEN-flock	QT	say.TE-PPG.PST

'A while ago, I was saying the book title "A song of sheep" as "A flock of sheep".'

Table (17) displays the number of tokens with the 2NPs, 3NPs, 4NPs and 5NPs construction in different constituent order variations. The symbol "X" stands for any constituent other than the subject (S) and the object (O) in a clause. The number of tokens declines as the number of constituents increases. This phenomenon is not surprising based on the fact that the short-term memory capacity is limited.[21]

Table (17) The number of tokens with 2NPs, 3NPs 4NPs and 5NPs constructions in different constituent ordering

Construction	Constituent Ordering								Total
	SO				OS				
	S-initial			X-initial		O-initial			
2NPs	SOV					OSV			
# of tokens	130					37			167
3NPs	SOXV	SXOV		XSOV	XOSV	OSXV	OSOV	OXSV	
# of tokens	13	32		29	3	12	2	2	93
4NPs		SXOXV	SXXOV	XSXOV				OXOSV	
# of tokens		1	10	4				1	16
5NPs			SXXXOV	XXSXOV	XOSOXV XOSSSV				
# of tokens			2	1	2				5
Total	188			39		54			281
	222				59				

[21] In her study on the difference in word order between women and men's speech, Shibamoto (1985) states that the number of constituent strings longer than five elements in her corpus is too small as data, and that long strings tend to contain extra elements such as repetitions, backchannels, pauses, and repairs to the extent that it is questionable to consider them as coherent sentences. Relevant are discussions on the constraint on the volume of information per speech unit (e.g. Givón 1975, Chafe 1987, 1994, DuBois 1987).

Table (17) shows that in every category of 2NPs, 3NPs, 4NPs and 5NPs clause, the SO order occurs more often than the OS order. For example, for 2NPs clauses, SO order outnumbers OS order by 130 to 37, for 3NPs, OS outnumbers OS by 74 (13 SOXV, 32 SXOV, and 29 XSOV) to 19 (3 XOSV, 14 OSXV and 2 OXSV), for 4 NPs, SO outnumbers OS by 15 (1 SXOSV, 10 SXXOV and 4 XSXOV) to 1 (1 OXOSV), and for 5NPs, SO outnumbers OS by 3 (2 SXXXOV and 1 XXSXOV) to 2 (1 XOSOSV and 1 XOSSSV). The table also shows that subject-initial order (the total of 188 tokens) is more frequent than object-initial order (the total of 54 tokens) or X-initial order (the total of 39 tokens), which is the least frequent. Among the tokens of the 3NPs construction with SO order, the constituent order of the SXOV construction is dominant. Among the tokens of the 3NPs construction with OS order, the OSXV order is the most frequent[22]. Out data includes tokens in the ≥3NPs construction that contain more than one NPs of the same grammatical role. For example, two Os in OS order with 4NPs construction, OXOSV, are coreferential. As far as referents are concerned, the OXOSV is reduced to OXSV. In OS order with 5NPs construction, two Os in XOSOXV and three Ss in XOSSSV are coreferential. If we consider only referents, the XOSOXV is XOSXV, and the XOSSSV is XOSV.[23]

[22] If one assumes syntactic movement, OSXV order is understood as a result of preposing the O in SXOV or SOXV. Similarly, the XSOV order, which is the most dominant order in X-initial ordering, can be considered as a result of fronting the X in SXOV or SOXV, and XOSV order as a result of fronting the X in OSXV or OXSV. Moreover, the XSXOV order in the 4NPs construction can be derived from SXXOV, since the sequence of XX is constantly observed in both S-initial and O-initial order.

[23] In addition, there are tokens containing two locatives or temporals, which may or may not be coreferential. We will see more data on the ≥3NPs construction in Chapter 4.

Tables (18) and (19) display the number of 2NPs, 3NPs, 4NPs and 5NPs tokens with SO order in each combination of postnominal markings on the subject and the object. Table (18) shows the data grouped by the particle of the subject, and Table (19) by the particle of the object. There are six different subject markings: the nominative *ga*, the adverbial particle *wa*, the zero-particle *ø*, the adverbial particle *mo* denoting 'also', the case particle *de* that is mostly in the form of *min'na de* 'everybody', and the genitive *no*. The genitive *no* is commonly employed to mark the subject instead of the nominative *ga* in relative clauses such as *watasi-no-yonda-hon* (1sg-GEN-read-book) 'the book that I read' for *watasi-ga-yonda-hon* (1sg-NOM-read-book) 'the book that I read'. For the object marking, there are four particles: the accusative *o*, the adverbial particle *wa*, the zero-particle *ø*, and the adverbial particle *mo*.

Table (18): The number of 2NPs, 3NPs, 4NPs and 5NPs tokens in SO order, sorted by the particle of the subject

		2NPs		≥3NPs							
Order variations		**SOV**		**SOXV**		**SXOV (SXOXV) (SXXOV) (SXXXOV)**		**XOSV (XSXOV) (XXSXOV)**		Total	
S	**O**	#	%	#	%	#	%	#	%	#	%
ga	wo	12	9.2	5	38.5	9, (3)	28.3	6, (2)	23.5	37	16.6
ga	wa	1	0.8	1	7.7					2	0.9
ga	ø	21	16.2	1	7.7	2	4.3	1, (1)	5.9	26	11.7
ga	mo	1	0.8							1	0.4
sub-total		*35*	*26.9*	*7*	*53.8*	*11 (3)*	*32.6*	*7 (3)*	*29.4*	*66*	*29.6*
wa	wo	26	20.0	2	15.4	6, (1)	15.2	2	5.9	37	16.6
wa	wa	3	2.3	1	7.7					4	1.8
wa	ø	15	11.5			3, (3)	13.0	2	5.9	23	10.3
wa	mo	3	2.3	1	7.7			1	2.9	5	2.2
sub-total		*47*	*36.2*	*4*	*30.8*	*9 (4)*	*28.3*	*5*	*14.7*	*69*	*30.9*
ø	wo	9	6.9			1, (1)	4.3	7	20.6	18	8.1
ø	wa	2	1.5							2	0.9
ø	ø	16	12.3	1	7.7	5, (4)	19.6	7, (2)	26.5	35	15.7
ø	mo	3	2.3			1	2.2			4	1.8
sub-total		*30*	*23.1*	*1*	*7.7*	*7 (5)*	*26.1*	*14 (2)*	*47.1*	*59*	*26.5*
mo	wo	5	3.8			1	2.2			6	2.7
mo	wa			1	7.7	1	2.2			2	0.9
mo	ø	5	3.8			1, (1)	4.3	2	5.9	9	4.0
mo	mo									0	0.0
sub-total		*10*	*7.7*	*1*	*7.7*	*3 (1)*	*8.7*	*2*	*5.9*	*17*	*7.6*
de	wo	3	2.3			1	2.2			4	1.8
de	wa									0	0.0
de	ø	5	3.8					1	2.9	6	2.7
de	mo									0	0.0
sub-total		*8*	*6.2*			*1*	*2.2*	*1*	*2.9*	*10*	*4.8*
no	ø					1	2.2			1	0.4
sub-total						*1*	*2.2*			*1*	*0.4*
Total		130	100	13	100	32 (13)	100	29 (5)	100	222	99

* The number in parentheses is the number of 4NPs or 5NPs tokens. S=subject, O=direct object, X=other constituent than subject and direct object.

Table (19) The number of 2NPs, 3NPs, 4NPs and 5NPs tokens in SO order, sorted by the particle of the direct object

Order variations		2NPs **SOV**		≥3NPs **SOXV**		**SXOV (SXOXV) (SXXOV) (SXXXOV)**		**XOSV (XSXOV) (XXSXOV)**		Total	
S	**O**	#	%	#	%	#	%	#	%	#	%
ga	wo	12	9.2	5	38.5	9, (3)	28.3	6, (2)	23.5	37	16.6
wa	wo	26	20.0	2	15.4	6, (1)	15.2	2	5.9	37	16.6
ø	wo	9	6.9			1, (1)	4.3	7	20.6	18	8.1
mo	wo	5	3.8			1	2.2			6	2.7
de	wo	3	2.3			1	2.2			4	1.8
sub-total		*55*	42.3	*7*	*53.8*	*18 (5)*	*52.2*	*15 (2)*	*50*	*102*	*45.7*
ga	wa	1	0.8	1	7.7					2	0.9
wa	wa	3	2.3	1	7.7					4	1.8
ø	wa	2	1.5							2	0.9
mo	wa			1	7.7	1	2.2			2	0.9
de	wa									0	0.0
sub-total		*6*	4.6	*3*	*23.1*	*1*	2.2	*0*	*0.0*	*10*	*4.5*
ga	ø	21	16.2	1	7.7	2	4.3	1, (1)	5.9	26	11.7
wa	ø	15	11.5			3, (3)	13.0	2	5.9	23	10.3
ø	ø	16	12.3	1	7.7	5, (4)	19.6	7, (2)	26.5	35	15.7
mo	ø	5	3.8			1, (1)	4.3	2	5.9	9	4.0
de	ø	5	3.8					1	2.9	6	2.7
no	ø					1	2.2			1	0.4
sub-total		*62*	47.7	*2*	*15.4*	*12 (8)*	*43.5*	*13 (3)*	*47.1*	*100*	*44.8*
ga	mo	1	0.8							1	0.4
wa	mo	3	2.3	1	7.7			1	2.9	5	2.2
ø	mo	3	2.3			1	2.2			4	1.8
mo	mo									0	0.0
de	mo									0	0.0
sub-total		*7*	5.4	*1*	*7.7*	*1*	*2.2*	*1*	*2.9*	*10*	*4.5*
Total		130	100	13	100	32 (13)	100	29 (5)	100	222	100

* The number in the parenthesis is the number of 4NPs or 5NPs tokens. S=subject, O=direct object, X=other cases than subject and object.

According to Table (18), SO tokens are evenly distributed (with a slight inclination to the subject marking by the particle *wa*) among subjects marked with the nominative *ga*, adverbial particle *wa* and zero-particle *ø* in both 2NPs and ≥3NPs constructions. On the other hand, Table (19) shows that object tokens with 2NPs construction in SO order are mostly marked with the accusative *o* (42.3%) or by zero-particle (47.7%). The same tendency is valid with ≥3NPs tokens in SO order, except that the object marking by the accusative *o* is much greater than by zero-particle *ø* in the ≥3NPs constructions. The SO tokens with objects marked with the adverbial particles *wa* or *mo* are scarce. There is no token of S-*mo* O-*mo*, S-*de* O-*wa*, or S-*de* O-*mo* sequences.

Tables (20) and (21) show the number of tokens with OS order for each particle combination. The data is displayed by the particle of the object in Table (20) and by the particle of the subject in Table (21).

Table (20): The number of 2NPs, 3NPs, 4NPs and 5NPs tokens in OS order, sorted by the particle of the direct object

		2NPs		≥3NPs							
Order variations		**OSV**		**OSXV OSOV**		**OXSV (OXOSV)**		**XOSV (XOSXXV)**		Total	
O	**S**	#	%	#	%	#	%	#	%	#	%
wo	ga	9	24.3	2	14.3			2	40.0	13	22.0
wo	wa	3	8.1	4	28.6					7	11.9
wo	ø	5	13.5					(1)	20.0	6	10.2
wo	mo	1	2.7	2	14.3					3	5.1
wo	de			1	7.1	1	33.3			2	3.4
sub-total		*18*	*48.6*	*9*	*64.3*	*1*	*33.3*	*2 (1)*	*60.0*	*31*	*52.5*
wa	ga	3	8.1	1	7.1			(1)	20.0	5	8.5
wa	wa			1	7.1					1	1.7
wa	ø	4	10.8	1	7.1					5	8.5
wa	mo	2	5.4							2	3.4
sub-total		*9*	*24.3*	*3*	*21.4*			*-1*	*20.0*	*13*	*22.0*
ø	ga	2	5.4			1	33.3			3	5.1
ø	wa										
ø	ø	5	13.5					1	20.0	6	10.2
ø	mo	2	5.4							2	3.4
ø	de			1	7.1	(1)	33.3			2	3.4
sub-total		*9*	*24.3*	*1*	*7.1*	*1 (1)*	*66.7*	*1*	*20.0*	*13*	*22.0*
mo	ga			1	7.1					1	1.7
mo	wa										
mo	ø										
mo	mo										
mo	de	1	2.7							1	1.7
sub-total		*1*	*2.7*	*1*	*7.1*					*2*	*3.4*
Total		37	100	14	100	2 (1)	100	3 (2)	100	59	100

* The number in parentheses is the umber of 4NPs or 5NPs tokens. S=subject, O=direct object, X=other constituent than subject and direct object.

According to Table (20), nearly half of objects in OS order are marked with *o* (i.e. 48.6%). The percentage of object marking by *o* is much higher in the 3NPs construction than in the 2NPs construction; the object in thirteen out of twenty-two tokens in the ≥3NPs constructions is marked with *o*. In OS order, the zero marking of objects and the *wa*-marking of objects are comparable. The object marking by the particle *mo* is significantly less frequent than other postnominal markings.

Table (21): The number of 2NPs, 3NPs, 4NPs and 5NPs in OS order, sorted by the particle of the subject

		2NPs		≥3NPs							
Order variations		**OSV**		**OSXV OSOV**		**OXSV (OXOSV)**		**XOSV (XOSXXV)**		Total	
O	**S**	#	%	#	%	#	%	#	%	#	%
wo	ga	9	24.3	2	14.3			2	40	13	22.0
wa	ga	3	8.1	1	7.1			(1)	20.0	5	8.5
ø	ga	2	5.4			1	33.3			3	5.1
mo	ga			1	7.1					1	1.7
sub-total		*14*	*37.8*	*4*	*28.6*	*1*	*33.3*	*2 (1)*	*60.0*	*22*	*37.3*
wo	wa	3	8.1	4	28.6					7	11.9
wa	wa			1	7.1				`	1	1.7
ø	wa										
mo	wa										
sub-total		*3*	*8.1*	*5*	*35.7*					*8*	*13.6*
wo	ø	5	13.5					(1)	20.0	6	10.2
wa	ø	4	10.8	1	7.1					5	8.5
ø	ø	5	13.5					1	20.0	6	10.2
mo	ø										
sub-total		*14*	*37.8*	*1*	*7.1*			*1 (1)*	*40.0*	*17*	*28.8*
wo	mo	1	2.7	2	14.3					3	5.1
wa	mo	2	5.4							2	3.4
ø	mo	2	5.4							2	3.4
mo	mo										
sub-total		*5*	*13.5*	*2*	*14.3*					*7*	*11.9*
wo	de			1	7.1	1	33.3			2	3.4
wa	de										
ø	de			1	7.1	(1)	33.3			2	3.4
mo	de	1	2.7							1	1.7
sub-total		*1*	*2.7*	*2*	*14.3*	*1 (1)*	*66.7*			*5*	*8.5*
Total		37	100	14	100	2 (1)	100	3 (2)	100	59	100

* The number in the parentheses is the number of 4NPs or 5NPs tokens. S=Subject, O=Direct Object, X=other constituent than subject and direct object

Table (21) shows that in OS order the subject marking by the nominative *ga* is the most frequent, while as far as the 2NPs construction is concerned, the zero marking of the subject is as frequent as the *ga*-marking. It is also observed that *wa*-marked subjects only occur with objects marked with the accusative *o* or the adverbial particle *wa*, not with other postnominal markings. There is no token in which both subject and object are marked with the particle *mo*, while 15 % of tokens are *ø-ø* marking (i.e. the co-occurring subject and object are both zero-marked).

Table (18) shows that in SO order the subject marking by the particle *wa* is more frequent than *ga*-marking or zero marking (30.9% for *wa*-marking, 30.0% for *ga*-marking and 26.5% for zero marking). Table (21) shows that the *wa*-marking of subjects in OS order is less frequent than *ga*-marking or zero marking (11.9% for *wa*-marking, 37.3% for *ga*-marking and 30.5% for zero marking). As for the *wa*-marking of objects, the comparison of Tables (19) and (20) tells us that the object marking by *wa* is much more common when the object is an initial constituent than when the object is not in initial position. This result suggests a relation between the *wa*-marking and clause-initial constituents.

2.4.1. The 2NPs construction

Table (22) summarizes the variety of 2NPs constructions and their constituent order variations which were presented in the previous section. The construction with a subject (S) and a direct object (O) is the most frequently used construction in the present data. The number refers to the number of tokens of each order variation in main and subordinate clauses. The verb-final order has two variations, S-initial and S-medial, and the verb-medial order has two variations, S-initial and S-final. The verb-initial order has only one variation, S-medial order.

Table (22) Constituent order variations and the number of tokens in the 2NPs construction

Verb-Final				Verb-Medial				Verb-Initial		# of tokens	
SXV		XSV		SVX		XVS		VSX		Total	
main	sub	main	sub	main	sub	main	sub	main	sub	main	sub
SOV		OSV		SVO		OVS		VSO		311	
130	99	37	17	13	2	10	2	1		191	120
SQV		QSV		SVQ		QVS				251	
147	44	33	13	2	0	10	2			192	59
SDV		DSV						VSD		32	
15	12	1	3					1	0	17	15
SCVi						CViS				83	
48	32					3	0			51	32
SLVm		LSVm		SVmL						156	
73	46	22	8	5	2					100	56
SLVi		LSVi				LViS				32	
11	7	6	7			1	0			18	14
SLVex (iru)		LSVex (iru)		SVexL (iru)		LVexS (iru)				48	
10	6	24	6	1	0	1	0			36	12
SLVex (aru)		LSVex (aru)		SVexL (aru)						155	
10	7	93	35	10	0					113	42
SOVi		OSVi				OViS				32	
11	8	5	5			3	0			19	13
		T-S-Adj		S-Adj-T						37	
		29	6	2	0					31	6
Ex-Sti-Adj		Sti-Ex-Adj								22	
7	6	5	4							12	10
Ex-Sti-CC		Sti-Ex-CC				Sti-CC-Ex		CC-Ex-Sti		52	
28	13	6	3			1	0	1	0	36	16
Total		Total		Total		Total		Total		Total	
770		368		37		33		3		1211	
490	280	261	107	33	4	29	4	3	0	816	395

Table (22) shows that among the verb-final orders, SXV (where the X is non-subject constituent) is most frequent, except for the construction with the existential verb *aru* 'exist' and *iru* 'be', and the Topic-Subject-Adjective construction. It indicates that the SXV order is dominant in terms of frequency. In the construction with the existential verb, XSV (where the X is locative) is the dominant order. In the verb-medial construction, the subject-initial SVX order is slightly more frequent than the subject-final XVS order, except the constructions involving intransitive verbs with quotative, complement and *ga*-marked objects. In the verb-initial construction, only VSX order is found in the present data.

2.4.2. The SOV construction: SOV vs. OSV orders

In this section, I will examine the statistics of SOV order and OSV order in the SOV construction. Example (72), which previously appeared as (64), is a token of SOV order, and (73), which previously appeared as (65), is a token of OSV order.

(72) SOV

7-14	anata	wa	o-mise	o	hira-ita
	2sg	TOP	POL-shop	ACC	open-PST

‘You opened up a shop.’

(73) OSV

8-19	anata	no	koto	o	Y-sensei	ga	home-te-iru
	2sg	GEN	matter	ACC	Y-teacher	NOM	applause-TE-be

‘Mr. Y applauds you.’

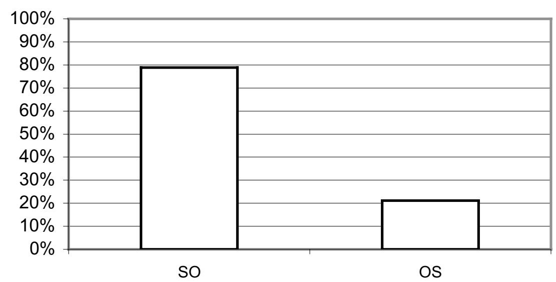
Table (23) shows the order variations of the 2NPs construction involving a subject and a direct object. The number in Table (23) was obtained by combining the number of tokens in main and subordinate clauses, presented in Tables (4) in section 2.3.1.1. and (22) in 2.4.1.

Table (23) The number of tokens in each constituent order variation of the 2NPs construction with a subject and an object.

Verb-Final				Verb-Medial				Verb-Initial		Total	
SOV		OSV		SVO		OVS		VSO			
#	%	#	%	#	%	#	%	#	%	#	%
229	73.6	54	17.4	15	4.8	12	3.9	1	0.3	311	100

As shown in Table (23), the SOV order (73.6%) is the most frequent order among the 2NPs clauses containing a subject and a direct object, followed by OSV (17.4%), OVS (4.8%), SVO (4%), and then VSO (0.3%). With respect to SO vs. OS orders, there are a total of 245 tokens of SO order (229 tokens of SOV, 15 tokens of SVO and one token of VSO) and a total of 66 tokens of OS order (54 tokens of OSV and 12 tokens of OVS). This result is displayed in Chart (1), indicating that SO order is more frequent (78.8%, i.e. 245/311 tokens) than OS order (21.2%, i.e. 66/311 tokens).

Chart (1) SO order vs. OS order



With respect to the frequency of the verb-final, verb-medial and verb-initial orders, the verb-final order prevails (i.e. 91.0%) as shown in Table (24).

Table (24) The frequency of occurrence: Verb-final vs. Verb-medial vs. Verb-initial

Orders	Verb-final			Verb-medial			Verb-initial			Total	
		#	%		#	%		#	%	#	%
SO	SOV	229	93.5%	SVO	15	6.1%	VSO	1	0.4%	245	100%
OS	OSV	54	81.8%	OVS	12	18.2%				66	100%
Total		283	91.0%		27	8.7%		1	0.3%	311	100%

Table (25) shows the difference in frequency between SO and OS orders.

Table (25) The frequency of occurrence: SO order vs OS order

Orders	Verb-final			Verb-medial			Verb-initial			Total	
		#	%		#	%		#	%	#	%
SO	SOV	229	80.9%	SVO	15	55.6%	VSO	1	100%	245	78.8%
OS	OSV	54	19.1%	OVS	12	44.4%				66	21.2%
Total		283	100%		27	100%		1	100%	311	100%

It shows that the difference in frequency between SO and OS orders is greater in verb-final orders (i.e. SOV and OSV) than the verb-medial orders (i.e. OVS and SVO). For example, in verb-final orders, the SO order, SOV, is more than four times more frequent

(80.9%) than the OS order, OSV (19.1%), while in verb-medial orders, the difference in frequency between SO order, SVO (55.6%) and OS order, OVS (44.4%) is smaller.

The above results indicate that with verb-final order, SOV order is more frequent than OSV order, but there is no significant difference in frequency of occurrence as to verb-medial OVS and SVO orders.

Herring and Paolillo (1995:178) measured the verb-finality in Sinhala and Tamil by counting the number of finite verb-final, verb-non-final and verbless clauses in their corpus. The results of their measurement are summarized in Tables (26) and (27).

Table (26) Herring and Paolillo's measurement of verb-finality for Sinhala

	X(V)	VX	verbless
% all finite clauses	66.5%	21.4%	12.0%
% verbal clauses only	75.6%	24.4%	-

Table (27) Herring and Paolillo's measurement of verb-finality for Tamil

	X(V)	VX	verbless
% all finite clauses	83.0%	4.0%	13.0%
% verbal clauses only	95.4%	4.6%	-

Since the data in this study only contain those tokens that have both subject and object in a clause, the current measurement of verb-finality for spoken Japanese, strictly speaking, cannot be compared to Herring and Paolillo's measurement in Sinhala and Tamil. Having said that, the 91.0% of verb-finality in the current data indicates rather rigid verb-finality as in Tamil rather than a non-rigidly verb-final language as in Sinhala.

Lastly, Table (28) shows that the number of main clauses is slightly higher than that of subordinate clauses in the present data. It also shows that the frequency difference between the main and subordinate clauses increases as the occurrence of the word order decreases. The difference is the least, for instance, in SOV order, the most frequent word order, and it is the most in VSO order, the least frequent word order.

Table (28) The number of tokens in main vs. subordinate clauses.

Order variations	Main		Subordinate		Total	
	#	%	#	%	#	%
SOV	130	56.8%	99	43.2%	229	100%
OSV	37	68.5%	17	31.5%	54	100%
OVS	10	83.3%	2	16.7%	12	100%
SVO	13	86.7%	2	13.3%	15	100%
VSO	1	100%	0	0%	1	100%
Total	191	61.4%	120	38.6%	311	100%

2.4.2.1. SOV vs. OSV: RD

Table (29) shows the token distribution of SOV order with respect to the referential distance (RD). The data is only concerned with nonsubordinate clauses.

Table (29) RD of Subject and Object in SOV order

	Subject		Object	
RD	# of tokens	%	# of Tokens	%
1	27	20.8%	16	12.3%
2	11	8.5%	3	2.3%
3	11	8.5%	3	2.3%
4 ≤ 20	35	26.9%	22	16.9%
20+	18	13.8%	12	9.2%
FM	28	21.5%	74	56.9%
Total	130	100%	130	100%

Table (29)[24] shows the concentration of objects on RD=FM while subjects spread over the categories of RD counts. The table also shows that a higher percentage of subjects than objects have an RD=1 (20.8% vs. 12.3%) and a higher percentage of objects than

---

[24] Mean RD for the subject in SOV is 10.3, and that for the object in SOV is 15.4. RD=20+ was counted as RD=20 and FM as RD=21.

subjects are first mentions (56.9% vs. 21.5%). The '74' first mention objects is noteworthy.

Table (30) shows the token distribution of OSV order with respect to RD.

Table (30) RD of Object and Subject in OSV order

	Object		Subject	
RD	# of tokens	%	# of Tokens	%
1	11	29.7%	4	10.8%
2	1	2.7%	4	10.8%
3	2	5.4%	0	0.0%
4 ≤ 20	2	5.4%	5	13.5%
20+	1	2.7%	4	10.8%
FM	20	54.1%	20	54.1%
Total	37	100%	37	100%

Table (30)[25] shows that the proportions of objects and subjects that are first mentions are equal (54.15% for both objects and subjects). In OSV clauses, in contrast to SOV clauses, a greater proportion of objects have an RD of 1 compared to subjects (29.7% vs. 12.3%). Non-first mention objects tend to have lower RD than higher RD, while non-first mention subjects split equally between higher and lower RD.

The subject in SOV clauses and the object in OSV clauses are the entities introduced earlier in a clause (hereafter Early-constituents), and the object in SOV clauses and the subject in OSV clauses are those introduced later in a clause (hereafter Late-constituents). With respect to the Early-constituents vs. Late-constituents, Tables (29) and (30) indicate a tendency for an entity that has an RD of 1 to be introduced earlier in a clause. The Early-constituent objects show a tendency to be either RD=1 or RD=FM.

[25] Mean RD for the object in OSV is 12.9, and that for the subject in OSV is 14.9. RD=20+ was counted as RD=20 and FM as RD=21.

Since more than 50% of the Late-constituent objects are first mentions, the property of first mentions might pertain to the object regardless of its position in a clause. As to the Late-constituents, both Late-constituent objects and subjects show a similar pattern of token distribution with respect to the RD, i.e. spreading across each grid of the RD count. In sum, the comparison of RD between subjects and objects in SOV and OSV orders suggests that the property of Early-constituents is more influential than the Late-constituents upon the choice of constituent ordering. This observation conforms to the claim, which was discussed in Chapter 1, that more recently mentioned entities tend to be introduced earlier in a string.

2.4.2.2. SOV vs. OSV: TP

Table (31) shows the number of subject and object tokens with respect to the Topic persistence (TP) in SOV order.

Table (31) TP of Subject and Object in SOV order

	Subject		Object	
TP	# of tokens	%	# of Tokens	%
0	24	18.5%	64	49.2%
1	31	23.8%	27	20.8%
2	19	14.6%	12	9.2%
3	20	15.4%	12	9.2%
≥4	36	27.7%	15	11.5%
Total	130	100%	130	100%

Table (31)[26] indicates the concentration of object tokens with TP=0, nonpersistent (i.e. the referent does not persist in the following discourse), and as the TP increases, the number of object tokens decreases. On the other hand, nearly 80% of subjects are

---

[26] Mean TP for the subject in SOV is 2.49, and that for the object in SOV is 1.22.

persistent (i.e. TP ≥ 1, the referent persist in the following discourse). That is to say, a higher percentage of objects than subjects are not persistent (49.2% vs. 18.5%), and more number of subjects than objects are persistent (81.5% vs. 50.8%).

Table (32)[27] shows the number of object and subject tokens in OSV order with respect to the TP.

Table (32) TP of Object and Subject in OSV order

	Object		Subject	
TP	# of tokens	%	# of Tokens	%
0	13	35.1%	12	32.4%
1	14	37.8%	9	24.3%
2	3	8.1%	5	13.5%
3	4	10.8%	5	13.5%
4≥	3	8.1%	6	16.2%
Total	37	100%	37	100%

In contrast to SOV order clauses, the proportions of subjects and objects in OSV clauses that are nonpersistent are comparable (35.1% vs. 32.4%). Similar to SOV clauses, the proportion of objects that are not persistent is greater than that of nonpersistent subjects in OSV clauses. That is to say, subjects tend to be more persistent than objects.

The comparison of Tables (31) and (32) tells us that when the referent is persistent, it is most likely the Early-constituent (i.e. the subject in SOV order) than the Late-constituent (i.e. the object in SOV order) in SOV order, but it is not the case in OSV order. Moreover, when the referent does not persist, it is most likely the object than the subject, whether it is the Early-constituent or the Late-constituents. This suggests that the

27 Mean TP for the object in OSV is 1.38, and that for the subject in OSV is 1.89.

property of cataphoric persistence is more significant in the default order (i.e. SOV order in Japanese) than in non-default order (i.e. OSV order in Japanese).

2.4.2.3. SOV vs. OSV: SW

Table (33) shows the syntactic weight (SW) of subjects and objects in SOV order.

Table (33) SW of subjects and objects in SOV order

	Subject		Object	
SW	# of tokens	%	# of Tokens	%
1	14	10.8%	30	23.1%
2	68	52.3%	41	31.5%
3	18	13.8%	17	13.1%
4	13	10.0%	11	8.5%
5	5	3.8%	8	6.2%
6	4	3.1%	8	6.2%
7	4	3.1%	2	1.5%
8	3	2.3%	3	2.3%
9	0	0.0%	3	2.3%
≥10	1	0.8%	7	5.4%
Total	130	100%	130	100%

Table (33) shows a similar token distribution pattern for subjects and objects with respect to the SW count. For example, the majority of tokens have an SW of one to four counts (i.e. 86.9%, i.e. 113/130 tokens, for subjects and 76.2%, i.e. 99/130 tokens, for objects), and when the SW is more than five, the number of tokens drastically drops. The table also shows that a higher percentage of objects than subject have an SW that are larger than five counts (13.1%, i.e. 17/130 tokens, for subjects and 23.8%, i.e. 31/130 tokens, for objects). Table (34) shows the syntactic weight (SW) of objects and subjects in OSV order.

Table (34) SW of objects and subjects in OSV order

	Object		Subject	
SW	# of tokens	%	# of Tokens	%
1	4	10.8%	10	27.0%
2	12	32.4%	13	35.1%
3	3	8.1%	3	8.1%
4	2	5.4%	6	16.2%
5	4	10.8%	1	2.7%
6	2	5.4%	1	2.7%
7	1	2.7%	0	0.0%
8	1	2.7%	0	0.0%
9	1	2.7%	1	2.7%
≥10	7	18.9%	2	5.4%
Total	37	100%	37	100%

Table (34) shows that a higher percentage of objects than subjects have an SW of higher than five counts (16/37 tokens = 43.2 % for objects and 5/37 tokens = 13.5% for subjects). The table also shows that a higher percentage of subjects than objects have an SW of one to four counts (32/37 tokens = 86.5% for subjects and 21/37 tokens = 56.8% for objects).

The comparison of Tables (33) and (34) tells us that a higher percentage of objects, regardless of their position in a clause, than subjects are syntactically heavier, and that a higher percentage of subjects, regardless of their position in a clause, than objects are syntactically lighter.

2.4.3. The choice of postnominal markings in the SOV construction

In the following sections, I will discuss the possible dependency of postnominal markings on the count of RD, TP and SW. There are five subject markings and four object markings found in the present data. Subjects are marked with the case particle *ga* (nominative), zero-particle *ø*, the adverbial particles *wa* and *mo*[28] and the case particle *de*.[29] Object markings are by the case particle *o* (accusative), zero-particle *ø*, and the adverbial particles *wa*, and *mo*. What we are looking at here is the speaker's choice over the postnominal markings. For instance, when a speaker is about to produce an SOV order clause, the RD of the subject is already determined, and the speaker have to choose a particular postnominal marker for the subject. If the speaker chooses a particular postnominal marker for a particular value of RD. there is a correlation between the choice of postnominal markings and the count of RD.

2.4.3.1. RD and postnominal markings

In this section, I will examine if there is any correlation between the choice of postnominal markings and the RD of subjects and objects. In 2.4.3.1.1. I will study the correlation in nonsubordinate clauses with SOV order, and in 2.4.3.1.2. that in nonsubordinate clauses with OSV order.

2.4.3.1.1. RD and the choice of postnominal markings: SOV order

Table (35) show the token distribution of subjects in SOV order in terms of RD, with reference to different postnominal markings.

---

[28] The particles *wa* and *mo* are labeled as discourse particles in the Japanese Language Engine (Kameyama 1995). Subjects marked with the particle *mo* tend to be higher in RD in my data; it is implied by the presence of presupposition indicated by the use of *mo*.

[29] In traditional Japanese grammar, the particle *de* is categorized as a case particle. In the Japanese Language Engine, it is classified as a semantic particle. (Kameyama 1995). The case particle *de* is used to mark quantifier subjects as in *san'nin-de* (three.people-by) 'with three people', or in the construction with the comitative marker *to* as in *ane-to-watasi-de* (my.elder.sister-COM-1sg-by) 'I with my sister'. The former usage tends to show the lower RD and the latter usage is linked to first mentions in my data.

Table (35) RD of Subject and Subject Marking in SOV order

RD	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	3	%	#	%	#	%
1	5	19.2	11	42.3	9	34.6	0	0.0	1	3.8	26	100
2	4	36.4	1	9.1	3	27.3	0	0.0	3	27.3	11	100
3	1	9.1	7	63.6	3	27.3	0	0.0	0	0.0	11	100
4 ≤ 20	6	16.7	19	52.8	5	13.9	6	16.7	0	0.0	36	100
20+	5	27.8	7	38.9	4	22.2	1	5.6	1	5.6	18	100
FM	14	50.0	2	7.1	6	21.4	3	10.7	3	10.7	28	100
Total	35	26.9	47	36.2	30	23.1	10	7.7	8	6.2	130	100

Table (35) shows that when the subject is a first mention, the speaker tends to use the nominative *ga*, and that when the RD of the subject is between 3 to 20+ (i.e. previously mentioned, but in decay), the particle *wa* tend to be used. When the subject has an RD of 1 and 2, there was no preference for a particular postnominal particle. Table (36) compares the use of the nominative *ga* and the particles other than *ga* for the first mention entities and for the entities previously mentioned (PM). The numbers in Table (36) were pooled from Table (35) to see whether the use of the nominative *ga* for the first mention entities has any significance. The result of Fisher Exact Test (p=0.0034)[30] indicates that the correlation between the use of *ga* and first mentions is statistically significant.

Table (36) The use of the nominative ga for first mention entities

	ga		not ga		Total	
	#	%	#	%	#	%
PM	21	20.6	81	79.4	102	100
FM	14	50.0	14	50.0	28	100
Total	21	16.2	95	73.1	130	89

[30] Fisher Exact Test calculates the probability of the occurrence by chance. Throughout this thesis, the p-value is calculated by Fisher Exact Test in http://www.matforsk.no/ola/fisher.htm. The significance level is set at p=0.05

Similarly, the numbers in Table (35) were pooled together to see if there is any significant correlation between the use of the particle *wa* and the previous mention. The result is shown in Table (37). The result of Fisher Exact Test (p=0.0002) indicates the statistical significance of the correlation.

Table (37) The use of the particle *wa* for first mention entities

	wa		not wa		Total	
	#	%	#	%	#	%
PM	45	44.1	57	55.9	102	100
FM	2	7.1	26	92.9	28	100
Total	45	34.6	83	63.8	130	98

In Charts (2) and (3), the distribution of tokens by different subject markings, shown in Table (35), is organized into two categories, previously mentioned entities (PM) and first mentions (FM). Charts (2) and (3) demonstrate clearly the correlation between the first mention subject and *ga*-marking and between the previously mentioned subject and *wa*-marking.

Chart (2) Distribution of tokens when the subject in SOV is previously mentioned

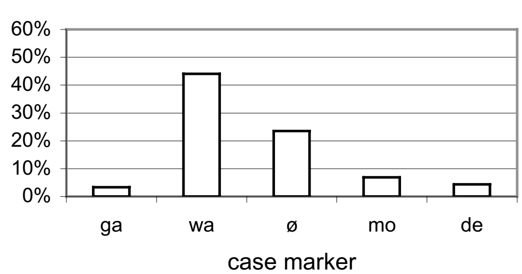


Chart (3) Distribution of tokens when the subject in SOV is first mention

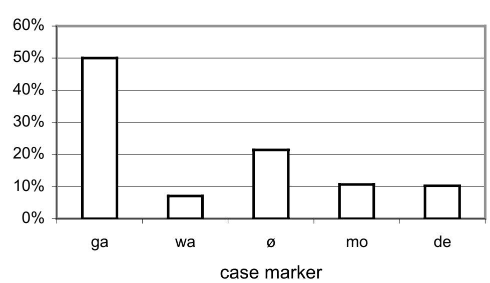


Table (38) shows the token distribution of objects in SOV order in terms of RD, with reference to different postnominal markings.

Table (38) RD of Object and Object Marking in SOV order

RD	Object wo		Object wa		Object ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
1	8	50.0	0	0.0	8	50.0	0	0.0	16	100
2	2	66.7	0	0.0	1	33.3	0	0.0	3	100
3	2	66.7	0	0.0	1	33.3	0	0.0	3	100
4 ≤ 20	10	45.5	1	4.5	11	50.0	0	0.0	22	100
20+	4	33.3	1	8.3	7	58.3	0	0.0	12	100
FM	29	39.2	4	5.4	34	45.9	7	9.5	74	100
Total	55	42.3	6	4.6	62	47.7	7	5.4	130	100

Table (38) shows no evidence that the choice of postnominal marker is dependent on the RD of the object. First, the accusative *o* and zero-particle *ø* mark 90% of objects in SOV order, and regardless of the RD of the object, the accusative *o* and zero-particle *ø* are the postnominal markings for the majority of objects. In order to see whether there is any difference between previously mentioned objects and first mention objects as to the postnominal markings, the numbers in Table (38) were pooled into two categories, when the object is previously mentioned (PM) and when the object is a first mention (FM). The result is shown in Table (39).

Table (39) Postnominal marking on the object in SOV order when O=PM vs. O=FM

RD	Object wo		Object wa		Object ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
PM	26	46.4	2	3.6	28	50.0	0	0.0	56	100
FM	29	39.2	4	5.4	34	45.9	7	9.5	74	100
Total	55	42.3	6	4.6	62	47.7	7	5.4	130	100

Similar to Table (38), Table (39) does not show any significant difference of postnominal markings on the object depending on the RD of the object, either PM or FM.[31]

2.4.3.1.2. RD and the choice of postnominal markings: OSV order

In Table (40) the OSV order tokens are grouped together according to the postnominal marking of the object with respect to the RD of the object.

Table (40) RD of Object and Object Marking in OSV order

RD	Object wo		Object wa		Object Ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
1	6	54.5	1	9.1	4	36.4	0	0.0	11	100
2	0	0.0	1	100.0	0	0.0	0	0.0	1	100
3	1	50.0	0	0.0	1	50.0	0	0.0	2	100
4≤20+	1	25.0	3	75.0	0	0.0	0	0.0	4	100
FM	10	52.6	4	21.1	4	21.1	1	5.3	19	100
Total	18	48.6	9	24.3	9	24.3	1	2.7	37	100

Table (40) shows an object-marking pattern in OSV order that is similar to the postnominal marking of the object in SOV order. The accusative *o* and zero-particle *ø* mark 72.9% of objects in OSV order (i.e. 27/37 tokens), and these two markings mark the majority of the object regardless of the RD of the object. The numbers in Table (40) were pooled together into two categories, when the object is previously mentioned (PM) and when the object is a first mention (FM). The result is shown in Table (41).

---

[31] Although it is irrelevant to the discussion if the choice of postnominal marking is dependent on the RD, the comparison of Tables (35) and (38) tells us about the property of a particular postnominal marker. For example, subjects in SOV order marked with the particle *mo* tend to be high in RD, as shown in Table (35). The *mo*-marked objects in SOV order are all first mentions, as shown in Table (38). This conforms to the discourse function of *mo*, i.e. the use of *mo* implies the presence of presupposition, and it activates the presupposed entities at the moment of utterance. Subjects marked with *wa* tend to have a high RD, as shown in Table (35), and *wa*-marked objects occur only when the RD is high, as shown in Table (38). It confirms the argument that *wa* is a marker of an entity that was already introduced in the previous discourse. The observed correlation between a higher RD and *wa*-marked entities does not support the claim that associates *wa* and given information (e.g. Mikami 1963, Kuno, 1973, Maynard 1980).

Table (41) Postnominal marking on the object in OSV order when O=PM vs. O=FM

RD	Object wo		Object wa		Object Ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
PM	8	44.4	5	27.8	5	27.8	0	0.0	18	100
FM	10	52.6	4	21.1	4	21.1	1	5.3	19	100
Total	18	48.6	9	24.3	9	24.3	1	2.7	37	100

Table (41) shows no statistical significance with respect to the dependency of the postnominal marking of the object on the RD of the object, either PM or FM.[32] The Table also shows a tendency that the speaker uses the nominative *ga* more often than other postnominal markings when the object in OSV is a first mention.

Table (42) shows the token distribution of subjects in OSV order in terms of RD, with respect to different postnominal markings.

Table (42) RD of Subject and Subject Marking in OSV Order

RD	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
1	1	25.0	0	0.0	2	50.0	0	0.0	1	25.0	4	100
2	2	50.0	1	25.0	0	0.0	1	25.0	0	0.0	4	100
3 ≤ 20+	2	22.2	1	11.1	5	55.6	1	11.1	0	0.0	9	100
FM	9	45.0	1	5.0	7	35.0	3	15.0	0	0.0	20	100
Total	14	37.8	3	8.1	14	37.8	5	13.5	1	2.7	37	100

Table (42) shows that the nominative *ga* and zero-particle *ø* mark 75.6% of subjects in OSV order (i.e. 28/37 tokens), and regardless of the RD of the subject, the nominative *ga* and zero-particle *ø* are the postnominal markers for the majority of subjects in OSV order. Table (43) shows the number of tokens in two categories, when the subject is previously mentioned (PM) and when the subject is a first mention (FM). The numbers were pooled from Table (42) to yield Table (43).

---

[32] The *mo* marking of the object in the OSV construction in the present data is only found when the object is a first mention. This result confirms that the particle *mo* encodes the presupposition.

Table (43) Postnominal marking on the subject in OSV order when S=PM vs. S=FM

RD	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
PM	5	29.4	2	11.8	7	41.2	2	11.8	1	5.9	17	100
FM	9	45.0	1	5.0	7	35.0	3	15.0	0	0.0	20	100
Total	14	37.8	3	8.1	14	37.8	5	13.5	1	2.7	37	100

The figures in Table (43) do not show any statistical significance as to the dependency of postnominal marking on the RD of the subject, either PM or FM.

2.4.3.1.3. RD and the choice of postnominal markings: summary

As to the correlation between particular particles and the RD of subjects or objects, only the correlation between the subject marking of the nominative *ga* and first mention and between the subject marking of the particle *wa* and previous mention was observed. There was no correlation between the object markings and the RD of objects.

The comparison of subject markings in SOV and OSV (cf. Table (35) vs. Tables (42) and (43)) tells us that in both SOV and OSV, when the subject is a first mention, the speaker tends to use *ga*-marking. The *ga*-marking on the subject in SOV order that is a first mention proved to be statistically significant. The preference for *ga*-marking of first mention subjects in OSV order (i.e. Late-constituent subjects), however, is not as strong as that for *ga*-marking of first mention subjects in SOV order (i.e. Early-constituent subjects).

There is a tendency that *wa*-marked subjects in SOV order (i.e. Early-constituent subjects) show lower RD. Such a tendency, however, was not observed among *wa*-marked subjects in OSV order (Late-constituent subjects). Likewise, the *wa* marking on objects, either in SOV order or OSV order, did not show such a tendency. This suggests that the claimed correlation between *wa*-marked entities and low RD is only valid when *wa*-marked constituents are clause-initial subjects.

2.4.3.2. TP and postnominal markings

In this section, I will discuss whether there is any correlation between the use of different particles and the count of TP of subjects and objects. In 2.4.3.2.1., I will examine the correlation in nonsubordinate clauses with SOV order, and in 2.4.3.2.2. that in nonsubordinate clauses with OSV order.

2.4.3.2.1. TP and the choice of postnominal markings: SOV order

Table (44) shows the TP of subjects in SOV order by the postnominal markings on the subject.

Table (44) TP of Subject and Subject Marking in SOV order

TP	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
0	9	37.5	6	25.0	5	20.8	3	12.5	1	4.2	24	100
1	7	22.6	8	25.8	9	29.0	2	6.5	5	16.1	31	100
2	6	31.6	7	36.8	4	21.1	2	10.5	0	0.0	19	100
3	3	15.0	10	50.0	5	25.0	1	5.0	1	5.0	20	100
≥4	10	27.8	16	44.4	7	19.4	2	5.6	1	2.8	36	100
Total	35	26.9	47	36.2	30	23.1	10	7.7	8	6.2	130	100

Table (44) shows no evidence of any dependency of the choice of postnominal markings on the TP of the subject. The tokens are distributed evenly across the count of TP, and there is no correlation between particular postnominal markings and the count of TP. The table also shows that when the subject is highly persistent, *wa*-marking prevails. In order to see whether there is any significant difference of postnominal markings between nonpersistent subjects and persistent subjects, the numbers in Table (44) were pooled together into two categories, when the subject is nonpersistent and when the subject is persistent in the following discourse. The result is shown in Table (45).

Table (45) Postnominal marking on the subject in SOV order when S is not persistent vs. when S is persistentP of Subject in SOV order

TP	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
nonpersistent	9	37.5	6	25.0	5	20.8	3	12.5	1	4.2	24	100
persistent	26	24.5	41	38.7	25	23.6	7	6.6	7	6.6	106	100
Total	35	26.9	47	36.2	30	23.1	10	7.7	8	6.2	130	100

Table (45) shows no statistically significant correlation. The table only shows a trend that when the subject is not persistent, the speaker uses the nominative *ga* more often than other particles, and when the subject is persistent, the particle *wa* is more often used than other postnominal markings.

Table (46) gives the number of tokens with respect to the TP of objects in SOV order by the different postnominal markings.

Table (46) TP of Object and Object Marking in SOV order

TP	Object wo		Object wa		Object ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
0	27	42.2	4	6.3	28	43.8	5	7.8	64	100
1	11	40.7	1	3.7	14	51.9	1	3.7	27	100
2	7	58.3	0	0.0	4	33.3	1	8.3	12	100
3	5	41.7	1	8.3	6	50.0	0	0.0	12	100
≥4	5	33.3	0	0.0	10	66.7	0	0.0	15	100
Total	55	42.3	6	4.6	62	47.7	7	5.4	130	100

Table (46) shows no evidence that the postnominal markings on the object in SOV order correlate to the count of TP. The accusative *o* and zero-particle *ø* mark 90% of objects in SOV order (i.e. 117/130 tokens), and regardless of the TP of the object, these two markers are the postnominal markings for the majority of objects in SOV order. Table (47) shows the number of object tokens in SOV order when the object is not persistent and when the object is persistent. The numbers were pooled from Table (46) to yield Table (47).

Table (47) Postnominal marking on the object in SOV order, when O is not persistent vs. when O is persistent.

TP	Object wo		Object wa		Object ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
nonpersistent	27	42.2	4	6.3	28	43.8	5	7.8	64	100
persistent	28	42.4	2	3.0	34	51.5	2	3.0	66	100
Total	55	42.3	6	4.6	62	47.7	7	5.4	130	100

Similar to Table (46), Table (47) shows no correlation between the postnominal markings and the count of TP on the object in SOV order.

2.4.3.2.2. TP and the choice of postnominal markings: OSV order

Table (48) shows the token distribution of objects in clauses with OSV order with respect to the TP, sorted by the different postnominal markings.

Table (48) TP of Objects and Object Marking in OSV order

TP	Object wo		Object wa		Object Ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
0	4	33.3	5	41.7	3	25.0	0	0.0	12	100
1	7	50.0	4	28.6	2	14.3	1	7.1	14	100
2	2	66.7	0	0.0	1	33.3	0	0.0	3	100
3	4	80.0	0	0.0	1	20.0	0	0.0	5	100
≥4	1	33.3	0	0.0	2	66.7	0	0.0	3	100
Total	18	48.6	9	24.3	9	24.3	1	2.7	37	100

Table (48) shows that there is no significant pattern of postnominal markings with respect to the TP count when the object is not persistent. When the object is persistent, the accusative *o* tends to be used. The table also show that the accusative *o* and zero-particle *ø* are the postnominal markers for the majority of objects in OSV order, and they are the only postnominal markers for persistent objects with TP≥2. In order to see whether there is any difference of postnominal markings on the object in OSV order between nonpersistent objects and persistent objects, the numbers in Table (48) were pooled together. The result is shown in Table (49) which shows the number of object tokens in OSV order in two categories, when O is not persistent and when O is persistent.

Table (49) Postnominal marking on the object in OSV order when O is not persistent vs. when O is persistent.

TP	Object wo		Object wa		Object Ø		Object mo		Total	
	#	%	#	%	#	%	#	%	#	%
nonpersistent	4	33.3	5	41.7	3	25.0	0	0.0	12	100
persistent	14	56.0	4	16.0	6	24.0	1	4.0	25	100
Total	18	48.6	9	24.3	9	24.3	1	2.7	37	100

Table (49) shows no statistically significant difference in postnominal markings depending on the TP, either nonpersistent or persistent. The table, however, shows a trend that the accusative *o* is more often used when the object is persistent.

Table (50) shows the token distribution of subjects in clauses with OSV order with respect to the TP, sorted by the different postnominal markings.

Table (50) TP of Subject and Subject marking in OSV order

TP	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
0	2	16.7	1	8.3	6	50.0	2	16.7	1	8.3	12	100
1	4	44.4	0	0.0	4	44.4	1	11.1	0	0.0	9	100
2	4	80.0	0	0.0	1	20.0	0	0.0	0	0.0	5	100
3	2	40.0	0	0.0	2	40.0	1	20.0	0	0.0	5	100
≥4	2	33.3	3	50.0	1	16.7	0	0.0	0	0.0	6	100
Total	14	37.8	4	10.8	14	37.8	4	10.8	1	2.7	37	100

The numbers in each cell in Table (50) are too small to draw any conclusion. The numbers in Table (50) were pooled together to yield Table (51), which shows the number of subject tokens in clauses with OSV order in two categories, when the subject is not persistent and when the subject is persistent.

Table (51) Postnominal marking on the subject in OSV order when S is not persistent vs. when S is persistent.

TP	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
nonpersistent	2	16.7	1	8.3	6	50.0	2	16.7	1	8.3	12	100
persistent	12	48.0	3	12.0	8	32.0	2	8.0	0	0.0	25	100
Total	14	37.8	4	10.8	14	37.8	4	10.8	1	2.7	37	100

Table (51) does not indicate the statistically significant dependency of the postnominal markings on the TP, either nonpersistent or persistent. The table only shows a trend that when the subject is nonpersistent, the zero-particle *ø* is used slightly more often than other postnominal markings, and when the subject is persistent, the nominative *ga* is more often used than other particles.

2.4.3.2.3. TP and the choice of postnominal markings: summary

The discussion in the preceding subsections 2.4.3.2.1. and 2.4.3.2.2. reveals that there is no dependency of the postnominal markings on the TP count. The trend, which falls short of statistical significance, was observed among subjects in SOV order; nonpersistent subjects tend to have the *ga*-marking and persistent subjects tend to have the *wa*-marking. This trend, however, is not consistent with the trend that subjects in OSV order show; nonpersistent subjects in OSV order tend to have zero marking, and persistent subjects in OSV order tend to have *ga*-marking. Even the trend indicates no correlation between particular particles and the TP, either persistent or nonpersistent.

2.4.3.3. SW and postnominal markings

In this section, I will examine whether there is any dependency of the postnominal markings on the SW count. In 2.4.3.3.1., I will study the correlation between particular postnominal markings and the SW of subjects and objects in SOV order, and in 2.4.3.3.2. the correlation in OSV order.

2.4.3.3.1. SW and the choice of postnominal markings: SOV order

Table (52) shows the token distribution of subjects in clauses with SOV order with respect to the SW count by their postnominal markings. Based on the statistics that the majority of subjects and objects concentrates on the SW count of one to four (cf. Tables (33) and (34) in 2.4.2.3), the tokens were pooled together into two categories, SW = 1 to 4 and SW≥5. The constituents with zero marking are by definition expected to be shorter.

Table (52) SW of Subject and Subject Marking in SOV order

SW	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
1-4	26	23.0	44	38.9	27	23.9	9	8.0	7	6.2	113	100
≥5	9	52.9	3	17.6	3	17.6	1	5.9	1	5.9	17	100
Total	35	26.9	47	36.2	30	23.1	10	7.7	8	6.2	130	100

Table (52) shows that when the subject has an SW of more than five counts, the speaker tends to use the nominative *ga*. Table (53) compares the use of the nominative *ga* and the particles other than *ga* for the syntactically heavier entities (i.e. SW≥5) and the syntactically lighter entities (i.e. SW≤4). The numbers in Table (52) were pooled together to yield Table (53).

Table (53) The use of the nominative ga for the syntactically heavier subject in SOV order

SW	ga		not ga		Total	
	#	%	#	%	#	%
1-4	26	23.0	87	77.0	113	100
≥5	9	52.9	8	47.1	17	100
Total	35	26.9	95	73.1	130	100

Table (53) shows that subjects with an SW≥5 are significantly more likely to be marked with *ga* than shorter subjects (p=0.016). The syntactically heavy entities tend to be nominalized clauses or nominals with modifiers. These clauses and modified nominals denote propositions rather than entities. They tend to have no referent or antecedent in the recent discourse since they contain more semantically complex information, and thus tend to be first mentions or to have high RD. In that sense, the syntactic heaviness corresponds to high RD or first mentions. This result confirms to the finding in section 2.4.2.1.1. that the use of the nominative *ga* correlates with first mention.

In order to see if the tendency for long subjects to be marked with *ga* is due to their being high-RD or first mentions, the eleven tokens of S-*ga* with SW≥5 in Table (53) were further grouped together according to the count of RD. Likewise, the RD of four tokens of S-*ga* in OSV order shown in Table (57) in section 2.4.3.3.2. below were also examined. Table (54) shows the RD of subjects marked with the nominative *ga* that have an SW greater than or equal to 5.

Table (54) RD of subjects marked with *ga* that have an SW≥5

	RD											
S-ga with SW≥5	1		3		9		20		FM		Total	
	#	%					#	%	#	%	#	%
S-ga in SOV	1	11.1	0	0.0	0	0.0	1	11.1	7	77.8	9	100
S-ga in OSV	1	25.0	1	25.0	1	25.0	0	0.0	1	25.0	4	100
Total	2	15.4	1	7.7	1	7.7	1	7.7	8	61.5	13	100

Table (54) shows that a significant percentage of longer subjects in SOV order that are marked with *ga* are first mentions. However, it is not the case for subjects in OSV order that are marked with *ga*. This result evidences the correlation between the high SW and first mention on *ga*-marked subjects in SOV.

Table (55) shows the number of object tokens in SOV order with respect to the SW, sorted by the postnominal markings.

Table (55) SW of Object and Object Marking in SOV order

SW	Object-wo		Object-wa		Object-ø		Object-mo		Total	
	#	%	#	%	#	%	#	%	#	%
1-4	42	42.4	3	3.0	51	51.5	3	3.0	99	100
≥5	13	41.9	3	9.7	11	35.5	4	12.9	31	100
Total	55	42.3	6	4.6	62	47.7	7	5.4	130	100

Table (55) shows no evidence that the choice of postnominal markings depends on the SW count.

2.4.3.3.2. SW and the choice of postnominal markings: OSV order

Table (56) shows the token distribution of objects and subjects in OSV order with respect to the SW count, sorted by the postnominal markings.

Table (56) SW of Object and Object Marking in OSV order

SW	Object-wo		Object-wa		Object-ø		Object-mo		Total	
	#	%	#	%	#	%	#	%	#	%
1-4	8	38.1	5	23.8	7	33.3	1	4.8	21	100
≥5	10	62.5	2	12.5	4	25.0	0	0.0	16	100
Total	18	48.6	7	18.9	11	29.7	1	2.7	37	100

Table (56) shows that in my data, objects with an SW greater than or equal to 5 are marked with the accusative *o* more often than objects with an SW less than 5 (62.5% vs. 38.1%), though this is not statistically significant (p=0.191)[33]. Table (57) shows the distribution of subject tokens in OSV order when SW=1 to 4 and when SW≥5, sorted by the postnominal markings.

---

[33] The numbers , 8 (*o*-marked objects with SW = 1≤4), 13 (objects that are marked by other than *o* with SW = 1≤4), 10 (*o*-marked objects with SW ≥ 5), and 6 (objects that are marked by other than *o* with SW≤5), were computed for Fisher Exact Test.

Table (57) SW of subject and Subject Marking in OSV order

SW	Subject ga		Subject wa		Subject ø		Subject mo		Subject de		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
1-4	10	31.3	3	9.4	14	43.8	4	12.5	1	3.1	32	100
≥5	4	80.0	0	0.0	0	0.0	1	20.0	0	0.0	5	100
Total	14	37.8	3	8.1	14	37.8	5	13.5	1	2.7	37	100

Table (57) shows a trend that subjects with an SW greater or equal to 5 are marked with the nominative *ga* more often than subjects with an SW less than 5 (80% vs. 31.3%), and this trend falls just short of statistical significance (p=0.057)[34].

2.4.3.3.3. SW and the choice of postnominal markings: summary

The data presented in the preceding subsection 2.4.3.3.1. and 2.4.3.3.2. indicates that there is no dependency of the object markings on the SW count. The *ga*-marking of subjects in SOV order that have an SW greater or equal to 5 proved to be statistically significant, while the *ga*-marking of subjects in OSV order with SW≥5 does not. The parallelism between the correlation of the *ga*-marking on subjects with first mentions and the correlation of the *ga*-marking on subjects with higher SW suggests the relation between the informational richness and structural complexity. Both correlations between the use of the nominative *ga* and first mention, and the *ga*-marking and high SW proved to be statistically significant for subjects in SOV order (Early-constituent subjects), while these correlations were trends for subjects in OSV order (Late-constituent subjects).

---

[34] The numbers , 10 (*ga*-marked subjects with SW = 1≤4), 13 (objects that are marked by other than *ga* with SW = 1≤4), 10 (*ga*-marked objects with SW ≥ 5), and 6 (objects that are marked by other than *ga* with SW≤5), were computed for Fisher Exact Test.

2.4.4. The 3NPs construction

There are seventy-five tokens of the 3NPs construction in SO order and nineteen tokens in OS order. A total of 93 tokens are also grouped together according to the position of noun phrases with respect to two other constituents, i.e. X-initial, X-medial and X-final[35] (X stands for the noun phrase other than the subject and the object in the 3NPs construction). Table (58) summarizes the tokens of the 3NPs construction in the seven ordering variations, SOXV, SXOV, XSOV, OSXV, OSOV, OXSV and XOSV.

---

[35] X-initial, X-medial and X-final are merely labels. Depending on what constituent we focus on, SOXV order, for example, can be S-initial, O-medial or X-final.

Table (58) The number of tokens in the 3NPs construction

	SO order	# of tokens	OS order	# of tokens
X-final	**SOXV**		**OSXV**	
	SODV	3		
	SOQV	5	OSQV	3
	SOLV	2	OSLV	4
	SOInsV	2	OSTV	2
	SOCV	1	OSCV	2
			OSAV	1
			OSOV	
			OSOV	2
	Sub-total	*13*	*Sub-total*	*14*
X-middle	**SXOV**		**OXSV**	
	SDOV	2		
	SQOV	1		
	SLOV	6		
	SInsOV	4	OInsSV	1
	SComOV	2		
	STOV	13	OTSV	1
	SAOV	4		
	Sub-total	*32*	*Sub-total*	*2*
X-initial	**XSOV**		**XOSV**	
	QSOV	2		
	LSOV	7	LOSV	2
	ComSOV	1		
	TSOV	13	TOSV	1
	PSOV	1		
	ASOV	1		
	ToSOV	4		
	Sub-total	*29*	*Sub-total*	*3*
	Total	74	Total	19

S=subject, O=direct object, D=dative, Q=quotative, L=locative, Ins=instrumental, C=complement, Com=comitative, T=temporal, P=purpose, A=adverbial clause or phrase and To=topic.

As Table (58) shows, the number of the different sorts of clauses is too small to draw any statistically significant conclusions. Subsequently, the observations that I make in this section are based on the small numbers and are not in general statistically significant.

2.4.4.1. SOXV

Sentence (74) below is an example of the 3NPs construction in SOXV order, which is subject-initial and X-final in SO order. In all examples of the 3NPs construction in sections 2.4.4.1.to 2.4.4.6., the subject is double-underlined, the object is underlined, and the X is marked with a dotted underline.

(74) SODV
8-4 SM-kantoku ga anata o zenbu o-tukurininatta eiga ni
SM-director NOM 2sg ACC all HON-make.HON.PST movie DAT

kiyoo-siteiru
cast-do.PPG

'The director SM casts you in all the movies that he made.'

Tables (59)-(61) show the RD, TP and SW of subjects, objects and Xs (i.e. other phrases than subjects and objects in the 3NPs constructions) in SXOV order.

Table (59) RD of subjects, objects and Xs in SOXV order

SOXV	Subject		Object		X	
RD	#	%	#	%	#	%
1	6	46.2%	2	15.4%	1	7.7%
2	0	0.0%	1	7.7%	1	7.7%
3	0	0.0%	0	0.0%	0	0.0%
4 ≤ 20	3	23.1%	1	7.7%	0	0.0%
20+	0	0.0%	1	7.7%	1	7.7%
FM	4	30.8%	8	61.5%	10	76.9%
Total	13	100%	13	100%	13	100%

Table (60) TP of subjects, objects and Xs in SOXV order

SOXV	Subject		Object		X	
TP	#	%	#	%	#	%
0	2	15.4%	6	46.2%	9	69.2%
1	2	15.4%	1	7.7%	2	15.4%
2	4	30.8%	3	23.1%	2	15.4%
3	2	15.4%	2	15.4%	0	0.0%
≥4	3	23.1%	1	7.7%	0	0.0%
Total	13	100%	13	100%	13	100%

Table (61) SW of subjects, objects and Xs in SOXV order

SOXV	Subject		Object		X	
SW	#	%	#	%	#	%
1-2	7	53.8%	4	30.8%	4	30.8%
3-4	3	23.1%	4	30.8%	7	53.8%
5-8	2	15.4%	2	15.4%	2	15.4%
≥9	1	7.7%	3	23.1%	0	0.0%
Total	13	100%	13	100%	13	100%

Table (59) shows that a higher percentage of clause-initial subjects in SOXV order have an RD of 1 than objects or Xs (46.2% for subjects, 15.4% for objects and 7.7% for Xs), while a higher percentage of objects and Xs than subjects are first mentions (30.7% for subjects, 61.5% for objects, and 76.9% for other phrases). Table (60) shows that a higher percentage of objects and Xs than subjects are nonpersistent (15.4% for subjects, 46.2% for objects, and 69.2% for Xs), while 85% of subjects are persistent. The table also shows that a higher percentage of subjects than objects and Xs have TP of greater than four (23.1% for subjects and 7.7% for objects and 0% for Xs). Table (61) shows that a higher percentage of objects than subjects or Xs have an SW of greater than nine counts (7.7% for subjects, 23.1% for objects and 0% for Xs), whereas a higher percentage of subjects than objects or Xs have an SW of one to two counts (53.8% for subjects and 30.8% for objects and Xs). The tendency that a higher percentage of subjects, either Early-constituents or Late-constituents, have smaller SW counts was observed in the data of the SOV construction (cf. section 2.4.2.3.).

2.4.4.2. SXOV

Sentence (75) is an example of the 3NPs construction in SXOV order, which is subject-initial and X-medial in SO order.

(75) SDOV

22-16	K.T. ga	hinin-tati	ni	o-kane	o	megumu-n-desu-yo
	K.T. NOM	untouchable-PL	DAT	POL-money	ACC	give-NMLZ-COP-FP

‘K.T. gives money to the untouchables.’

Tables (62)-(64) show the RD, TP and SW of subjects, objects and Xs in SXOV order, respectively.

Table (62) RD of subjects, objects and Xs in SXOV order

SXOV	Subject		X		Object	
RD	#	%	#	%	#	%
1	11	34.4%	3	9.4%	4	12.5%
2	2	6.3%	0	0.0%	4	12.5%
3	1	3.1%	2	6.3%	0	0.0%
4 ≤ 20	9	28.1%	4	12.5%	4	12.5%
20+	3	9.4%	1	3.1%	2	6.3%
FM	6	18.8%	22	68.8%	18	56.3%
Total	32	100%	32	100%	32	100%

Table (63) TP of subjects, objects and Xs in SXOV order

SXOV	Subject		X		Object	
TP	#	%	#	%	#	%
0	9	28.1%	23	71.9%	12	37.5%
1	9	28.1%	6	18.8%	10	31.3%
2	5	15.6%	3	9.4%	6	18.8%
3	2	6.3%	0	0.0%	1	3.1%
≥4	7	21.9%	0	0.0%	3	9.4%
Total	32	100%	32	100%	32	100%

Table (64) SW of subjects, objects and Xs in SXOV order

SXOV	Subject		X		Object	
SW	#	%	#	%	#	%
1-2	18	56.3%	13	40.6%	18	56.3%
3-4	10	31.3%	9	28.1%	7	21.9%
5-8	3	9.4%	10	31.3%	5	15.6%
≥9	1	3.1%	0	0.0%	2	6.3%
Total	32	100%	32	100%	32	100%

Similar to Table (59), Table (62) shows that a higher percentage of subjects than objects or Xs have an RD of 1 (33.3% for subjects, 9.1% for Xs and 12.1% for objects), while a higher percentage of objects and Xs than subjects are first mentions (21.2% for subjects, 69.7% for Xs and 57.6% for objects). Similar to Table (60), Table (63) shows that a higher percentage of subjects than objects or Xs have TP≥4 (21.2% for subjects, 0% for Xs 9.1% and for objects), and a higher percentage of objects and Xs than subjects are nonpersistent (30.3% for subjects, 72.7% for Xs and 36.4% for objects). Table (64) shows that a higher percentage of Xs than subjects or objects have an SW of greater than five counts (12.1% for subjects, 21.3% for objects and 33% for Xs).

2.4.4.3. XSOV

Sentence (76) is an example of the 3NPs construction in XSOV order, which is X-initial in SO order.

(76) LSOV
13-9 hooboo no mise kara boku syootaizyoo moratteru-n-desu
here.and.there GEN store from 1sg.male invitation.letter receive.RES-NMLZ-COP
'I received the invitation from many stores in all directions.'

Tables (65)-(67) show the RD, TP and SW of subjects, objects and Xs in XSOV order.

Table (65) RD of subjects, objects and Xs in XSOV order

XSOV	X		Subject		Object	
RD	#	%	#	%	#	%
1	4	13.8%	8	27.6%	2	6.9%
2	1	3.4%	3	10.3%	0	0.0%
3	0	0.0%	2	6.9%	1	3.4%
4 ≤ 20	5	17.2%	7	24.1%	5	17.2%
20+	1	3.4%	1	3.4%	1	3.4%
FM	18	62.1%	8	27.6%	20	69.0%
Total	29	100%	29	100%	29	100%

Table (66) TP of subjects, objects and Xs in XSOV order

XSOV	X		Subject		Object	
TP	#	%	#	%	#	%
0	20	69.0%	7	24.1%	20	69.0%
1	5	17.2%	8	27.6%	4	13.8%
2	1	3.4%	2	6.9%	3	10.3%
3	1	3.4%	6	20.7%	0	0.0%
≥4	2	6.9%	6	20.7%	2	6.9%
Total	29	100%	29	100%	29	100%

Table (67) SW of subjects, objects and Xs in XSOV order

XSOV	X		Subject		Object	
SW	#	%	#	%	#	%
1-2	11	37.9%	18	62.1%	19	65.5%
3-4	9	31.0%	8	27.6%	5	17.2%
5-8	8	27.6%	2	6.9%	4	13.8%
≥9	1	3.4%	1	3.4%	1	3.4%
Total	29	100%	29	100%	29	100%

Table (65) shows that a higher percentage of objects and Xs than subjects are first mentions (69% for object, 62.1% for Xs and 27.6% for subjects, while a higher percentage of subjects than Xs or objects have an RD of 1 (27.6% for subjects, 13.8% for Xs and 6.9% for objects). The percentage of Xs that have an RD of 1 in XSOV order, 13.8%, is higher than that of Xs that have an RD of 1 in SOXV or SXOV orders (13.8% for XSOV order, 7.7% for SOXV order and 9.1% for SXOV order). Table (66) shows that a higher percentage of Xs and objects than subjects are nonpersistent (69% for Xs and objects and 24.1% for subjects), while a higher percentage of subjects than Xs or objects have a TP of greater than three (41.4% for subjects, 10.3% for Xs and 6.9% for objects). Similar to Table (64), Table (67) shows that a higher percentage of Xs than subjects or objects have an SW of greater than five counts (31% for Xs, 10.3% for subjects and 17.2% for objects). This result contrasts with Xs in SOXV order (cf. Table (61) in 2.4.4.1.) that have the least percentage of tokens with SW≥5 among subjects, objects and Xs (31% for Xs with SW≥5 in XSOV order, 33.3% for Xs with SW≥5 in SXOV order and 15.4% for Xs with SW≥5 in SOXV order).

2.4.4.4. OSXV[36]

Sentence (77) is an example of the 3NPs construction in OSXV order, which is object-initial and X-final in OS order.

(77) OSTV

14-3	sore	o	watasi	ga	syoogaku-sei	no
	it	ACC	1sg	NOM	elementary.school-student	GEN

tei-gakunen	no	toki	ni	yonde
low-grade	GEN	time	LOC	read-TE

'I read it when being a lower grader in grammar school.'

Tables (68)-(70) show the RD, TP, and SW of subjects, objects and Xs in OSXV and OSOV orders.

Table (68) RD of subjects, objects and Xs in OSXV order

OSXV	Object		Subject		X	
RD	#	%	#	%	#	%
1	6	42.9%	0	0.0%	2	14.3%
2	1	7.1%	2	14.3%	0	0.0%
3	2	14.3%	2	14.3%	0	0.0%
4 ≤ 20	2	14.3%	4	28.6%	2	14.3%
20+	0	0.0%	3	21.4%	0	0.0%
FM	3	21.4%	3	21.4%	10	71.4%
Total	14	100%	14	100%	14	100%

Table (69) TP of subjects, objects and Xs in OSXV order

OSXV	Object		Subject		X	
TP	#	%	#	%	#	%
0	4	28.6%	3	21.4%	10	71.4%
1	4	28.6%	2	14.3%	3	21.4%
2	3	21.4%	1	7.1%	0	0.0%
3	0	0.0%	2	14.3%	1	7.1%
≥4	3	21.4%	6	42.9%	0	0.0%
Total	14	100%	14	100%	14	100%

[36] For the statistic purpose, I will include 2 tokens with OSOV order in OSXV order category.

Table (70) SW of subjects, objects and Xs in OSXV order

OSXV	Object		Subject		X	
SW	#	%	#	%	#	%
1-2	7	50.0%	9	64.3%	8	57.1%
3-4	2	14.3%	4	28.6%	2	14.3%
5-8	3	21.4%	1	7.1%	4	28.6%
≥9	2	14.3%	0	0.0%	0	0.0%
Total	14	100%	14	100%	14	100%

Table (68) shows that a higher percentage of objects than subject or Xs have an RD of 1 (42.9% for objects 14.3% for Xs and 0% for subjects), while a higher percentage of Xs than objects or subjects are first mentions (71.4% for Xs, 21.4% for objects and subjects). Objects in object-initial OSXV order exhibit a higher percentage of tokens with RD=1 than objects in object-final SXOV and XSOV orders, and objects in object-medial SOXV order (42.9% for OSXV order, 12.1% for SXOV order, 6.9% for XSOV order and 15.4% for SOXV order). Moreover, and objects in object-initial OSXV order exhibit a lower percentage of first mentions than objects in object-final SXOV and XSOV orders, and objects in object-medial SOXV order (21.4% for OSXV order, 57.6% for SXOV order, 69% for XSOV order and 61.5% for SOXV order). Similar to Tables (60), (63) and (66), Table (69) shows that a higher percentage of Xs than objects and subjects are nonpersistent (71.4% for Xs, 28.6% for objects and 21.4% for subjects). The proportion of objects that are nonpersistent in OSXV order was lower than that of objects in SOXV, SOXV and XSOV orders (28.6% for OSXV order, 46.2% for SOXV order, 36.4% for SXOV order and 69% for XSOV order), indicating than a lower percentage of clause-initial objects than non clause-initial objects are nonpersistent. Table (70) shows that a higher percentage of objects than subjects and Xs have an SW of greater than five counts (35.7% for objects 28.6% for Xs and 7.1% for subjects). The proportion of objects that have an SW of greater than five counts was higher in OSXV order than in SXOV and XSOV orders (35.7% for OSXV order, 21.3% for SXOV order, 17.2% for XSOV order and 38.5% for SOXV order).

2.4.4.5. OXSV

Sentence (78) is an example of the 3NPs construction in OXSV order, which is object-initial and X-medial in OS order.

(78) OInsSV
20-14 nanya tte no oosaka-ben de anata ga itte-rassyaru
what QT NMLZ Osaka-dialect INS 2sg NOM say-HON
'You said 'What' in Osaka dialect.'

Tables (71)-(73) show the RD, TP and SW of subjects, objects and Xs in OXSV order.

Table (74) RD of subjects, objects and Xs in OXSV order

OXSV	Object		X		Subject	
RD	#	%	#	%	#	%
1	1	50.0%	0	0.0%	0	0.0%
4 ≤ 20	1	50.0%	1	50.0%	2	100.0%
20+	0	0.0%	0	0.0%	0	0.0%
FM	0	0.0%	1	50.0%	0	0.0%
Total	2	100%	2	100%	2	100%

Table (72) TP of subjects, objects and Xs in OXSV order

OXSV	Object		X		Subject	
TP	#	%	#	%	#	%
0	2	100.0%	2	100.0%	0	0.0%
3	0	0.0%	0	0.0%	1	50.0%
≥4	0	0.0%	0	0.0%	1	50.0%
Total	2	100%	2	100%	2	100%

Table (73) SW of subjects, objects and Xs in OXSV order

OXSV	Object		X		Subject	
SW	#	%	#	%	#	%
1-2	0	0.0%	0	0.0%	1	50.0%
3-4	1	50.0%	2	100.0%	1	50.0%
5-8	0	0.0%	0	0.0%	0	0.0%
≥9	1	50.0%	0	0.0%	0	0.0%
Total	2	100%	2	100%	2	100%

Table (71) shows that only objects have an RD of 1, while Xs and subjects do not. Table (72) shows that objects and Xs are nonpersistent, while subjects are persistent. Table (73) shows that objects have an SW of greater than nine counts, while the SW of all Xs and subjects have an SW count from one to four.

2.4.4.6. XOSV

Sentence (79) is an example of the 3NPs construction in XOSV order, which is X-initial in OS order.

(79) LOSV
8-19 sono naka ni anata no koto Y-sensei ga
its inside LOC 2sg GEN matter Y-teacher NOM

totemo homete-rassyaru-no-ne
very praise.TE-HON-FP-FP

'In it, Mr. Y praises you.'

Tables (74)-(76) show the RD, TP and SW of the subject, object and X in XOSV order.

Table (74) RD of subjects, objects and Xs in XOSV order

XOSV	X		Object		Subject	
RD	#	%	#	%	#	%
1	1	33.3%	0	0.0%	0	0.0%
2	0	0.0%	0	0.0%	2	66.7%
3	0	0.0%	1	33.3%	0	0.0%
4 ≤ 20	1	33.3%	1	33.3%	0	0.0%
20+	0	0.0%	0	0.0%	0	0.0%
FM	1	33.3%	1	33.3%	1	33.3%
Total	3	100%	3	100%	3	100%

Table (75) TP of subjects, objects and Xs in XOSV order

XOSV	X		Object		Subject	
TP	#	%	#	%	#	%
0	2	66.7%	1	33.3%	1	33.3%
2	0	0.0%	1	33.3%	1	33.3%
3	1	33.3%	1	33.3%	0	0.0%
5	0	0.0%	0	0.0%	1	33.3%
Total	3	100%	3	100%	3	100%

Table (76) SW of subjects, objects and Xs in XOSV order

XOSV	Object		X		Subject	
SW	#	%	#	%	#	%
1-2	2	66.7%	1	33.3%	2	66.7%
3-4	1	33.3%	1	33.3%	1	33.3%
≥9	0	0.0%	1	33.3%	0	0.0%
Total	3	100%	3	100%	3	100%

Table (74) shows that only the X has an RD of 1, while the object and the subjects do not. Tables (75) and (76) show that the three constituents have a comparable TP and SW.

2.4.4.7. Summary

Table (77) summarizes the proportion of tokens that have an RD of 1, that are previously mentioned (PM) and that are first mentions (FM). The symbol ">" reads as "a higher proportion of", "<" as "a lower proportion of" and "=" as "an equal proportion of". For example, "S > O > X" in the row of SOXV order and in the column of RD = 1 reads as "A higher proportion of subjects than objects and Xs (other phrases than subjects or objects) have an RD of 1, and a higher proportion of objects than Xs have an RD of 1".

Table (77) Summary: Relative proportion of tokens with respect to the RD

Order	# of tokens	RD=1	PM	FM
SOXV	13	S > O > X	S > O > X	X > O > S
SXOV	32	S > O > X	S > O > X	X > O > S
OSXV	14	O > X > S	O>X, S>X, O=S	X>O, X>S, O=S
OXSV	2	O>X, O>S, X=S	O>X, S>X, O=S	X>O, X>S, O=S
XSOV	29	S > X > O	X > S > O	O > S > X
XOSV	3	X>O, X>S, O=S	X = O = S	X = O = S

Table (77) shows that in other five orders than XSOV order, a higher percentage of clause-initial constituents, either subjects, objects or other phrases, have an RD of 1 than other constituents in the same clause. The table also shows that a lower percentage of subjects than objects or Xs, regardless of their position in a clause, are first mentions. The results indicate a possibility of low RD as a parameter for detecting the constituent that comes in the clause-initial position.

Table (78) summarizes the proportion of tokens that are not persistent in the following discourse, that are persistent in the following discourse and that have a TP of greater than four.

Table (78) Summary: Relative proportion of tokens with respect to the TP

Order	# of tokens	Nonpersistent	Persistent	TP ≥ 4
SOXV	13	X > O > S	S > O > X	S > O > X
SXOV	32	X > O > S	S > O > X	S > O > X
OSXV	14	X > O > S	S > O > X	S > O > X
OXSV	2	O>S, X>S, O=X	S>O, S>X, O=X	S > O > X
XSOV	29	X>S, O>S, X=O	S>X, S>O, X=O	S>X, S>O, X=O
XOSV	3	X>O, X>S, O=S	O>X, S>X, O=S	S>X, S>O, X=O

Table (78) shows that a higher proportion of subjects than objects and Xs in the same clause, regardless their position in a clause, is highly persistent. The table also shows that a higher proportion of Xs than subjects and objects in the same clause, regardless their position in a clause, is nonpersistent. Moreover, a higher proportion of objects than Xs in the same clause is persistent. In object-initial orders, a higher proportion of subjects than objects is persistent. The result suggests the possibility of using high TP as a parameter for detecting the occurrence of subjects.

Table (79) summarizes the proportion of tokens that have an SW of one to four counts, that have an SW of five to eight counts and that have an SW of greater than nine.

Table (79) Summary: Relative proportion of tokens with respect to the SW

Order	# of tokens	SW = 1~4	SW ≥5	SW ≥ 9
SOXV	13	S > O > X	O > S > X	O > S > X
SXOV	32	S > O > X	X > O > S	O > S > X
OSXV	14	S > X > O	O > X > S	O>S, O>X, S=X
OXSV	2	X>O, S>O, X=S	O = X = S	O>S, O>X, S=X
XSOV	29	S > O > X	X > O > S	X = S = O
XOSV	3	O>X, S>X, O=S	X = O = S	X>O, X>S, O=S

Table (79) shows that except in X-initial orders, a higher proportion of objects than that of subjects and Xs in the same clause has an SW of greater than nine. In X-initial orders, a higher proportion of Xs than that of subjects and objects has an SW of greater than five. The link between the syntactic heaviness and clause-initial constituents are more pronounced among clause-initial objects and Xs than clause-initial subjects. The table also shows that when objects precede Xs, such as in OXSV, SOXV and OSXV orders, the proportion of objects that have an SW of greater than five is higher than that of Xs. This result indicates the possibility of using syntactic heaviness as a parameter to detect among non-subject constituents what comes in the clause-initial position, or to detect which constituent of two non-subject clause-internal constituents, such as O vs. X, can come earlier.

2.4.5. Postnominal particle and position

The present data indicates that postnominal marking does not correlate with constituent ordering except in the case of the particle *wa*. Table (80) gives the number of constituents sorted by the postnominal marking with respect to their position in a clause. “Clause-initial” means that marking occurs on a clause-initial NP. “2nd position” means that marking occurs on an NP that is the second NP in the clause. “3rd position” means that marking occurs on an NP that is the third NP in the clause. "4th position" means that marking occurs on an NP that is the fourth NP in the clause, and "5th position" on an NP that is the fifth NP in the clause. The total number of clause-initial or 2nd position NPs (281) is the total number of tokens in the 2NPs, 3NPs, 4NPs and 5NPs constructions in the present data. The total number of 3rd position NPs (114) is the total number of tokens with the 3NPs, 4NPs and 5NPs constructions. The total number of 4th position NPs (21) is the total number of the 4NPs and 5NPs construction tokens, and that of 5th position NPs (5) is the number of tokens with the 5NPs construction (see Table (17) in 2.4 for the number of the 2NPs, 3NPs, 4NPs and 5NPs construction in the present data).

Table (80) Particles and position

Particles	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
ga	56	62.9	27	30.3	4	4.5	1	1.1	1	1.1	89	100
wo	27	20.3	65	48.9	35	26.3	5	3.8	1	0.8	133	100
wa	84	77.1	24	22.0	1	0.9	0	0.0	0	0.0	109	100
ø	65	29.5	107	48.6	36	16.4	10	4.5	2	0.9	220	100
mo	18	46.2	18	46.2	3	7.7	0	0.0	0	0.0	39	100
de	17	33.3	19	37.3	13	25.5	2	3.9	0	0.0	51	100
ni	6	20.7	9	31.0	11	37.9	3	10.3	0	0.0	29	100
to	3	17.6	5	29.4	8	47.1	0	0.0	1	5.9	17	100
kara	2	33.3	2	33.3	2	33.3	0	0.0	0	0.0	6	100
made	1	100	0	0.0	0	0.0	0	0	0	0	1	100
others	2	25.0	5	62.5	1	12.5	0	0.0	0	0.0	8	100
total	281	40.0	281	40.0	114	16.2	21	3.0	5	0.7	702	100

Table (80) shows no significant correlation between a particular particle and the position in a clause except that 77.1% of constituents marked by *wa* and 62.9% of constituents marked by *ga* occur at the clause-initial position. The study on the postnominal marking and the count of RD, TP and SW in the SOV construction in the above section 2.4.3. indicates that *wa*-marked constituents are mostly referential, thus active, and first mention *wa*-marked constituents often occur when the mental representation of the constituents is active in the speaker's mind. The presence of Potential Interference is most frequent observed for *wa*-marked constituents.[37] Zero-particle is often used for entities that demonstrate characteristics of cohesion to their adjacent constituents; there are cases where a zero-marked Early-constituent subject is bound to its adjacent adverb to yield one semantic unit, or the cohesiveness of a zero-marked constituent to its verb often anchors the constituent in the preverbal position as in the light verb construction.

The data in Table (38) were broken down according to the grammatical role of the NP to create Table (81)-(83). Tables (81)-(83) show the number of subjects, objects and noun phrases other than subjects or objects by their postnominal marking with respect to their position in a clause.

Table (81) Particles and position: subjects

Particle of Subject	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
ga	56	62.9	27	30.3	4	4.5	1	1.1	1	1.1	89	100
wa	64	83.1	13	16.9	0	0.0	0	0.0	0	0.0	77	100
ø	43	55.8	31	40.3	3	3.9	0	0.0	0	0.0	77	100
mo	15	62.5	9	37.5	0	0.0	0	0.0	0	0.0	24	100
de	9	60.0	4	26.7	1	6.7	1	6.7	0	0.0	15	100
others	1	100.0	0	0.0	0	0.0	0	0.0	0	0.0	1	100
total	188	66.4	84	29.7	8	2.8	2	0.7	1	0.4	283	100

[37] This observation supports such a view that the use of *wa* is essentially generalizable in terms of contrastiveness (e.g. Clancy and Downing 1987, Shibatani 1990, Shimojo 2005).

Table (82) Particles and position: direct objects

Particle of Object	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
wo	27	20.3	65	48.9	35	26.3	5	3.8	1	0.8	133	100
wa	13	54.2	10	41.7	1	4.2	0	0.0	0	0.0	24	100
ø	12	10.3	65	56.0	27	23.3	10	8.6	2	1.7	116	100
mo	2	16.7	8	66.7	2	16.7	0	0.0	0	0.0	12	100
total	54	18.9	148	51.9	65	22.8	15	5.3	3	1.1	285	100

Table (83) Particles and position: other phrases

Particle of other phrases	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
wa	7	0.0	1	12.5	0	0.0	0	0.0	0	0.0	8	13
ø	10	37.0	11	40.7	6	0.0	0	0.0	0	0.0	27	78
mo	1	33.3	1	33.3	1	33.3	0	0.0	0	0.0	3	100
de	8	22.2	15	41.7	12	33.3	1	2.8	0	0.0	36	100
ni	6	20.7	9	31.0	11	37.9	3	10.3	0	0.0	29	100
to	3	17.6	5	29.4	8	47.1	0	0.0	1	5.9	17	100
kara	2	33.3	2	33.3	2	33.3	0	0.0	0	0.0	6	100
made	1	100.0	0	0.0	0	0.0	0	0.0	0	0.0	1	100
others	1	14.3	5	71.4	1	14.3	0	0.0	0	0.0	7	100
total	39	29.1	49	36.6	41	30.6	4	3.0	1	0.7	134	100

Table (81) shows that regardless of postnominal markings, subjects tend to be clause-initial. Table (82) shows that regardless of postnominal markings, objects tend to occur in the 2nd position except objects marked with *wa* which do not show a particular preference for the 2nd position. Table (83) shows that *wa* and *mo* markings of noun phrases other than subjects or objects are significantly less frequent than other postnominal markings. The comparison of Tables (81), (82) and (83) indicates that the particle is irrelevant to determining constituent ordering. For example, *mo*-marked subjects tend to be clause-initial as shown in Table (81), *mo*-marked objects tend to occur in the 2nd position as shown in Table (82), and other phrases marked with *mo* appear in the clause-initial, 2nd and 3rd position as shown in Table (83). Rather a particular particle, when it associates with a particular grammatical role, seems to exhibit particular characteristics in discourse.

In sum, the particles marking correlates with the grammatical roles, and the particles themselves do not appear to be a relevant factor for ordering choice. The exception is the case of the particle *wa*, and I will test the significance of *wa*-marking with respect to the constituent ordering choice in Chapter Three.

2.4.6. Reference form and position

In this section, I will examine whether constituents with different reference forms display different characteristics in terms of the constituent position in a clause. Table (84) shows the number of constituents with different reference forms with respect to their position in a clause. The total number of clause-initial or 2nd position NPs (281) is the total number of the 2NPs and 3NPs construction tokens in the present data. The total number of 3rd position NPs (114) is the total number of tokens with the 3NPs, 4NPs and 5NPs construction. The total number of 4th position NPs (21) is the total number of the 4NPs and 5NPs construction tokens, and that of 5th position NPs (5) is the number of tokens with the 5NPs construction (see Table (17) in 2.4 for the number of the 2NPs, 3NPs, 4NPs and 5NPs construction in the present data).

Table (84) Reference forms and position

Reference From	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
Demonstrative	61	60.4	27	26.7	12	11.9	1	1.0	0	0.0	101	100
1st person pronoun	56	68.3	25	30.5	1	1.2	0	0.0	0	0.0	82	100
2nd person pronoun	20	57.1	12	34.3	3	8.6	0	0.0	0	0.0	35	100
3rd person pronoun	1	50.0	1	50.0	0	0.0	0	0.0	0	0.0	2	100
Quantifier	14	37.8	19	51.4	3	8.1	1	2.7	0	0.0	37	100
Reflexive	3	60.0	1	20.0	1	20.0	0	0.0	0	0.0	5	100
Nouns	126	28.9	193	44.3	93	21.3	19	4.4	5	1.1	436	100
Adverbials	0	0.0	3	75.0	1	25.0	0	0.0	0	0.0	4	100
total	281	40.0	281	40.0	114	16.2	21	3.0	5	0.7	702	100

Table (84) indicates that there is no particular reference form bound to a particular position in a clause. However, there is a tendency that demonstratives, first and second person pronouns and reflexives are introduced earlier in a clause. A more detailed examination reveals that reference forms are more likely to be bound to a particular grammatical role and its position. The data in Table (84) were further broken down according to the grammatical role of the NP to create Tables (85)-(87). Tables (85)-(87) show the number of subjects, objects, and other phrases in different reference forms with respect to their position in a clause. "Demonstratives" comprise demonstrative pronouns[38], nouns modified with demonstrative adjectives[39] and noun phrases containing demonstrative adverbs[40]. "Quantifiers" means constituents containing quantifiers such as *san-nin* (three-counter for person) 'three people'. Reflexives means constituents containing the reflexive *zibun* 'self'.

---

[38] Demonstrative pronouns are *kore* 'this', *sore* 'that', *are* 'that over there' *koko* 'here', *soko* 'there', *asoko* 'over there', *kotti* 'this way', *sotti* 'that way' *atti* 'that way over there', *kotira* 'this way (Polite)', *sotira* 'that way (Polite)' and *atira* 'that way over there (Polite)'

[39] Demonstrative adjectives are *kono* 'this', *sono* 'that', *ano* 'that over there', *kon'na* 'this sort of', *son'na* 'such' and *an'na* 'that sort of'.

[40] Demonstrative adverbs are *koo* 'like this', *soo* 'like that, so' and *aa* 'like that'. They are often used in the form of *koo-iu* (like this-say) 'this type of' such as in *koo-iu-hito* (like this-say-person) "people like this', *soo-iu* (so-say) 'such' such as in *soo-iu-mono* (so-say-thing) 'such a thing' and *aa-iu* (like that-say) 'like that' such as in *aa-iu-koto* (like that-say-matter) 'matters like that'.

Table (85) Reference form of subjects and their position

Subject	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
Demonstrative	16	80.0	4	20.0	0	0.0	0	0.0	0	0.0	20	100
1st person pronoun	56	71.8	21	26.9	1	1.3	0	0.0	0	0.0	78	100
2nd person pronoun	20	69.0	7	24.1	2	6.9	0	0.0	0	0.0	29	100
3rd person pronoun	1	50.0	1	50.0	0	0.0	0	0.0	0	0.0	2	100
Quantifier	14	45.2	16	51.6	0	0.0	1	3.2	0	0.0	31	100
Reflexive	3	0.0	1	20.0	1	20.0	0	0.0	0	0.0	5	40
Nouns	78	66.1	34	28.8	4	3.4	1	0.8	1	0.8	118	100
Adverbials	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
total	188	66.4	84	29.7	8	2.8	2	0.7	1	0.4	283	100

Table (85) shows that regardless of their reference forms, subjects are most likely clause-initial.

Table (86) Reference form of direct objects and their position

Object	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
Demonstrative	29	58.0	14	28.0	6	12.0	1	2.0	0	0.0	50	100
1st person pronoun	0	0.0	4	100.0	0	0.0	0	0.0	0	0.0	4	100
2nd person pronoun	0	0.0	4	80.0	1	20.0	0	0.0	0	0.0	5	100
3rd person pronoun	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
Quantifier	0	0.0	3	75.0	1	25.0	0	0.0	0	0.0	4	100
Reflexive	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
Nouns	25	11.3	123	55.4	57	25.7	14	6.3	3	1.4	222	100
Adverbials	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
total	54	18.9	148	51.9	65	22.8	15	5.3	3	1.1	285	100

Table (86) shows that demonstrative objects are most likely to occur in initial position.

Table (87) Reference form of other phrases and their position

Other Phrases	Clause-initial		2nd position		3rd position		4th position		5th position		Total	
	#	%	#	%	#	%	#	%	#	%	#	%
Demonstrative	16	51.6	9	29.0	6	19.4	0	0.0	0	0.0	31	100
1st person pronoun	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
2nd person pronoun	0	0.0	1	100.0	0	0.0	0	0.0	0	0.0	1	100
3rd person pronoun	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
Quantifier	0	0.0	0	0.0	2	100.0	0	0.0	0	0.0	2	100
Reflexive	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0
Nouns	23	24.0	36	37.5	32	33.3	4	4.2	1	1.0	96	100
Adverbials	0	0.0	3	75.0	1	25.0	0	0.0	0	0.0	4	100
total	39	29.1	49	36.6	41	30.6	4	3.0	1	0.7	134	100

Table (87) shows that demonstrative constituents other than subjects and objects are significantly most likely to occur in initial position (Fisher Exact Test p=0.002).

Tables (85)-(87)[41] show that when demonstrative constituents are not subjects or objects, they are significantly most likely to occur in initial position. I will discuss the statistical significance of non-subject constituents that are demonstratives and the possible use of demonstratives as a factor to implement an algorithm to predict the choice of word order in Chapter 3. Tables (85) and (86) indicate that first[42] and second person pronouns are bound to the grammatical role of subject than to the position in a clause. In sum, there is no significant correlation between a specific reference form and the position in a clause except the use of demonstrative.

---

[41] The tables also show the correlation between 1st and 2nd person pronouns and subjects, which is reported in Shimojo (2005:69).

[42] With respect to the use of first singular person pronouns which are commonly dropped in Japanese when the referent is highly active, the present data suggests that the actual distance between the current occurrence of an entity and its last explicit mention seems to be relevant to the usage of first singular personal pronouns. In passing, the average distance indicated by the present data is RD = 10, i.e. 10 clauses.

Chapter 3

The choice of constituent ordering: SOV vs. OSV

In Chapter 1, three measurable variables that the speakers might depend on in choosing a particular constituent order over another were identified: referential distance (RD), topic persistence (TP), and syntactic weight (SW). The method of measuring the variables of each token was established in Chapter 2. The goal of this chapter is to devise an algorithm that can predict the ordering choice between SOV and OSV. In order to achieve this goal, I will first start examining in what way the three measurable variables can work for predicting SOV order versus OSV order, and vice versa.

My database yielded 130 tokens of SOV order and 37 tokens of OSV order (see the breakdown of tokens in terms of particle marking in Tables (18)-(22) in Section 2.4.). All tokens contain a lexical subject and a lexical direct object. A total of 167 tokens comprise the study in this chapter.

The relative measurement of a variable in each token is important in that it tells us when a particular order is chosen over another and under what conditions. The categorization of RD, TP, and SW in terms of relative measurement is shown below. “S” stands for the subject, and “O” for the direct object (hereafter “object”, otherwise specified).

Referential distance (RD)

S > O RD of S is higher than RD of O
S = O RD of S and O is the same
S < O RD of O is higher than RD of S

For example RD = FM is higher than RD = 20+, RD = 20+ is higher than RD = 20, 10…1.

Topic persistence (TP)

S > O TP of S is higher than TP of O
S = O TP of S and O is the same
S < O TP of O is higher than RD of S

Syntactic weight (SW)

S	>	O	syntactic weight of S	is higher than	syntactic weight of O
S	=	O	syntactic weight of S	is the same as	syntactic weight of O
S	<	O	syntactic weight of O	is higher than	syntactic weight of S

3.1. Single variable

In this section, I will examine how the single variables, referential distance (RD), topic persistence (TP), and syntactic weight (SW), predict SOV order versus OSV order.

3.1.1. Referential distance (RD)

Table (1) shows the relative measurement of RD in SOV and OSV orders.

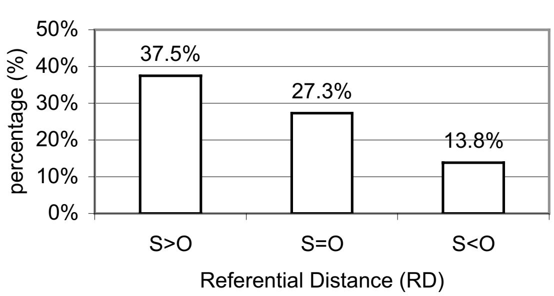
Table (1) Relative RD of S and O in SOV and OSV orders

RD	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	25	62.5	15	37.5	40	100
S=O	24	72.7	9	27.3	33	100
S<O	81	86.2	13	13.8	94	100
Total	130	77.8	37	22.2	167	100

Table (1) indicates that SOV order is more frequent than OSV order in all three categories: S<O (RD) where the RD of the subject is lower than that of the object; S=O

(RD) where the RD of the subject and the object is the same; and S>O (RD) where the RD of the subject is higher than that of the object. These results suggest that SOV is a default order and OSV is a marked order. The table also shows that the choice of SOV order is the highest when the RD of the subject is lower than the RD of the object, next highest when the RDs are the same for the subject and the object, and lowest when the RD of the object is higher than the RD of the subject. These results are displayed in Chart (1) created from Table (1) to show the percentage of the occurrence of OSV order as opposed to SOV order in three categories of relative measurement of the RD between the S and O. The chart shows the occurrence of OSV over SOV is highest when the RD of S > the RD of O, next highest when the RD of S = the RD of O, and lowest when the RD of S < the RD of O.

Chart (1) Occurrence of OSV over SOV



These results indicate a relationship between the initial position and lower RD, which conforms to the Japanese speaker's given-before-new preference, evidenced by the experimental results of Ferreira and Yoshita (2003) (cf. Section 1.3.5.). Moreover, the differences between S>O (RD) and S<O (RD) in Table (1) is statistically significant (p=0.004), while the difference between S>O (RD) and S=O (RD) as well as the difference between S=O (RD) and S<O (RD) are not (p=0.454) for the former and p=0.1074 for the latter).

The relationship between the initial position and lower RD, shown in Table (1), as well as the statistical significance of the difference between S>O (RD) and S<O (RD) suggest that the relative RD of constituents is relevant to the choice of constituent ordering between SOV and OSV. That is, between the subject and the object in a given SOV construction clause[1], the constituent with a lower RD tends to appear earlier in a clause.

Now we observed the pattern between the factor RD and the earlier position. In what follows, I will show how algorithms based upon the pattern can predict one order over another. It is important to formulate an algorithm based on the pattern that the data indicates because algorithms can tell exactly under what conditions a particular linguistic form occurs, and under what conditions this particular linguistic form does occur. That is to say, algorithms define the pattern and the outcome of the pattern. To discuss the patterns is not enough: To state that 86.2% of SOV order occur when the S has a lower RD than the O, or there is a tendency for a constituent with a lower RD to appear first in a clause, is "to make an actuarial statement" (Tomlin 1995:521). These statements appeal to probability, but do not indicate the occurrences of particular linguistic forms in the real speech context. Functional linguists recognize the importance of the form-function correspondence that predictions by algorithms can project. In addition, it is important to

---

[1] The SOV construction denotes the syntactic structure of a clause containing a subject, a direcct objecct and a verb. The SOV construction clauses in my data exhibit five variations of word order, SOV, OSV, SVO, OVS and VSO (cf. Section 2.3).

formulate an algorithm for each pattern that each factor demonstrates so that we can compare the relative relevance of competing factors. Building algorithms and applying them to the data provides a clearer picture for understanding why we speak the way we do.

Having said that, based on the association between the initial position and a lower RD, shown in Table (1), we will formulate an algorithm using the RD as a factor to predict the choice of constituent ordering between SOV and OSV. The question remains as to which way the number of S=O (RD) should be merged, whether with S<O or with S>O (RD). Merging the number of S=O (RD) is necessary for an algorithm to work since the algorithm has to be applied to the whole data. Let us compare the prediction rate through two possible algorithms, Algorithm (1a) and Algorithm (1b). In Algorithm (1a), the number of S=O (RD) is merged with S>O (RD), and in Algorithm (1b) with S<O (RD).

Algorithm 1a: If the RD of S < RD of O, then SOV; else OSV.
Algorithm 1b: If the RD of S ≤ RD of O, then SOV, else OSV.

Algorithm (1a) states that if the RD of the subject is lower than that of the object, the order will be SOV. The algorithm also states that if the RD of the object is lower than the RD of the subject, or the RDs are the same for the subject and the object, the order will be OSV. Algorithm (1b) states that if the RD of the subject is lower than that of the object, or if the RD of the subject and the object are the same, the order will be SOV. The algorithm also states that if the RD of the subject is higher than that of the object, the order will be OSV.

When applying Algorithm (1a) to the 167 tokens in the present data, it correctly predicts 81 tokens in SOV order that are S<O (RD), 9 tokens in OSV order that are S=O (RD) and 15 tokens in OSV order that are S>O (RD). Thus, a total of 105 tokens are correctly predicted by Algorithm (1a). The result is summarized in Table (2). Table (2) shows that the ordering choice between SOV and OSV is correctly predicted by Algorithm (1a) in 62.9% of the data.

Table (2) Predictions for Algorithm 1a

RD	Correct Predictions		Incorrect Predictions		Total
S<O	SOV	81	OSV	13	94
O≤S	OSV	24	SOV	49	73
	Total	105	Total	62	167
	Percent	62.9%	Percent	37.1%	100%

When applying Algorithm (1b) to the 167 tokens in the present data, it correctly predicts 81 tokens in SOV order that are S<O (RD), 24 tokens in SOV order that are S=O (RD) and 15 tokens in OSV order that are S>O (RD). Thus, a total of 120 tokens are correctly predicted by Algorithm (1b). The result is summarized in Table (3). Table (3) shows that the ordering choice between SOV and OSV is correctly predicted by Algorithm (1b) in 71.9% of the data.

Table (3) Predictions for Algorithm 1b

RD	Correct Predictions		Incorrect Predictions		Total
S≤O	SOV	105	OSV	22	127
O<S	OSV	15	SOV	25	40
	Total	120	Total	47	167
	Percent	71.9%	Percent	28.1%	100%

The comparison of Tables (2) and (3) indicates that Algorithm (1b) has a higher correct prediction rate (62.9% for Algorithm (1a) and 71.9% for Algorithm (1b)). In addition, since the default order is SOV (in terms of frequency), it would be more natural to say that when the RD of the S and O are the same, the order will be SOV as Algorithm (1b) states than to say the order will be OSV as Algorithm (1a) states.

There is a third algorithm to be considered. Namely the algorithm that has a higher correct prediction rate than either Algorithm (1a) or (1b). I will call it Algorithm (1c).

Algorithm 1c: Always use SOV, never use OSV.

Since the percentage of tokens with SOV order in the present data is 77.8%, the correct prediction rate of Algorithm (1c) is always 77.8%. Because SOV order is more common than OSV order for any value of RD, TP or SW, none of these factors alone provides the basis for an algorithm. Despite the fact that it has a higher correct prediction rate than Algorithm (1a) or (1b), Algorithm (1c) is to be dismissed.

3.1.2. Topic persistence (TP)

Table (4) shows the relative measurement of TP in SOV and OSV orders.

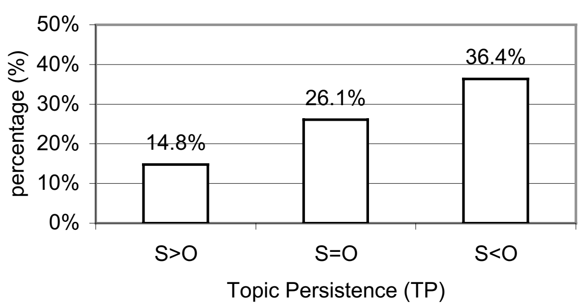
Table (4) Relative TP of S and O in SOV and OSV orders

TP	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	75	85.2	13	14.8	88	100
S=O	34	73.9	12	26.1	46	100
S<O	21	63.6	12	36.4	33	100
Total	130	77.8	37	22.2	167	100.0

Table (4) shows that the occurrence of SOV order is higher than OSV order in all three categories: S>O (TP) where the TP of the subject is higher than that of the object; S=O (TP) where the TP is the same for the subject and the object; and S<O (TP) where the TP of the subject is lower than that of the object. Table (4) also indicates that the choice of SOV order is highest when the TP of the subject is higher than the TP of the

object, next highest when the TP is the same for the subject and the object, and lowest when the TP of the object is higher than that of the subject. These results are displayed in Chart (2), created from Table (4), where the occurrence of OSV order as opposed to SOV order in the three categories is shown. Chart (2) shows that the choice of OSV is highest when TP of S < TP of O, next highest when TP of S = TP of O, and lowest when TP of S > TP of O.

Chart (2) Occurrence of OSV as opposed to SOV



While the difference between S>O (TP) and S<O (TP) is statistically significant (p=0.0125), the differences between S>O (TP) and S=O (TP) and between S<O (TP) and S=O (TP) are not (p=0.1598 for S>O (TP) and S=O (TP) and p=0.457 for S<O (TP) and S=O (TP)). Although the raw number of OSV tokens are comparable in all three categories: S>O (TP), S=O (TP) and S<O (TP), the percentage of occurrence for OSV clauses when S<O (TP) is higher than when S>O (TP), i.e. 36.4% for S<O (TP) and 14.8% for S>O (TP). That is, between the subject and the object in the SOV construction with SOV and OSV order, the constituent with higher TP tends to appear in the initial position. The association between the initial position and higher TP and its statistical significance provides the basis for an algorithm using the TP as a factor.

Now we can posit two possible algorithms:

Algorithm 2a: If the TP of S > TP of O, then SOV; else OSV.
Algorithm 2b: If the TP of S ≥ TP of O, then SOV, else OSV.

The number of S=O (TP) is merged with S<O (TP) in Algorithm (2a), and in Algorithm (2b) it is merged with S>O (TP). Algorithm (2a) states that if the TP of the subject is higher than that of the object, the order will be SOV, and if the TP of the object is higher than the TP of the subject or if the TPs are the same for the subject and the object, the order will be OSV. Algorithm (2b) states that if the TP of the subject is higher than that of the object, or if the TP is the same for the subject and the object, the order will be SOV, and if the TP of the object is higher than the TP of the subject, the order will be OSV.

When Algorithm (2a) is applied to the 167 tokens in the data, 75 tokens of SOV order with S>O (TP), 12 tokens of OSV order with S=O (TP) and 12 tokens of OSV order with S<O (TP) are correctly predicted. The result of these predictions is summarized in Table (5). Table (5) shows that the ordering choice between SOV and OSV is correctly predicted by Algorithm (2a) in 59.3% of the data.

Table (5) Predictions for Algorithm 2a

TP	Correct Predictions		Incorrect Predictions		Total
S>O	SOV	75	OSV	13	88
O≥S	OSV	24	SOV	55	79
	Total	99	Total	68	167
	Percent	59.3%	Percent	40.7%	100%

When Algorithm (2b) is applied to the 167 tokens in the data, 75 tokens of SOV order with S>O (TP), 34 tokens of SOV order with S=O (TP), and 12 tokens of OSV order with S<O (TP) are correctly predicted. The result of these predictions is

summarized in Table (6). Table (6) shows that Algorithm (2b) correctly predicts the ordering choice between SOV and OSV in 72.5% of the data.

Table (6) Predictions for Algorithm 2b

TP	Correct Predictions		Incorrect Predictions		Total
S≥O	SOV	109	OSV	25	134
O>S	OSV	12	SOV	21	33
	Total	121	Total	46	167
	Percent	72.5%	Percent	27.5%	100%

The comparison of Tables (5) and (6) indicates that Algorithm (2b) has a higher correct prediction rate (59.9% for Algorithm (2a) and 72.5% for Algorithm (2b)). In addition, since the default order is SOV, it would be more natural to say that when the TP of the S and O are the same, the order will be SOV as Algorithm (2b) states than to say the order will be OSV as Algorithm (2a) states.

The correct prediction rate by Algorithm (2b) is similar to that of Algorithm (1b), i.e. 71.9% for Algorithm (1b) and 72.5% for Algorithm (2b). However, as with Algorithm (1b), the success rate for Algorithm (2b) is less than the algorithm that says to use only SOV and never use OSV (Algorithm (1c)), so it is of limited value.

The single variables RD and TP might not be strong factors alone, but they might become so when both variables are combined. I will look at the effects of these two combined variables in Section 3.2.

3.1.3. Syntactic weight (SW)

Table (7) shows the number of tokens of SOV and OSV orders in five categories according to the relative SW of S and O. For Table (7), the following categorizations apply.

Syntactic weight (SW)

S >> O	syntactic weight of S	–	syntactic weight of O	=	4 or more
S > O	syntactic weight of S	–	syntactic weight of O	=	$1 \leq 3$
S = O	syntactic weight of S	–	syntactic weight of O	=	0
S < O	syntactic weight of O	–	syntactic weight of S	=	$1 \leq 3$
S << O	syntactic weight of O	–	syntactic weight of S	=	4 or more

Table (7) Relative Syntactic Weight of S and O in SOV and OSV orders

SW	SOV		OSV		Total	
	#	%	#	%	#	%
S>>O	7	70.0	3	30.0	10	100
S>O	51	86.4	8	13.6	59	100
S=O	27	77.1	8	22.9	35	100
S<O	25	75.8	8	24.2	33	100
S<<O	20	66.7	10	33.3	30	100
Total	130	77.8	37	22.2	167	100

In all five categories, S>>O (SW), S>O (SW), S=O (SW), S<O (SW), and S<<O (SW), the occurrence of SOV is higher than OSV. This result suggests that SOV is a default order and OSV is a marked order. The choice of SOV is the highest when S>O (SW), and the choice of OSV is the highest when S<<O (SW). Table (7) also shows a clear trend that the occurrence of SOV increases as the subject becomes longer than the object, and the choice of OSV increases as the object becomes longer than the subject. Table (8) presents the same data in Table (7) in three categories instead of five categories.[2] I will repeat here the three way categorization specified in the beginning of Chapter 3.

Syntactic weight (SW)

S > O	syntactic weight of S	is higher than	syntactic weight of O
S = O	syntactic weight of S	is the same as	syntactic weight of O
S < O	syntactic weight of O	is higher than	syntactic weight of S

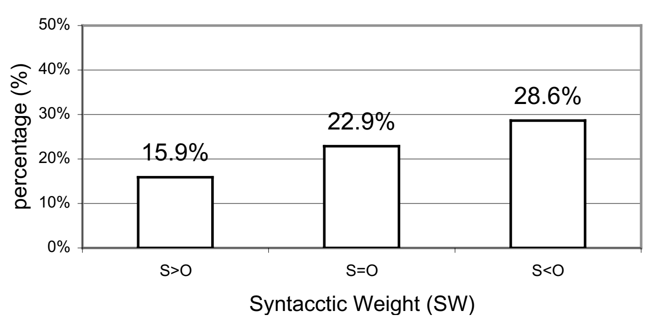
[2] The three way categorization can eliminate the anomality of 3 tokens (30%) in the category of S>>O (SW) which can be interpreted as a statistical glitch.

Table (8) Relative Syntactic Weight of S and O in SOV and OSV orders

SW	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	58	84.1	11	15.9	69	100
S=O	27	77.1	8	22.9	35	100
S<O	45	71.4	18	28.6	63	100
Total	130	77.8	37	22.2	167	100

Table (8) shows that the occurrence of SOV order is highest when the SW of the subject is higher than the SW of the object, next highest when the SW is the same for the subject and the object, and lowest when the SW of the object is higher than the SW of the subject. However, none of the differences in Table (8) is statistically significant. I.e. the difference between S>O (SW) and S<O (SW) is p=0.094, the difference between S>O (SW) and S=O (SW) is p=0.427, and the difference between S<O (SW) and S=O (SW) is p=0.636. The data in Table (8) was used to yield Chart (3), which shows the occurrence of OSV order as opposed to SOV order. Chart (3) indicates a clear trend that the occurrence of OSV order decreases as the subject grows longer than the object.

Chart (3) Occurrence of OSV order as opposed SOV order



These results shown in Table (8) and Chart (3) indicate the relationship between the initial position and higher SW. This relationship provides the basis for an algorithm using the SW as a factor to predict the ordering choice between SOV and OSV. That is, between the subject and the object in a given SOV construction clause with SOV and OSV order, the constituent with higher SW tends to occur in the initial position.

Now we consider two possible algorithms using the SW.

Algorithm 3a: If the SW of S > SW of O, then SOV; else OSV.
Algorithm 3b: If the SW of S ≥ SW of O, then SOV, else OSV.

In Algorithm (3a), the number of S=O (SW) is merged with S<O (SW), and in Algorithm (3b) it is merged with S>O (SW). Algorithm (3a) states that if the SW of the subject is higher than that of the object, the order will be SOV, and if the SW of the object is higher than the SW of the subject or if the subject and the object have the same SW, the order will be OSV. Algorithm (3b) states that if the SW of the subject is higher than the SW of the object, or if the subject and the object have the same SW, the order will be SOV, and the order will be OSV only when the SW of the object is higher than the SW of the subject.

When Algorithm (3a) is applied to the 167 tokens in the data, it correctly predicts the ordering choice between SOV and OSV for a total of 84 tokens, in 58 tokens of SOV with S>O (SW), 8 tokens of OSV with S=O (SW) and 18 tokens of OSV with S<O (SW). Table (9) summarizes the result of these predictions. Table (9) shows that Algorithm (3a) correctly predicts the ordering choice between SOV and OSV in 50.3% of the data.

Table (9) Predictions for Algorithm 3a

SW	Correct Predictions		Incorrect Predictions		Total
S>O	SOV	58	OSV	11	69
O≥S	OSV	26	SOV	72	98
	Total	84	Total	83	167
	Percent	50.3%	Percent	49.7%	100%

When Algorithm (3b) is applied to the 167 tokens in the data, it correctly predicts 58 tokens of SOV order with S>O (SW), 27 tokens of SOV order with S=O (SW) and 18 tokens of OSV order with S<O (SW). The result of the predictions is summarized in Table (10). Table (10) shows that Algorithm (3b) correctly predicts the ordering choice between SOV and OSV for 61.7% of the data.

Table (10) Predictions for Algorithm 3b

SW	Correct Predictions		Incorrect Predictions		Total
S≥O	SOV	85	OSV	19	104
O>S	OSV	18	SOV	45	63
	Total	103	Total	64	167
	Percent	61.7%	Percent	38.3%	100%

The comparison of Tables (9) and (10) indicates that Algorithm (3b) has a higher correct prediction rate (50.3% for Algorithm (3a) and 61.7% for Algorithm (3b)). In addition, since the default order is SOV, it would be more natural to say that when the SW of the S and O are the same, the order will be SOV than to say the order will be OSV. The result of correct predictions by Algorithm (3b) supports Hawkins' argument that in left-branching languages such as Japanese, preposing a heavy constituent facilitates the language processing, as well as Yamashita and Chang's experimental results.

The correct prediction rate by Algorithm (3b) is the lowest among the success rates of the algorithm using the single variables RD, TP and SW (72.5% for Algorithm (2b) using the TP, 71.9% for Algorithm (1b) using the RD and 61.7% for Algorithm (3b) using the SW). Moreover, the success rate of Algorithm (3b) is lower than that of Algorithm (1c) that says "Always use SOV and never use OSV". As far as the correct prediction rate is concerned, Algorithm (1c) is superior to any algorithm that we have discussed so far.

Algorithm (1c), however, does not provide a satisfactory basis for choosing between SOV and OSV. The problem with all three algorithms based on single variables (1b, 2b, and 3b) is none of these variables individually can tell us when OSV order is used, since SOV is more common than OSV for all three conditions for all three variables. There is no real correspondence of form and function.

3.2. Two variables

In the previous section, we learned that algorithms using the single variables RD, TP, and SW could predict the ordering choice between SOV and OSV orders to some extent. However, the algorithms using the RD, TP or SW were not as successful as the algorithm that says “Always use SOV, never use OSV” (i.e. Algorithm 1c), which has a 78.1% of success rate of predictions. In this section, I will examine whether the interaction of two variables can produce a better result in terms of predicting the ordering choice between SOV and OSV. The interaction I will look at is between RD and TP, RD and SW and TP and SW.

3.2.1. RD and TP

Table (11) shows the interaction of RD and TP. The table gives the number of tokens in SOV and OSV orders and the percentage of occurrence of OSV order as opposed to SOV order under the nine conditions. The nine conditions are based on the interaction of relative RD and SW: (1) S>O (RD) and S<O (TP), (2) S>O (RD) and S=O (TP), (3) S>O (RD) and S>O (TP), (4) S=O (RD) and S<O (TP), (5) S=O (RD) and S=O (TP), (6) S=O (RD) and S>O (TP), (7) S<O (RD) and S<O (TP), (8) S<O (RD) and S=O

(TP), and (9) S<O (RD) and S>O (TP). Table (11) reads, for example, that for the condition (1) S>O (RD) and S <O (TP), i.e. when the RD of the subject is higher than that of the object and when the TP of the subject is lower than that of the object, the number of tokens in SOV order (first column) is 7, the number of tokens in OSV order (second column) is 8, the total number of tokens in SOV and OSV orders (third column) is 15, and the percentage of occurrence of OSV order as opposed to SOV order (fourth column) is 53.3%.

Table (11) Interaction of RD and TP

	S>O (RD)				S=O (RD)				S<O (RD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	7	8	15	53.3	2	3	5	60.0	12	1	13	7.7
S=O	3	3	6	50.0	9	3	12	25.0	22	6	28	21.4
S>O	15	4	19	21.1	13	3	16	18.8	47	6	53	11.3

It should be noted that the individual numbers in Table (11) are small, and thus most of the differences are not statistically significant.

Table (11) shows that under the condition where the RD of S = the RD of O, the occurrence of OSV order is highest when TP of O > TP of S (i.e. 60% for S<O (TP), 25% for S=O (TP), and 18.8% for S>O (TP)). This result indicates that when the RD is not relevant, the TP plays a role in determining the word order. The table also shows that under the condition where TP of S = TP of O, the occurrence of OSV order is highest when the RD of S > the RD of O (i.e. 50% for S>O (RD), 25% for S=O (RD), and 21.4% for (S<O (RD)). This result indicates that when the TP is not relevant, the RD plays a role in determining the word order. Table (11) also indicates that the higher TP and the lower RD are complementarily working for predicting SOV order. For instance, under the

condition where the RD of S > the RD of O (i.e. when the RD favors OSV), the occurrence of SOV order is highest when the TP of S > TP of O (i.e. when the TP favors SOV), i.e. 78.9% for S>O (TP), 50% for S=O (TP) and 46.7% for S<O (TP). Similarly, under the condition where the TP of S < the TP of O (i.e. when the TP favors OSV), the occurrence of SOV order is highest when the RD of S < the RD of O (i.e. when the RD favors SOV), i.e. 92.3% for S<O (RD), 46.7% for S>O (RD) and 40% for S=O (RD). Likewise, under the condition when the TP of S > the TP of O (i.e. when the TP favors SOV), the occurrence of OSV is highest when the RD of S > the RD of O (i.e. the RD favors OSV), i.e. 21.2 % for S>O (RD), 18.8% for S=O (RD) and 11.3% for S<O (RD). It is not the case, however, for the TP in predicting OSV order. For example, under the condition when the RD of S < the RD of O (i.e. when the RD favors SOV), the occurrence of OSV is lowest when the TP of O > the TP of S (the TP favors OSV), i.e. 7.7% for S<O (TP), 11.3% for S>O (TP) and 21.4% for S=O (TP).

Now we will try to formulate the algorithm to predict the ordering choice between SOV and OSV. The first algorithm we will try is the one that simply combines Algorithms (1b) and (2b), formulated in section 3.1. The combined Algorithms are restated as Algorithm (4) as follows.

Algorithm (4): If the RD of S ≤ the RD of O, or if the TP of S ≥ the TP of O, then SOV; else OSV.

Algorithm (4) states that if the RD of the subject is lower than or the same as the RD of the object, or if the TP of the subject is higher than or the same as the TP of the object, the order will be SOV, and elsewhere the order will be OSV. Thus, Algorithm (4) predicts OSV order only when the RD of S > the RD of O and the TP of S > the TP of O. Table (12) illustrates the word orders that Algorithm (4) predicts.

Table (12) Word Order that Algorithm (4) predicts

<table>
<tr><td colspan="2" rowspan="2">Algorithm (4)</td><td colspan="3">RD</td></tr>
<tr><td>S>O</td><td>S=O</td><td>S<O</td></tr>
<tr><td rowspan="3">TP</td><td>S<O</td><td>OSV</td><td colspan="2" rowspan="3">SOV</td></tr>
<tr><td>S=O</td><td rowspan="2"></td></tr>
<tr><td>S>O</td></tr>
</table>

The result of predictions by Algorithm (4) is given in Table (13). Table (13) shows that Algorithm (4) correctly predicts the ordering choice between SOV and OSV in 78.4% of the data. The correct prediction rate by Algorithm (4) is higher than 77.8% by Algorithm (1c) that states “Always use SOV, never use OSV”.

Table (13) Predictions for Algorithm (4)

Algorithm (4)	Correct	Incorrect	Total
Predicts SOV	123	29	152
Predicts OSV	8	7	15
Total	131	36	167
Percent	78.4%	21.6%	100%

Now we consider a second algorithm using RD and TP.

Algorithm (5): Use SOV if
(a) RD of S < RD of O or
(b) TP of S > TP of O o
(c) RD of S = RD of O and TP of S = TP of O
Otherwise, use OSV.

Algorithm (5) says that if either RD or TP favors SOV, then SOV is used, and if neither RD nor TP favors SOV, then SOV is used, and that OSV is used only if one of the two factors favors OSV while neither favors SOV. Algorithm (5) is based on the idea that since SOV is the default order, it makes more sense if it is predicted when either measure favors SOV or when neither measure favors either SOV or OSV, and OSV is used only if some factor favors OSV without anything favoring SOV. Table (14) shows the word orders predicted by Algorithm (5).

Table (14) Word Order that Algorithm (5) predicts

Algorithm (5)		RD		
		S>O	S=O	S<O
TP	S<O	OSV	OSV	SOV
	S=O	OSV	SOV	SOV
	S>O	SOV	SOV	SOV

The result of predictions by Algorithm (5) is summarized in Table (15).

Table (15) Predictions for Algorithm (5)

Algorithm (5)	Correct	Incorrect	Total
Predicts SOV	118	23	141
Predicts OSV	14	12	26
Total	132	35	167
Percent	79.0%	21.0%	100%

Algorithm (5) correctly predicts the ordering choice between SOV and OSV in 79% of the data. The success rate of predictions for Algorithm (5) is higher than that for Algorithm (4) (79% for Algorithm (5) and 78.4% for Algorithm (4)). Moreover, the correct prediction rate by Algorithm (5) is higher than that by Algorithm (1c) that says "Always use SOV, never use OSV" (77.8% for Algorithm (1c)).

3.2.2. RD and SW

Table (16) shows the number of tokens in SOV and OSV orders and the percentage of occurrence of OSV order as opposed to SOV order under the nine conditions defined by the interaction of relative RD and SW. The nine conditions are: (1) S>O (RD) and S<O (SW), (2) S>O (RD) and S=O (SW), (3) S>O (RD) and S>O (SW), (4) S=O (RD) and S<O (SW), (5) S=O (RD) and S=O (SW), (6) S=O (RD) and S>O

(SW), (7) S<O (RD) and S<O (SW), (8) S<O (RD) and S=O (SW), and (9) S<O (RD) and S>O (SW).

Table (16) Interaction of RD and SW

	S>O (RD)				S=O (RD)				S<O (RD)			
SW	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	3	6	50.0	6	5	11	45.5	36	10	46	21.7
S=O	8	5	13	38.5	4	1	5	20.0	15	2	17	11.8
S>O	14	7	21	33.3	14	3	17	17.6	30	1	31	3.2

Table (16) shows that under the condition when the RD of S and O are the same, the occurrence of OSV order is higher when the SW of O is higher than the SW of S (45.5% for S<O (SW), 20% for S=O (SW) and 17.6% for S>O (SW)). This result indicates that when the RD is not relevant, the SW play a role in determining the word order. The table also shows that under the condition when the SW of S and O are the same, the occurrence of OSV order is highest when the RD of O is smaller than the RD of S (38.5% for S>O (RD), 20% for S=O (RD), and 11.8% for S<O (RD)). This result indicates that when the SW is not relevant, the RD plays a role in determining the word order. Table (16) indicates that the lower RD and the higher SW are complementarily working for predicting the ordering choice between SOV and OSV. For instance, under the condition when the RD of S is higher than the RD of O (i.e. when the RD favors OSV order), the occurrence of SOV is highest when the SW of S is higher than the SW of O (i.e. when the SW favors SOV), i.e. 66.7% for S>O (SW), 61.5% for S=O (SW) and 50% for S<O (SW). Similarly, under the condition when the SW of S is lower than the SW of O (i.e. the SW favors OSV order), the occurrence of SOV is highest when the RD of S is lower than the RD of O (the RD favors SOV), i.e. 78.3% for S<O (RD), 54.5% for S=O (RD)

and 50% for S>O (RD). Moreover, under the condition when the RD of S is lower than the RD of O (i.e. when the RD favors SOV), the occurrence of OSV is highest when the SW of O is higher than the SW of S (i.e. when the SW favors OSV), i.e. 21.7% for S<O (SW), 11.8% for S=O (SW) and 3.2% for S>O SW), and when the SW of S is higher than the SW of O (i.e. when the SW favors SOV), the occurrence of OSV is highest when the RD of O is lower than the RD of S (i.e. when the RD favors OSV), i.e. 33.3% for S>O (RD), 17.6% S=O (RD) and 3.2% for S<O (RD).

Now we will try an algorithm that simply combines Algorithms (1b) and (3b). The combined Algorithms are restated as Algorithm (6) as follows.

Algorithm (6): If the RD of S ≤ the RD of O, or if the SW of S ≥ the SW of O, then SOV; else OSV.

Algorithm (6) states that if the RD of the subject is lower than or the same as the RD of the object, or if the SW of the subject is higher than or the same as the SW of the object, the order will be SOV, and elsewhere OSV. That is, Algorithm (6) implies “Use OSV only when the RD of S > the RD of O and the SW of S < the SW of O”. Table (17) shows the word orders predicted by Algorithm (6).

Table (17) Word Order that Algorithm (6) predicts

Algorithm (6)		RD		
		S>O	S=O	S<O
SW	S<O	OSV		
	S=O		SOV	
	S>O			

The result of predictions by Algorithm (6) is summarized in Table (18).

Table (18) Predictions for Algorithm (6)

Algorithm (4)	Correct	Incorrect	Total
Predicts SOV	127	34	161
Predicts OSV	3	3	6
Total	130	37	167
Percent	77.8%	22.2%	100%

Table (18) shows that Algorithm (6) correctly predicts the ordering choice between SOV and OSV in 77.8% of the data. The success rate of prediction for Algorithm (6) is the same as that for Algorithm (1c) that states "Always use SOV, never use OSV".

Let us consider a second algorithm that is analogous to Algorithm (5).

Algorithm (7): Use SOV if
(a) RD of S < RD of O or
(b) SW of S > SW of O or
(c) RD of S = RD of O and SW of S = SW of O
Otherwise, use OSV.

Algorithm (7) says that if either RD or SW favors SOV, then SOV is used, and if neither RD nor SW favors SOV, then SOV is used, and that OSV is used only if one of the two factors favors OSV while neither favors SOV. The word orders predicted by Algorithm (7) are illustrated in Table (19).

Table (19) Word Order that Algorithm (7) predicts

Algorithm (7)		RD		
		S>O	S=O	S<O
SW	S<O	OSV	OSV	SOV
	S=O	OSV	SOV	SOV
	S>O	SOV	SOV	SOV

Applying the predictions shown in Table (19) to the 167 tokens in Table (16), we can calculate the number of correct and incorrect predictions by Algorithm (7). The prediction of Algorithm (7) is summarized in Table (20).

Table (20) Predictions for Algorithm (7)

Algorithm (7)	Correct	Incorrect	Total
Predicts SOV	113	24	137
Predicts OSV	13	17	30
Total	126	41	167
Percent	75.4%	24.6%	100%

Table (20) shows that Algorithm (7) correctly predicts the ordering choice between SOV and OSV in 75.4% of the data. This prediction rate is lower than that by Algorithm (6) (75.4% for Algorithm (7) and 77.8% for Algorithm (6)). Moreover, the success rate of predictions for Algorithm (7) is lower than that for Algorithm (1c) that says to only use SOV and never to use OSV (77.8% for Algorithm (1c)). In sum, the algorithms using the RD and SW together are less successful than those using the RD and TP together.

3.2.3. TP and SW

Table (21) shows the number of tokens in SOV and OSV orders and the percentage of occurrence of SOV order as opposed to SOV order under the nine conditions defined by the interaction of relative TP and SW. The nine conditions are: (1) S>O (SW) and S<O (TP), (2) S>O (SW) and S=O (TP), (3) S>O (SW) and S>O (TP), (4) S=O (SW) and S<O (TP), (5) S=O (SW) and S=O (TP), (6) S=O (SW) and S>O (TP), (7) S<O (SW) and S<O (TP), (8) S<O (SW) and S=O (TP), and (9) S<O (SW) and S>O (TP).

Table (21) Interaction of TP and SW

	S>O (SW)				S=O (SW)				S<O (SW)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	12	3	15	20.0	4	6	10	60.0	5	3	8	37.5
S=O	11	3	14	21.4	8	0	8	0.0	15	9	24	37.5
S>O	35	5	40	12.5	15	2	17	11.8	25	6	31	19.4

Table (21) shows that under the condition when the SW of S and O are the same, the occurrence of OSV is highest when the TP of O is higher than the TP of S (60% for S<O (TP), 11.8% for S>O (TP) and 0% for S=O (TP)). This result indicates that when the SW is not relevant, the TP plays a role in determining the word order. The table also shows that under the condition when the TP of S and O are the same, the occurrence of OSV is highest when the SW of O is higher than the SW of S (37.5% for S<O (SW), 21.4% for S>O (SW) and 0% for S=O (SW). This result indicates that when the TP is not relevant, the SW plays a role in determining the word order. In addition, Table (21) indicates that the higher SW and the higher TP are working complementarily for predicting the ordering choice between SOV and OSV. For example, under the condition when the TP of S is higher than the TP of O (i.e. when the TP favors SOV), the occurrence of OSV order is highest when the SW of O is higher than the SW of S (i.e. when the SW favors OSV), i.e. 19.4% for S<O (SW), 12.5% for S>O (SW) and 11.8% for S=O (SW). Similarly, under the condition when the TP of S is lower than the TP of O (i.e. when the TP favors OSV), the occurrence of SOV is highest when the SW of S is higher than the SW of O (i.e. when the SW favors SOV), i.e. 80% for S>O (SW), 62.5% for S<O (SW) and 40% for S=O (SW). Moreover, under the condition when the SW of S is lower than the SW of O (i.e. the SW favors OSV), the occurrence of SOV is highest when the TP of S is higher than the TP of O (i.e. when the TP favors SOV), i.e. 80.6% for S>O (TP), 62.5% for S<O (TP) and for S=O (TP). Also under the condition when the SW of S is higher than the SW of O (i.e. the SW favors SOV order), the occurrence of OSV is higher

when the TP of O is higher than the TP of S than when the TP of O is lower than the TP of S (i.e. when the TP favors OSV), i.e. 20% for S<O (TP) and 11.4% for S>O (TP).

We will now examine the algorithm that combines Algorithms (2b) and (3b). The combined Algorithms (2b) and (3b) are restated as Algorithm (8) here.

Algorithm (8): If the SW of S ≥ the SW of O, or if the TP of S ≥ the TP of O, then SOV; else OSV.

Algorithm (8) states that if the SW of the subject is higher than or the same as the SW of the object, or if the TP of the subject is higher than or the same as the TP of the object, the order will be SOV, and elsewhere OSV. Algorithm (8) predicts OSV only when the SW of S < the SW of O and the TP of S < the TP of O. Table (22) shows the word orders predicted by Algorithm (8).

Table (22) Word Order that Algorithm (8) predicts

Algorithm (8)		SW		
		S>O	S=O	S<O
TP	S<O	SOV		OSV
	S=O			
	S>O			

By projecting Table (22) over the data in Table (21), we can obtain the number of correct and incorrect predictions by Algorithm (8). The result is summarized in Table (23).

Table (23) Predictions for Algorithm (8)

Algorithm (8)	Correct	Incorrect	Total
Predicts SOV	125	34	159
Predicts OSV	3	5	8
Total	128	39	167
Percent	76.6%	23.4%	100%

Table (23) shows that Algorithm (8) predicts the word order correctly in 76.6% of the data.

This correct prediction rate is lower than that of Algorithm (1c) that states “Always use SOV, never use OSV” (76.6% for Algorithm (8) and 77.8% for Algorithm (1c)).

Let us consider an alternative algorithm that is analogous to Algorithms (5) and (7).

Algorithm (9): Use SOV if
(a) TP of S > TP of O or
(b) SW of S > SW of O or
(c) TP of S = TP of O and
SW of S = SW of O
Otherwise, use OSV.

Algorithm (9) says that if either TP or SW favors SOV, then SOV is used, and if neither TP nor SW favors SOV, then SOV is used, and that OSV is used only if one of the two factors favors OSV while neither favors SOV. Table (24) illustrates the word orders predicted by Algorithm (9).

Table (24) Word Order that Algorithm (9) predicts

Algorithm (9)		SW		
		S>O	S=O	S<O
TP	S<O	SOV	OSV	OSV
	S=O	SOV	SOV	OSV
	S>O	SOV	SOV	SOV

Applying the predictions shown in Table (19) to the 167 tokens in the data shown in Table (21), we can calculate the number of correct and incorrect predictions by Algorithm (9). The result is summarized in Table (25).

Table (25) Predictions for Algorithm (9)

Algorithm (9)	Correct	Incorrect	Total
Predicts SOV	106	19	125
Predicts OSV	18	24	42
Total	124	43	167
Percent	74.3%	25.7%	100%

Table (25) shows that Algorithm (9) correctly predicts the choice of constituent ordering between SOV and OSV for 74.3% of the data. This success rate of prediction is lower than that of Algorithm (8) (76.6% for Algorithm (8)). The correct prediction rates of both Algorithms (8) and (9) are lower than that of Algorithm (1c) that says "Always use SOV, never use OSV". Moreover, the algorithm using the SW and TP together did not do as well as those using the RD and TP together or the RD and SW together.

3.3. Three variables

In the previous section, the algorithms using two variables, RD and TP, RD and SW, and TP and SW, can predict the ordering choice between SOV and OSV with a higher success rate than the algorithms using the single variables. In this section, I will examine how successfully the algorithm using the combination of the three variables, RD, TP, and SW, can predict the ordering choice between SOV and OSV.

3.3.1. RD, TP and SW

Tables (26) to (28) show the interaction of relative RD and TP under the three conditions of relative SW; Table (26) considers the case when the subject is longer than the object, (27) when the SW of the subject and the object is the same, and (28) when the subject is shorter than the object.

Table (26) Interaction of RD and TP when the SW of S > the SW of O

	S>O (RD)				S=O (RD)				S<O (RD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	4	1	5	20.0	1	2	3	66.7	7	0	7	0.0
S=O	1	3	4	75.0	4	0	4	0.0	6	0	6	0.0
S>O	9	3	12	25.0	9	1	10	10.0	17	1	18	5.6

Table (27) Interaction of RD and TP when the SW of S = the SW of O

TP	S>O (RD)				S=O (RD)				S<O (RD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	2	4	6	66.7	1	1	2	50.0	1	1	2	50.0
S=O	1	0	1	0.0	2	0	2	0.0	5	0	5	0.0
S>O	5	1	6	16.7	1	0	1	0.0	9	1	10	10.0

Table (28) Interaction of RD and TP when the SW of S < the SW of O

TP	S>O (RD)				S=O (RD)				S<O (RD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	3	4	75.0	0	0	0	n/a	4	0	4	0.0
S=O	1	0	1	0.0	3	3	6	50.0	11	6	17	35.3
S>O	1	0	3	0.0	3	2	5	40.0	21	4	25	16.0

Now we try the algorithm that combines Algorithms (1b), (2b) and (3b). The combined Algorithms are restated as Algorithm (10).

Algorithm (10): If the RD of S ≤ the RD of O, if the TP of S ≥ the TP of O,
or if the SW of S ≥ the SW of O, then SOV; else OSV.

Algorithm (10) states: (i) if the RD of the subject is lower than RD of object, or if the RD of the subject and the object is the same, the order will be SOV and elsewhere OSV, or (ii) if the TP of the subject is higher than the TP of the object, or if the TP of the subject and the object is the same, the ordering is SOV and elsewhere OSV, or (iii) if the SW of the subject is higher than the SW of the object, or if the SW of the subject and the object is the same, the ordering is SOV and elsewhere OSV. Thus, Algorithm (10) implies that it predicts OSV if the RD of the subject is higher than the RD of the object, if the TP of the subject is lower than the TP of the object, and if the SW of the subject is lower than the SW of the object.

Applying Algorithm (10) to the 167 tokens in the data yields the result shown in Table (29).

Table (29) Predictions for Algorithm (10)

Algorithm (10)	Correct	Incorrect	Total
Predicts SOV	129	34	163
Predicts OSV	3	1	4
Total	132	35	167
Percent	79.0%	21.0%	100%

Table (29) shows that Algorithm (10) correctly predicted 79% of the data. This success rate is higher than any of the correct prediction rates of algorithms that we have discussed in the previous sections including Algorithm (1c) that states "Always use SOV, never use OSV', i.e. 79% for Algorithm (5) using the RD and TP, 77.8% for Algorithm (6) using the RD and SW, 76.6% for Algorithm (8) using the SW and TP, and 77.8% for Algorithm (1c).

Let us consider another algorithm based on the same concept for Algorithms (5), (7) and (9). Algorithm (10) that we considered above predicts SOV if the RD, TP or SW does not favors either SOV or OSV even when one of the three favors OSV. It is not logical to say that the order is SOV when some factor is saying OSV. Moreover, we need an algorithm that states exactly when it predicts SOV and when it predicts OSV. Therefore, I will try to formulate an alternative algorithm that can mend the disadvantage of Algorithm (10).

Algorithm (11): Use SOV if
(i) RD of S < RD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or
(iv) RD of S = RD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (11) says that if RD, TP or SW favors SOV, then SOV is used, and if neither RD, TP nor SW favors SOV, then SOV is used, and that OSV is used only if one of the three factors favors OSV while none of them favors SOV.

Tables (30)-(32) illustrate the word orders predicted by Algorithm (11). Table (30) corresponds to Table (26), Table (31) to Table (27), and Table (32) to Table (28).

Table (30) SW of S > SW of O: Word Order that Algorithm (11) predicts

Algorithm (11) SW of S > SW of O		RD S>O	RD S=O	RD S<O
TP	S<O	SOV	SOV	SOV
TP	S=O	SOV	SOV	SOV
TP	S>O	SOV	SOV	SOV

Table (31) SW of S = SW of O: Word Order that Algorithm (11) predicts

Algorithm (11) SW of S = SW of O		RD S>O	RD S=O	RD S<O
TP	S<O	OSV	OSV	SOV
TP	S=O	OSV	SOV	SOV
TP	S>O	SOV	SOV	SOV

Table (32) SW of S < SW of O: Word Order that Algorithm (11) predicts

Algorithm (11) SW of S < SW of O		RD S>O	RD S=O	RD S<O
TP	S<O	OSV	OSV	SOV
TP	S=O	OSV	OSV	SOV
TP	S>O	SOV	SOV	SOV

Applying the predictions shown in Tables (30)-(32) to the 167 tokens in the data in Tables (26)-(28) yields the results of predictions presented in Table (33).

Table (33) Predictions for Algorithm (11)

Algorithm (11)	Correct	Incorrect	Total
Predicts SOV	121	26	147
Predicts OSV	11	9	20
Total	132	35	167
Percent	79.0%	21.0%	100%

Table (33) shows that Algorithm (11) correctly predicts the ordering choice between SOV and OSV for 79% of the data. This success rate is the same as that of Algorithm (10). Moreover, the success rates of both Algorithms (10) and (11) are higher than the correct prediction rate of Algorithm (1c) that says "Always use SOV, never use OSV". Furthermore, the correct prediction rate of 79% by Algorithm (10) and (11) is the highest among the algorithms that we have discussed in sections 3.1 through 3.3.1.

While Algorithm (10) makes more correct predictions than Algorithm (11) for the default order SOV (129 SOV tokens for Algorithm (10) and 121 SOV tokens for Algorithm (11)), Algorithm (11) makes more correct predictions than Algorithm (10) for the unmarked order OSV (11 OSV tokens for Algorithm (11) and 3 OSV tokens for Algorithm (10)). Moreover, Algorithm (10) does not predict OSV unless all three factors RD, TP and SW favor OSV, while Algorithm (11) predicts OSV when one of RD, TP or SW favors OSV without any of these factors favoring SOV. Algorithm (11) makes more sense than Algorithm (10) from what the data shows and from a logical point of view.

3.3.2. The relative strength of the three variables RD, TP and SW

In this section, I will examine the relative strength of the three variables RD, TP and SW. The relative strength is determined by comparing the correct prediction rates of each variable. For example, when the RD favors SOV and the SW favors OSV while the TP favors neither, if the token is a SOV clause, it is the RD that plays a role in determining the word order, but not the SW. Consequently, we can say that the RD is stronger than the SW.

First, we will examine the cases where two variables are in conflict. Table (34) shows the number of tokens in SOV or OSV orders when the RD favors one order and the TP favors another. The difference in Table (34) is not statistically significant.

Table (34) The number of tokens when the RD and TP are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
RD favors SOV, TP favors OSV	12	92.3	1	7.7	13	100
RD favors OSV, TP favors SOV	15	78.9	4	21.1	19	100

Table (34) shows that the RD makes correct predictions over the TP for 92.3% of the cases in conflict while the TP makes correct predictions over the RD for 78.9% of the cases in conflict.

Table (35) shows the number of tokens in SOV or OSV orders when the RD favors one order and the SW favors another. The difference in Table (35) is not statistically significant.

Table (35) The number of tokens when the RD and SW are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
RD favors SOV, SW favors OSV	36	78.3	10	21.7	46	100
RD favors OSV, SW favors SOV	14	66.7	7	33.3	21	100

Table (35) shows that the RD makes correct predictions over the SW for 78.3% of the cases in conflict whereas the SW makes correct predictions over the RD for 66.7% of the cases in conflict.

Table (36) shows the number of tokens in SOV and OSV when the TP favors one order and the SW favors another. The difference in Table (36) is not statistically significant.

Table (36) The number of tokens when the TP and SW are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
TP favors SOV, SW favors OSV	25	80.6	6	19.4	31	100
TP favors OSV, SW favors SOV	12	80.0	3	20.0	15	100

Table (36) shows that the TP makes correct predictions over the SW for 80.6% of the cases in conflict, and the SW makes correct predictions over the TP for 80% of the cases in conflict.

Next, we will look at the case where two variables conflict with one variable. Table (37) shows the number of SOV and OSV tokens when the RD favors one order while both TP and SW favor another. The difference in Table (37) is not statistically significant.

Table (37) The number of tokens when the RD and the TP & SW are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
RD favors SOV, TP and SW favor OSV	4	100.0	0	0.0	4	100
RD favors OSV, TP and SW favor SOV	9	75.0	3	25.0	12	100

Table (37) shows that the algorithm based on the RD makes correct predictions over the algorithm based on the TP and SW for 100% of the cases in conflict, and the algorithm based on the TP and SW makes correct predictions over the algorithm based on the RD for 75% of the cases in conflict.

Table (38) shows the number of SOV and OSV tokens when the TP favors one order while both RD and SW favor another. The difference in Table (38) is not statistically significant.

Table (38) The number of tokens when the TP and the RD & SW are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
TP favors SOV, RD and SW favor OSV	1	100.0	0	0.0	1	100
TP favors OSV, RD and SW favor SOV	7	100.0	0	0.0	7	100

Table (38) shows that the correct prediction rate for the algorithm based on the TP that overrides the algorithm based on the RD and SW is comparable to the rate for the algorithm based on the RD and SW that overrides the algorithm based on the TP.

Table (39) shows the number of SOV and OSV tokens when the SW favors one order while both RD and TP favor another. The difference in Table (39) is not statistically significant.

Table (39) The number of tokens when the SW and the RD & TP are in conflict

factors in conflict	SOV		OSV		Total	
	#	%	#	%	#	%
SW favors SOV, RD and TP favor OSV	4	80.0	1	20.0	5	100
SW favors OSV, RD and TP favor SOV	21	84.0	4	16.0	25	100

Table (39) shows that the algorithm based on the SW makes correct predictions over the algorithm based on the RD and TP for 80% of the cases in conflict while the algorithm based on the RD and TP makes correct predictions over the algorithm based on the SW for 84% of the cases in conflict.

Table (40) summarizes the results. The symbol ">" in A>B indicates that A is a better predictor than B, and "=" in A=B indicates that the relative strength between A and B is equal.

Table (40) The relative strength of factors

Tables	Relative Strength
34	RD > TP
35	RD > SW
36	TP > SW
37	RD > TP & SW
38	TP = RD & SW
39	RD & TP > SW

The data in Tables (34)-(39) yields results that the RD has a higher prediction rate than the TP which has a higher prediction rate than the SW. Table (40) shows that the RD is a better predictor than the TP or SW, either as a single variable or when it is combined with the TP. Due to the fact that the number of conflict cases my data produced was so small, I could not get statistically significant results in this area. Within my data, the result did indicate that the RD is overall the strongest, followed by the TP, and then by the SW. The fact that the number of conflict cases was so small suggests that those factors are part of a system functioning together rather than acting as an individual factor.

3.4. Initial constituents in NPs

In my data, more than half of the objects in OSV order are first mentions. As I went through the data, I realized that there were quite a few tokens where a clause-initial noun phrase had a high RD, but where the noun phrase contained a noun phrase at the beginning of the larger noun phrase which had a lower RD. Clauses (1) and (2) below are examples of these tokens. In (1), the initial constituent is a proper name.

(1)

22-12	K-san	no	koe	o	atasi	ga	mane-site-mita	kedo
	K-Mr	GEN	voice	ACC	1sg	NOM	impersonation-doTE-see.PST	though

‘I mimicked K’s voice, though.’

The object in (1) is in the form of Noun 1 *no* Noun 2 (N1 GEN N2) 'N2 of N1'.[3] The object as a whole is a first mention, the initial constituent of the object noun phrase, *K-san* 'Mr. K', has an RD of 1. In (2) the initial constituent is a demonstrative adjective.

(2)

20-29	sono	e	o	boku	kaita-n-desu	yo
	that	painting	ACC	1sg.male	paint.PST-NMLZ-COP	FP
	'I drew the picture of that.'					

The object noun phrase in (2) is led by a demonstrative adjective *sono* 'that'. Although the object as a whole 'the picture of that' is a first mention, the referent of the initial constituent of the object noun phrase, *sono*, is mentioned in the immediately preceding clause.

Examples (1) and (2) demonstrate the high activation status of the clause-initial constituents of the clause-initial noun phrases. In order to determine if the highly activated status of the initial constituents in clause-initial noun phrases has any influence on the ordering choice, I will introduce a notion of "Extended referential distance" (hereafter ERD).

Extended referential distance (ERD) is defined as follows.

If the NP is simple, the ERD is the RD.
If the NP is compound, the ERD is the RD of the modifier.
If the NP is complex, the ERD is the lowest RD of any NP inside the clause.

"Simple nouns" refer to nouns, demonstrative nouns, or pronouns. "Compound noun phrases" are the noun phrases where the head noun and the dependent noun are connected by the genitive *no* or the noun phrases modified by demonstrative adjectives or

---

[3] Japanese is a dependent marking language (Nichols 1986).

demonstrative adverbs. "Complex noun phrases" are nominalized phrases or noun phrases containing relative clauses. The initial noun phrase in Clauses (1) and (2) are examples of compound noun phrases. The ERD for (1) is the RD of *K-san*, "Mr. K", and for (2) the ERD is the RD of *sono* 'that'. Clauses (3) and (4) are examples of complex noun phrases. The ERD of (3), for example, would be the RD of *M* "M" or the RD of *zyoyuu* "actress", whichever lower.

(3) Complex Noun Phrase: Nominalized Phrase

M	ga	zyoyuu	ni	naroo	to	omotteita	no
M	NOM	actress	DAT	become.VOL	QT	think.PPG.PST	NMLZ

"That M was thinking of becoming an actress."

(4) Complex Noun Phrase: Relativized Noun Phrase

suki-na	hito	ga	yareba-ii
like-PrNom	people	NOM	do.BE-good

'Those who like (it) should do'.

Among tokens containing compound or complex noun phrases, there are cases where the ERD is relevant to making correct predictions for the choice between SOV and OSV. I will explain this phenomenon, using Example (1) which I repeat here as (5). Clause (5) is an OSV token.

(5)

22-12	K-san	no	koe	o	atasi	ga	manesite-mita	kedo
	K-Mr	GEN	voice	ACC	1sg	NOM	imitate.TE-see.PST	though

'I mimicked K's voice.'

The object noun phrase, *K-san no koe* "Mr. K's voice" is a first mention. The subject *atasi* "I" has an RD of 20+. The relative RD of the subject and the object in (5) favors SOV order because the RD of S < the RD of O. Thus, the algorithm using the RD as a factor incorrectly predicts the word order of (5). However, when the ERD is taken into account, the algorithm using the ERD as a factor correctly predicts the word order of (5) because the ERD of the object is the RD of the initial constituent of the object noun phrase, *K-san* 'Mr. K', which is RD=1, and the ERD of the subject is the RD of the subject, which is RD=20+ (i.e. the ERD of O < the ERD of S).

Table (41) shows the number compound or complex subjects or objects in SOV and OSV orders.

Table (41) The number of tokens containing compound or complex NPs

Type of NPs	compound or complex NP		simple noun		Total	
	#	%	#	%	#	%
S in SOV	29	22.3	101	77.7	130	100
O in SOV	58	44.6	72	55.4	130	100
O in OSV	23	62.2	14	37.8	37	100
S in OSV	10	27.0	27	73.0	37	100

Table (41) shows that compound or complex NPs outnumber simple nouns only with the O in OSV. Seventeen compound or complex subjects and objects in SOV order occur in the same token. Seven compound or complex objects and subjects in OSV occur in the same token. That is, 70 tokens with SOV (i.e. 29+58-17=70) and 26 tokens with OSV (i.e. 23+10-7=26) contain either the subject or the object that is compound or complex noun phrase, or both subject and object that are compound or complex noun phrases. Table (42) shows the number of SOV and OSV tokens that contain compound or complex NPs.

Table (42) The number of tokens in SOV and OSV orders, sorted by the types of noun phrases

Type of noun phrases	SOV		OSV		Total	
	#	%	#	%	#	%
simple noun	60	84.5	11	15.5	71	100
compound or complex noun phrase	70	72.9	26	27.1	96	100
Total	130	77.8	37	22.2	167	100

Table (42) shows 53.8% of SOV tokens (70/130 tokens) and 70.1% of OSV tokens (26/37 tokens) contain compound or complex noun phrases. It means that when we measure ERD instead of RD, 57.5% of the data (96/167 tokens) will have a different outcome of relative measurements between constituents.

Table (43) shows the relative measurement of ERD of the subject and the object in tokens containing compound or complex noun phrases. For comparison, the relative measurement of RD for the same data is shown in Table (44).

Table (43) Relative ERD of S and O in SOV and OSV tokens containing compound or complex noun phrase subjects or objects.

ERD	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	14	45.2	17	54.8	31	100
S=O	16	66.7	8	33.3	24	100
S<O	40	97.6	1	2.4	41	100
Total	70	72.9	26	27.1	96	100

Table (44) Relative RD of S and O in SOV and OSV tokens containing compound or complex noun phrase subjects or objects.

RD	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	15	60.0	10	40.0	25	100
S=O	11	61.1	7	38.9	18	100
S<O	46	83.6	9	16.4	55	100
Total	72	73.5	26	26.5	98	100

Table (43) shows that when the ERD of S < the ERD of O, the word order is overwhelmingly SOV (40 out of 41 cases), and when the ERD of S > the ERD of O, it is more likely to be OSV (by 17 to 14). Tables (43) and (44) show that while SOV order occurs more frequently than OSV even when the RD favors OSV, OSV order occurs more frequently than SOV when the ERD favors OSV. These results suggest the potential use of ERD as a factor for predicting the constituent ordering choice.

Table (45) shows the relative ERD of the subject and the object in 167 tokens in the data. For comparison, the relative RD of the subject and the object in 167 tokens is shown in Table (46)

Table (45) Relative ERD of S and O in SOV and OSV orders

ERD	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	26	54.2	22	45.8	48	100
S=O	29	74.4	10	25.6	39	100
S<O	75	93.8	5	6.3	80	100
Total	130	77.8	37	22.2	167	100

Table (46) Relative RD of S and O in SOV and OSV orders

RD	SOV		OSV		Total	
	#	%	#	%	#	%
S>O	25	62.5	15	37.5	40	100
S=O	24	72.7	9	27.3	33	100
S<O	81	86.2	13	13.8	94	100
Total	130	77.8	37	22.2	167	100

Table (45) shows that SOV order occurs more frequently than OSV at any value of relative ERD, suggesting that SOV is the default order. The table also shows that the proportion of SOV is highest when the ERD of S < the ERD of O (93.8%), next highest when the ERDs are the same for the S and O (74.4%), and lowest when the ERD of S < the ERD of O (54.2%). The difference between S=O (ERD) and S<O (ERD) is statistically significant ($p=0.0060$) while the differences between S>O (ERD) and S=O (ERD), and between S>O (ERD) and S<O (ERD) are not. However, similarly to the results from the RD measurement, the data in Table (45) shows the relationship between the initial element and low ERD.

Now, let us formulate an algorithm that is analogous to Algorithm (1b), but using the ERD instead of the RD.

Algorithm (12): If the ERD of S ≤ ERD of O, then SOV, else OSV.

When applying Algorithm (12) to the 167 tokens in the data, Algorithm (12) correctly predicts the word order in 74.3% of the data. The result of predictions of Algorithm (12) is summarized in Table (47).

Table (47) Predictions for Algorithm (12)

Algorithm (12)	Correct	Incorrect	Total
Predicts SOV	104	15	119
Predicts OSV	22	26	48
Total	126	41	167
Percent	75.4%	24.6%	100%

The success rate of predictions for Algorithm (12) is higher than for Algorithm (1b) using the RD as a factor (75.4% for Algorithm (12) and 71.9% for Algorithm (1b)). It is, however, still lower than Algorithm (1c), which states "Use only SOV, never use OSV", and which has a 77.8% success rate.

Now we examine the interaction of the three variables, ERD, TP and SW. Table (48) shows the interaction of relative ERD and TP under the condition when the subject is longer than the object, Table (49) when the SWs are the same for the subject and the object, and Table (50) when the object is longer than the subject.

Table (48) Interaction of ERD and TP when the SW of S > the SW of O

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	3	1	4	25.0	2	2	4	50.0	7	0	7	0.0
S=O	1	3	4	75.0	4	0	4	0.0	6	0	6	0.0
S>O	8	3	11	27.3	8	1	9	11.1	19	1	20	5.0

Table (49) Interaction of ERD and TP when the SW of S = the SW of O

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	2	5	7	71.4	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	2	0	2	0.0	5	0	5	0.0
S>O	5	1	6	16.7	1	0	1	0.0	9	1	10	10.0

Table (50) Interaction of ERD and TP when the SW of S < the SW of O

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	3	4	75.0	0	0	0	n/a	4	0	4	0.0
S=O	4	3	7	42.9	4	5	9	55.6	7	1	8	12.5
S>O	1	3	4	75.0	8	2	10	20.0	16	1	17	5.9

The algorithm we will apply to the 167 tokens presented in Tables (48)-(50) is an algorithm using the ERD, TP and SW, which I will call Algorithm (13). Algorithm (13) is analogous to Algorithm (11).

Algorithm (13): Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or
(iv) ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

When we apply the predictions of Algorithm (11) illustrated in Tables (30)-(32)[4] to the 167 tokens shown in Tables (48)-(50), Algorithm (13) yields an 80.9% success rate of predictions. The result of predictions by Algorithm (13) is summarized in Table (51).

Table (51) Predictions for Algorithm (13)

Algorithm (13)	Correct	Incorrect	Total
Predicts SOV	118	21	139
Predicts OSV	16	12	28
Total	134	33	167
Percent	80.2%	19.8%	100%

Table (51) shows that Algorithm (13) correctly predicts the word order for 80.2% of the data. This success rate is higher than that of Algorithm (10) and Algorithm (11) using the

[4] The predictions of Algorithm (11) illustrated in Tables (30)-(32) are valid for Algorithm (13); only it is the ERD instead ot the RD.

RD, TP and SW (79.2% for both Algorithms). In addition, Algorithm (13) makes more correct predictions than Algorithm (1c), which states "Always use SOV, never use OSV".

3.5. Particle *wa*

My statistical data shows that when the subject is marked with the particle *wa*, the order is overwhelmingly SOV order, and when the direct object is marked with *wa*, the order is most likely to be OSV order (see Tables (18)-(21) in Section 2.4. in Chapter 2). The property of *wa*-marked constituents has been discussed in terms of thematic readings and contrastive readings (e.g. Kuno 1973). Some researchers (e.g. Clancy and Downing 1987, Shimojo 2005) claim that the contrastive reading is the primary function of *wa*.

Some instances of wrong predictions by Algorithm (13) contain the subject marked with *wa* or the object marked with *wa*. In what follows, I will examine whether *wa*-marking as a factor has an influence on the constituent ordering choice.

3.5.1. Objects marked with *wa*

Table (52) shows the number of tokens with objects marked with *wa* and objects with a marking other than *wa* in SOV and OSV orders.

Table (52) The number of tokens with O = *wa* and O ≠ *wa*

Marking	SOV		OSV		Total	
	#	%	#	%	#	%
O = *wa*	6	40.0	9	60.0	15	100.0
O ≠ *wa*	124	81.6	28	18.4	152	100.0
total	130	77.8	37	22.2	167	100.0

Table (52) shows that the use of OSV as opposed to SOV is significantly higher with *wa*-marked objects than with objects not marked with *wa* ($p=0.0009$).

Based on the analysis of conversational data in Japanese, using Givón's referential distance measurement, Shimojo (2005) argues that *wa*-marked objects are contrastive, regardless of their position or the measurement of RD. In order to determine whether there is any correlation between *wa*-marked objects and RD in my data, I have categorized the tokens into four different ranges of RD. Table (53) shows the RD of the objects for all 167 tokens in the data when the object is marked with *wa* and when not marked with *wa*.

Table (53) RD of Object when the O is marked with *wa* and when not marked with *wa.*

	RD									
Marking	1-4		5≤10		11≤20+		FM		Total	
	#	%	#	%	#	%	#	%	#	%
O = wa	2	13.3	2	13.3	3	20.0	8	53.3	15	100
O ≠ wa	38	25.0	12	7.9	16	10.5	86	56.6	152	100
Total	40	24.0	14	8.4	19	11.4	94	56.3	167	100

Table (53) shows no statistically significant relationship between the *wa*-marking of objects and RD. The numbers in Table (53) are further divided into SOV and OSV tokens in order to see if the data shows any difference in SOV and OSV clauses. Table (54) shows the object token distribution in SOV clause and Table (55) in OSV clause.

Table (54) RD of O in SOV when the O is marked with *wa* and when not marked with *wa.*

	RD									
SOV	1-4		5≤10		11≤20+		FM		Total	
	#	%	#	%	#	%	#	%	#	%
O = wa	0	0.0	0	0.0	2	33.3	4	66.7	6	100
O ≠ wa	26	21.0	12	9.7	16	12.9	70	56.5	124	100
Total	26	20.0	12	9.2	18	13.8	74	56.9	130	100

Table (55) RD of O in OSV when the O is marked with *wa* and when not marked with *wa.*

OSV	RD									
	1-4		5≤10		11≤20+		FM		Total	
	#	%	#	%	#	%	#	%	#	%
O = wa	2	22.2	2	22.2	1	11.1	4	44.4	9	100
O ≠ wa	12	42.9	0	0.0	0	0.0	16	57.1	28	100
Total	14	37.8	2	5.4	1	2.7	20	54.1	37	100

The trend in the percentages for first mentions in Tables (54) and (55) suggest a relationship between the word order and the *wa*-marking for FM, i.e. a larger number of non-initial constituents marked with *wa* (i.e. the O-*wa* in SOV) than initial constituents marked with *wa* (i.e. the O-*wa* in OSV) are first mentions. While the relation is not statistically significant, the very small number of tokens of objects marked by *wa* in SOV clauses would make it very difficult to achieve statistical significance.

Now I will formulate an algorithm based on the statistics in Table (52). The algorithm is Algorithm O-*wa*.

Algorithm O-*wa*: If the object is marked with the particle *wa*, then OSV. Algorithm O-*wa* only applies to a set of data that contains tokens with a *wa*-marked object. Algorithm O-*wa*, when it is applied to the 15 tokens containing a *wa*-marked object, correctly predicts the word order of 9 OSV tokens and incorrectly predicts the word order of 6 SOV tokens in Table (52).

In order to achieve the maximal correct prediction rate, I will try to integrate Algorithm O-*wa* with Algorithm (13). The mechanism of integrating the two algorithms is to apply Algorithm O-*wa* to a set of data containing an object marked with *wa* (i.e. 15 tokens) first, and then to apply Algorithm (13) to the rest of the data (i.e. 152 tokens that do not contain a *wa*-marked object). In this way, Algorithm O-*wa* overrides Algorithm

(13). The algorithm that I will try is Algorithm (14). Algorithm (14) consists of two parts (A) and (B).

Algorithm (14):

(A) Use OSV if the O is marked with *wa*.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

(A) predicts correctly 9 OSV tokens, and incorrectly predicts 6 SOV tokens. Now, we will apply (B) to the 169 tokens that do not contain an object marked with *wa*. Tables (56)-(58) shows the number of SOV and OSV tokens according to the interaction of relative ERD, TP and SW for the 169 tokens.

Table (56) Interaction of ERD and TP when the SW of S > the SW of O for the 152 tokens which do not contain the object marked with wa

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	0	3	0.0	2	2	4	50.0	6	0	6	0.0
S=O	1	3	4	75.0	4	0	4	0.0	6	0	6	0.0
S>O	8	2	10	20.0	8	1	9	11.1	19	0	19	0.0

Table (57) Interaction of ERD and TP when the SW of S = the SW of O for the 152 tokens which do not contain the object marked with wa

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	2	5	7	71.4	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	2	0	2	0.0	5	0	5	0.0
S>O	5	1	6	16.7	1	0	1	0.0	8	0	8	0.0

Table (58) Interaction of ERD and TP when the SW of S < the SW of O for the 152 tokens which do not contain the object marked with *wa*

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	2	3	66.7	0	0	0	n/a	4	0	4	0.0
S=O	4	2	6	33.3	2	3	5	60.0	7	0	7	0.0
S>O	1	3	4	75.0	8	2	10	20.0	14	1	15	6.7

When we apply (B) to the 169 tokens in Tables (56)-(58), we obtain the result of predictions shown in Table (59).

Table (59) Predictions for (B) in Algorithm (14)

Predictions	Correct	Incorrect	Total
Predicts SOV	114	16	130
Predicts OSV	12	10	22
Total	126	26	152
Percent	82.9%	17.1%	100%

When integrating the results of prediction by (A) and (B) in Algorithm (14), the result of predictions for Algorithm (14) is obtained. The result is shown in Table (60).

Table (60) Predictions for Algorithm (14)

Algorithm (14)	Correct	Incorrect	Total
Predicts SOV	114	16	130
Predicts OSV	21	16	37
Total	135	32	167
Percent	80.8%	19.2%	100%

The change from Algorithm (13) to Algorithm (14) adds five new instances of correct predictions of OSV, but adds four new instances of incorrect predictions of OSV. In sum, the result of Algorithm (14) is one token better from Algorithm (13). Table (60) shows that Algorithm (14) correctly predicts the word order for 80.8% of the data, and this success rate is higher than an 80.2% success rate of Algorithm (13).

3.5.2. Subjects marked with *wa*

Table (61) shows the number of tokens in SOV and OSV clauses that contain a subject marked with *wa* and a subject marked with other particles than *wa*.

Table (61) The number of tokens with S = *wa* and S ≠ *wa*

Marking	SOV		OSV		Total	Total
	#	%	#	%	#	%
S = wa	47	94.0	3	6.0	50	100
S ≠ wa	83	70.9	34	29.1	117	100
Total	130	77.8	37	22.2	167	100

Table (61) shows that the percentage of subjects in SOV clauses as opposed to in OSV clauses is significantly higher when the subject is marked with *wa* than when the subject is marked with other particle than *wa* (p=0.0008). In order to see whether there is any relationship between *wa*-marking on subjects and the RD, the data in Table (61) was organized according to the value of RD for the subject into four categories, RD = 1-4, RD = 5≤10, RD = 11≤20+ and first mention (FM). The result is shown in Table (62).

Table (62) RD of Subject when the S is marked with *wa* and when not marked with *wa.*

	RD									
Marking	1-4		5≤10		11≤20+		FM		Total	
	#	%	#	%	#	%	#	%	#	%
S = wa	24	48.0	12	24.0	11	22.0	3	6.0	50	100
S ≠ wa	39	33.3	13	11.1	20	17.1	45	38.5	117	100
Total	63	37.7	25	15.0	31	18.6	48	28.7	167	100

Table (62) shows that *wa*-marked subjects tend to be non-first mention more often than subjects not marked with *wa* (p=0.0000). The data in Table (62) was further grouped together in SOV and OSV tokens in order to see whether the relationship between *wa*-marking on the subject and non-first mention is significant in both SOV and OSV. Table (63) shows the result in SOV clauses, and Table (64) in OSV clauses.

Table (63) The number of first mentions and non-first mentions in SOV clauses when the subject is marked with *wa*

SOV	Non-First Mentions		First Mentions		Total	
	#	%	#	%	#	%
S = wa	45	95.7	2	4.3	47	100
S ≠ wa	57	68.7	26	31.3	83	100
Total	102	78.5	28	21.5	130	100

Table (64) The number of first mentions and non-first mentions in OSV clauses when the subject is marked with *wa*

OSV	Non-First Mentions		First Mentions		Total	
	#	%	#	%	#	%
S = wa	2	66.7	1	33.3	3	100
S ≠ wa	15	44.1	19	55.9	34	100
Total	17	45.9	20	54.1	37	100

The percentage of non-first mention subjects marked with *wa* is significantly higher than first mention subjects marked with *wa* in SOV clauses (p=0.0002), while it is not in OSV clauses. This result suggests that non-first mention is rather relevant to the clause-initial constituents than to *wa*-marking of subjects.

Given the results in Table (61), we can hypothesize an algorithm using *wa*-marking of subjects as a factor. The algorithm is Algorithm S-*wa*.

Algorithm S-*wa*: If the subject is marked with *wa*, then SOV.

Algorithm S-*wa* only applies to a set of data that contains tokens with a *wa*-marked subject. When we apply Algorithm S-*wa* to the 50 tokens that contain a *wa*-marked subject, the algorithm correctly predicts the word order for 47 SOV tokens and incorrectly predicts for 3 OSV tokens.

We will use the same strategy that we used for Algorithm O-*wa* in order to achieve the maximal correct prediction rate for the algorithm which involves the S-*wa* as a factor. I will try to integrate Algorithm S-*wa* with Algorithm (13). The revised algorithm is Algorithm (15). Algorithm (15) consists of two parts, (A) and (B). The first part (A) is Algorithm S-*wa,* and the second part (B) is Algorithm (13). The two algorithms are combined in a way that Algorithm S-*wa* overrides Algorithm (13). The process of applying Algorithm (15) is to apply (A) to a set of data containing the subject marked with *wa* (i.e. 50 tokens) first, and then to apply (B) to the rest of the data (i.e. 117 tokens).

Algorithm (15):

(A) Use SOV if the S is marked with *wa*.
Otherwise,
(B) Use SOV if
- (i) ERD of S < ERD of O or
- (ii) TP of S > TP of O or
- (iii) SW of S > SW of O,
- or (iv)
  - ERD of S = ERD of O and
  - TP of S = TP of O and
  - SW of S = SW of O.

Otherwise, use OSV.

(A) predicts correctly the word order for 47 SOV tokens and incorrectly predicts for 3 OSV tokens. Now, we will apply (B) to the 134 tokens that do not contain a subject marked with *wa*. Tables (65)-(67) show the number of SOV and OSV tokens according to the interaction of relative ERD, TP and SW for the 117 tokens.

Table (65) Interaction of ERD and TP when the SW of S > the SW of O for the 134 tokens which do not contain the subject marked with *wa*

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	1	4	25.0	2	2	4	50.0	1	0	1	0.0
S=O	1	3	4	75.0	4	0	4	0.0	2	0	2	0.0
S>O	4	3	7	42.9	6	1	7	14.3	15	1	16	6.3

Table (66) Interaction of ERD and TP when the SW of S = the SW of O for the 134 tokens which do not contain the subject marked with *wa*

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	1	5	6	83.3	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	1	0	1	0.0	2	0	2	0.0
S>O	3	0	3	0.0	1	0	1	0.0	5	1	6	16.7

Table (67) Interaction of ERD and TP when the SW of S < the SW of O for the 134 tokens which do not contain the subject marked with *wa*

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	1	3	4	75.0	0	0	0	n/a	4	0	4	0.0
S=O	2	3	5	60.0	3	5	8	62.5	4	1	5	20.0
S>O	0	3	3	100.0	6	1	7	14.3	9	0	9	0.0

When we apply (B) to the 117 tokens in Tables (65)-(67), we obtain the result of predictions shown in Table (68).

Table (68) Predictions for (B) in Algorithm (15)

Predictions	Correct	Incorrect	Total
Predicts SOV	75	18	93
Predicts OSV	16	8	24
Total	91	26	117
Percent	77.8%	22.2%	100%

When we merge the results of predictions of (A) and (B) together, we can obtain the result of predictions for Algorithm (15). The result is shown in Table (69).

Table (69) Predictions for Algorithm (15)

Predictions	Correct	Incorrect	Total
Predicts SOV	122	21	143
Predicts OSV	16	8	24
Total	138	29	167
Percent	82.6%	17.4%	100%

Algorithm (15) adds four correct predictions of SOV that were incorrect prediction of OSV by Algorithm (13), resulting in four more correct predictions than Algorithm (13). Table (69) shows that Algorithm (15) correctly predicts the constituent ordering for

82.6% of the data. The success rate of 82.6% is higher than 80.2% by Algorithm (13) or 80.8% by Algorithm (14).

3.5.3. S-*wa* and O-*wa*

In this section, I will try an algorithm using the subject marked with *wa* and the object marked with *wa*. In my data, there are 3 tokens in which both subject and object are marked with *wa*. Considering that SOV is the default order for *wa*-marked constituents, and the 3 tokens in which both subject and object are marked with *wa* are all SOV clauses, we will formulate an algorithm that predicts SOV if both S and O are marked with *wa*. The algorithm we will try is Algorithm (16).

Algorithm (16)

(A) Use SOV if the S is marked with *wa* or
Use OSV if the O is marked with *wa* or
Use SOV if both S and O are marked with *wa*.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Table (70) shows the number of tokens containing *wa*-marked constituents.

Table (70) The number of tokens containing wa-marked constituents

	S = wa	O = wa	both S & O = wa	Total
SOV	44	3	3	50
OSV	3	9	0	12
Total	47	12	3	62

When applying (A) in Algorithm (16) is applied to the 62 tokens containing *wa*-marked constituents, (A) in the algorithm correctly predicts a total of 47 SOV tokens (i.e. 44 SOV and 3 SOV tokens containing both S and O marked with *wa*), and 9 OSV tokens containing a *wa*-marked object marked. It also incorrectly predicts 3 SOV tokens containing a *wa*-marked object and 3 OSV tokens containing a *wa*-marked subject. Table (71) shows the predictions by (A) in Algorithm (16).

Table (71) Predictions for(A) in Algorithm (16) when applied to the 62 tokens containing wa-marked constituents

Prediction	Correct	Incorrect	Total
Predicts SOV	47	3	50
Predicts OSV	9	3	12
Total	56	6	62
Percent	90.3%	9.7%	100%

Now we will apply Algorithm (13), which is (B) in Algorithm (16), to the 105 tokens that do not contain *wa*-marked constituents. Tables (72)-(74) show the interaction of ERD, TP and SW for the 105 tokens.

Table (72) Interaction of ERD and TP when the SW of S > the SW of O for the 105 tokens that do not contain wa-marked constituents

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	0	3	0.0	2	2	4	50.0	1	0	1	0.0
S=O	1	3	4	75.0	4	0	4	0.0	2	0	2	0.0
S>O	4	2	6	33.3	6	1	7	14.3	15	0	15	0.0

Table (73) Interaction of ERD and TP when the SW of S = the SW of O for the 105 tokens that do not contain wa-marked constituents

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	5	6	83.3	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	1	0	1	0.0	2	0	2	0.0
S>O	3	0	3	0.0	1	0	1	0.0	5	0	5	0.0

Table (74) Interaction of ERD and TP when the SW of S < the SW of O for the 105 tokens that do not contain wa-marked constituents

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	2	3	66.7	0	0	0	n/a	4	0	4	0.0
S=O	2	2	4	50.0	2	2	4	50.0	4	0	4	0.0
S>O	0	3	3	100	6	1	7	14.3	8	0	8	0.0

Applying Algorithm (13) to the 105 tokens shown in Tables (72)-(74) yields the result of predictions shown in Table (75).

Table (75) Predictions for Algorithm (13) when applied to the 105 tokens that do not contain wa-marked constituents

Prediction	Correct	Incorrect	Total
Predicts SOV	74	13	87
Predicts OSV	11	7	18
Total	85	20	105
Percent	81.0%	19.0%	100%

When we merge the results of predictions by (A) and by (B) in Algorithm (16), we can obtain the result of predictions for Algorithm (16). The result is shown in Table (76).

Table (76) Predictions for Algorithm (16)

Algorithm (16)	Correct	Incorrect	Total
Predicts SOV	121	16	137
Predicts OSV	20	10	30
Total	141	26	167
Percent	84.4%	15.6%	100%

The change from Algorithm (13) to Algorithm (16) adds 3 SOV tokens and 4 OSV tokens of correct predictions. Table (76) shows s 84.4% of success rate for Algorithm (16), which is the highest so far. This success rate is higher than 80.8% for Algorithm (14) that uses the *wa*-marked object with Algorithm (13) or 82.6% for Algorithm (15) that uses the *wa*-marked subject with Algorithm (13).

3.6. Cohesion of the object and the verb

The cohesion between a constituent and its verb is discussed in Miyajima (1964) in reference to the scope of the constituent (see section 1.2.4). The scope of the constituent concerns with the semantic link between an argument and its predicate. When a constituent occurs with a specific verb with a high frequency, the scope of the constituent is narrow, i.e. the cohesion between the constituent and the verb is strong. On the other hand, when there are no such co-occurring restrictions, the scope is broad, i.e. the cohesion between the constituent and the verb is weak. In my data, the cohesion between the object and the verb is observed. Cohesive objects and verbs are often called idiomatic expressions (e.g. Fujii and Ono 2000). The cohesion of the object and the verb exhibits characteristics that the semantic unit of the object and the verb conveys a single concept, and that the object is often abstract and unspecific, thus non-referential[5]. Fujii and Ono (2000) also observed these characteristics in their study of zero marking on the direct object in Japanese. The parameter to determine whether the object and the verb are idiomatic is to test them using the WH question and answer pair. Consider the question and answer pairs in (6) and (7).

(6)

Q1:	nani	o	tukatta	no
	What	ACC	use.PST	Q
	'What did you use?'			

A1:	enpitu	o	tukatta.
	pencil	ACC	use.PST
	'I used a pencil.'		

---

[5] Fujii and Ono also have observed the tendency of zero marking on the object in idiomatic predicates.

A2?	ki	o	tukatta[6]
	Mind	ACC	use.PST
	'I was attentive.'		

(7)

Q2:	nani	o	tukatta	tte
	What	ACC	use.PST	QT
	'What did you say you used?'			

A2:	ki	o	tukatta
	Mind	ACC	use.PST
	'I was attentive.'		

A1 is a natural response to the question, 'What did you use?'. A2 is only possible in a specific context that a question provides. In (7), for example, the questioner Q uses Q2 when the questioner, Q, heard a part of an utterance by the respondent, A, and asked A to repeat the part of the utterance that Q missed. In other words, the referent of the object *nani* 'what' in Q2 is already active in the speaker's mind, and most importantly the speaker believes that it is also activated in the hearer's mind. In such a context, A2 is a natural response. When an expression only makes sense in a particular context, we can identify them as idiomatic expression.

Table (77) shows a list of objects and verbs in idiomatic predicates found in my data. There were no idiomatic predicates in OSV order in the present data.

[6] Sentence (6)-A2 is grammatical, only it is anomalous in the context.

Table (77) Cohesive objects and verbs: idiomatic predicates

Cohesive objects and verbs: idiomatic predicates			# of tokens		
Object	Verb	meaning	SOV	3NPs	Total
ki 'spirit'	*tukau* 'to use'	to be attentive	2	2	4
tyuui 'attention'	*harau* 'to pay'	to be careful	1		1
kasa 'umbrella'	*sasu* 'to pierce'	to use umbrella	1		1
kyoomi 'interest'	*motu* 'to carry'	to be interested in		2	2
kimoti 'feeling'	*soroeru* 'to put in order'	to be harmonious		1	1
koe 'voice'	*kakeru* 'hang'	to invite		1	1
meeru 'mail'	*utu* 'to strike'	to send mail		1	1
sei 'vigor'	*tukeru* 'attach'	to be more energetic		1	1
tikara 'power'	*awaseru* 'put together'	to cooperate		1	1
teikoo 'resistence'	*motu 'to carry'*	to resist		1	1
tosi 'age'	*toru* 'to take'	to age		1	1
		Total	4	11	15

In addition to the idiomatic predicates identified in (77), another type of cohesion was observed. It is the construction with verbal nouns and the verb *suru* 'do'. Verbal nouns are nominals that have the "ability to be verbalized with *suru*" (Uehara 1998:132), and the verb *suru* 'do' is called a light verb (e.g. Grimshaw and Mester 1988) when it occurs with verbal nouns. The light verb construction is defined as "OV compounds in which the combination of the direct object nominal and the verb is lexicalized" (Fujii and Ono 2000:9).

The verbal noun in the light verb construction can appear with or without the accusative *o*[7]. Example (8) shows the case when the verbal noun is marked with the accusative *o*, and Example (9) shows the case when the verbal noun is marked with the zero-particle. Moreover, there are cases where the compound of a verbal noun and the verb *suru* takes a direct object as in (10).

[7] Researchers such as Martin (1975) and Fujii and Ono (2000) differentiate the verbal nouns in the light verb construction that take the accusative *o*, VN-*o*-*suru*, (the unincorporated form in Fujii and Ono's phrasing) or zero marking, VN-*suru*, (the incorporated form in Fujii and Ono's phrasing), and those which appear only in VN-*suru*. Fujii and Ono report that in their data of spontaneous, informal conversations in Japanese, 77% of the verbal nouns in the light verb construction were in the incorporated form.

(8)	taroo	ga	benkyoo	o	suru	
	Taro	NOM	study	ACC	do	
	'Taro studies.'					

(9)	taroo	ga	benkyoo	suru	
	Taro	NOM	study	do	
	'Taro studies.'				

(10)	taroo	ga	eigo	o	benkyoo	suru
	Taro	NOM	English	ACC	study	do
	'Taro studies English.'					

Uehara (1998) analyzes the light verb construction, N-*o*-*suru* and N-*suru* in terms of a continuum from more compositional to more lexical sequences. According to Uehara, the N-*o*-*suru* construction is syntactically analyzable, but the N-*suru* construction loses its syntactical analyzability[8]. Moreover, the verb *suru* (the accentual contour Low-High) in the N-*o*-*suru* construction loses its accentual contour pattern only on an irregular basis

---

[8] Uehara uses tests to show that the N-*o-suru* construction is syntactically analyzable and the N-*suru* construction is not. The tests he proposes are the adverbial insertion test and the nominal modifier test. (1) and (2) are examples of the adverbial insertion test, and (3) and (4) demonstrate the nominal modifier test. Examples (1) to (4) are from Uehara (1998:151-2).

(1)	kopii	o	kyoo	suru	
	copy	ACC	today	do	
	'(I will) make a copy today.'				

(2)??	seikoo	kyoo	suru	
	success	today	do	
	'(I will) succeed today.'			

(3)	hon	no	kopii	o	suru
	book	GEN	copy	ACC	do
	'(I will) xerox a book.'				

(4)???	sakusen	no	seikoo-suru
	stratagem	GEN	success-do
	'(I will) be successful in my stratagem.'		

The N-*o*-*suru* construction, *kopii*-*o*-*suru* 'to make a copy', in (1) and (3) allows the insertion of a temporal adverbial, kyoo 'today', and a genitive *no*. On the other hand, the N-*suru* construction, *seikoo*-*suru* 'to success', in (2) and (4) does not allows the insertion.

in order to assimilate to the accentual pattern of the preceding mora. On the other hand, *suru* in the N-*suru* construction loses its accentual contour in most cases, although, Uehara claims that, discourse factors such as emphasis can override this phonological assimilation.

In this study, verbal nouns such as *benkyoo* ‘study’ in (8) and (9) are counted as direct objects because in the former example, it is marked with the accusative *o*, and in the latter example, the noun *benkyoo* ‘study’ is a sole entity that bears objecthood in the clause. In contrast, verbal nouns such as *benkyoo* ‘study’ in (10) are not counted as direct objects because there is a syntactically well-marked direct object in the same clause, i.e. *eigo* ‘English’.

The data in this study was sorted according to the occurrence of the idiomatic predicates and the light verb construction.[9] Table (78) shows the number of tokens with idiomatic predicates and the light verb construction in SOV and OSV.

Table (78) The occurrence of idiomatic predicates and the light verb construction in SOV and OSV clauses

parameter	SOV		OSV		Total	
	#	%	#	%	#	%
Light verb construction	35	85.4	6	14.6	41	100
Idiomatic predicates	4	100	0	0	4	100

In order to see whether the parameter, cohesive objects (i.e. idiomatic predicates and the light verb construction), is statistically significant, the number of idiomatic predicates and the light verb construction in Table (78) are put together as cohesive objects. The results are shown in Table (79).

[9] See Appendix A for the complete list of cohesive objects in the light verb construction identified in the current data.

Table (79) The number of tokens with and without cohesive objects

Factor	SOV		OSV		Total	
	#	%	#	%	#	%
with cohesive objects	39	86.7	6	13.3	45	100
without cohesive objects	91	74.6	31	25.4	122	100
Total	130	77.8	37	22.2	167	100

The difference between the number of tokens with cohesive objects and without cohesive objects is not statistically significant (p=0.140). Table (79), however, shows that 26.9% of the data contains a cohesive object (i.e. 45 out of 167), and that the proportion of SOV occurrence is slightly higher when the object is cohesive than when it is not cohesive.

Based on the result in Table (79), we can try an algorithm using the cohesive object as a factor.

Algorithm Cohesion: If the S occurs with cohesive O, then SOV.

Algorithm Cohesion only applies to a set of data that contains cohesive objects. When we apply Algorithm Cohesion to the 45 tokens that contain a cohesive object, Algorithm Cohesion correctly predict the word order for 39 SOV tokens, and incorrectly predicts for 6 OSV tokens.

Now we will integrate Algorithm Cohesion with Algorithm (13). The integrated algorithm is Algorithm (17).

Algorithm (17):

(A) Use SOV if the S occurs with cohesive O.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (17) consists of Algorithm Cohesion as (A) and Algorithm (13) as (B). In Algorithm (17), (A) overrides (B). The part (A) of Algorithm (17), henceforth (17A), correctly predicts the constituent ordering for 39 SOV tokens, and incorrectly predicts for 6 OSV tokens. The next step is to apply the part (B) of Algorithm (17), henceforth (17B), to the 122 tokens that do not contain a cohesive object. Tables (80)-(82) show the interaction of relative ERD, TP and SW for the 122 tokens.

Table (80) Interaction of ERD and TP when the SW of S > the SW of O for the 122 tokens that do not contain a cohesive object.

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	2	0	2	0.0	2	2	4	50.0	4	0	4	0.0
S=O	1	3	4	75.0	3	0	3	0.0	4	0	4	0.0
S>O	6	2	8	25.0	4	1	5	20.0	12	1	13	7.7

Table (81) Interaction of ERD and TP when the SW of S = the SW of O for the 122 tokens that do not contain a cohesive object.

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	4	5	80.0	0	0	0	n/a	2	0	2	0.0
S=O	1	0	1	0.0	0	0	0	n/a	3	0	3	0.0
S>O	4	1	5	20.0	1	0	1	0.0	7	1	8	12.5

Table (82) Interaction of ERD and TP when the SW of S < the SW of O for the 122 tokens that do not contain a cohesive object.

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	3	4	75.0	0	1	1	n/a	3	0	3	0.0
S=O	2	3	5	60.0	4	4	8	50.0	4	1	5	20.0
S>O	1	3	4	75.0	6	1	7	14.3	13	1	14	7.1

When we apply (B) in Algorithm (16) to the 122 tokens presented in Tables (80)-(82), it yields the result of predictions shown in Table (83).

Table (83) Predictions for (B) in Algorithm (17)

Predictions	Correct	Incorrect	Total
Predicts SOV	82	17	99
Predicts OSV	14	9	23
Total	96	26	122
Percent	78.7%	21.3%	100%

When we combine the results of predictions by (A) and (B) of Algorithm (16), we can obtain the prediction result for Algorithm (17). Table (84) shows the result.

Table (84) Predictions for Algorithm (17)

Algorithm (17)	Correct	Incorrect	Total
Predicts SOV	121	23	144
Predicts OSV	14	9	23
Total	135	32	167
Percent	80.8%	19.2%	100%

Algorithm (17) adds 3 correct predictions of SOV and 2 incorrect predictions of OSV, resulting in one token better than Algorithm (13). Table (84) shows that Algorithm (17) correctly predicts the word order for 80.8% of the data, which is the same as Algorithm (14) involving the O-*wa* as a factor. An 80.8% success rate is lower than 84.4% by Algorithm (16) involving the S-*wa* and O-*wa* as factors and 82.6% by Algorithm (15) involving the S-*wa* as a factor.

Table (85) presents relative measurement of the three factors, SW, TP and ERD, for the 45 tokens with cohesive objects.

Table (85) Relative measurement of the three factors for 45 tokens with cohesive objects

	# of tokens	order	SW	TP	ERD
a	1	SOV	S>O	S<O	S>O
b	3	SOV	S>O	S<O	S<O
c	1	SOV	S>O	S=O	S=O
d	2	SOV	S>O	S=O	S<O
e	2	SOV	S>O	S>O	S>O
f	4	SOV	S>O	S>O	S=O
g	7	SOV	S>O	S>O	S<O
h	1	SOV	S=O	S<O	S>O
i	2	SOV	S=O	S=O	S=O
j	2	SOV	S=O	S=O	S<O
k	1	SOV	S=O	S>O	S>O
L	2	SOV	S=O	S>O	S<O
m	1	SOV	S<O	S<O	S<O
n	2	SOV	S<O	S=O	S>O
o	3	SOV	S<O	S=O	S<O
p	2	SOV	S<O	S>O	S=O
q	3	SOV	S<O	S>O	S<O
r	1	OSV	S>O	S<O	S>O
s	1	OSV	S>O	S>O	S>O
t	1	OSV	S=O	S<O	S>O
u	1	OSV	S=O	S<O	S<O
v	1	OSV	S<O	S=O	S=O
w	1	OSV	S<O	S>O	S=O

Algorithm (17A) correctly predicts the order in 39 SOV tokens and incorrectly predicts the order in 6 OSV tokens. Table (85) shows that the 6 OSV tokens are favored by one or more of the three factors, SW, TP or ERD. Suppose we posit Algorithm (17a), which says that if the S occurs with cohesive objects, then use SOV unless one or more of the three other factors favor OSV, in which case we use SOV. Table (86) shows the result of predictions by Algorithm (17a) for the 45 tokens containing cohesive objects.

Table (86) Predictions for Algorithm (17a) for 45 tokens with cohesive objects

Algorithm (17a)	Correct	Incorrect	Total
Predicts SOV	20	0	20
Predicts OSV	6	19	25
Total	26	19	45
Percent	57.8%	42.2%	100%

Among the 39 SOV tokens containing cohesive objects, there are 19 tokens in which one or more of the three factors favor OSV. While Algorithm (17a) inverts the 6 OSV tokens that Algorithm (17) incorrectly predicts into correct predictions (cf. (r)-(w) in Table (85)), it also inverts 19 SOV tokens that Algorithm (17) correctly predicts into incorrect predictions (cf. (a), (b), (e), (h), (k), (m)-(q) in Table (85)). The success rate of 57.8% shown in Table (86) is lower than 86.7% success rate of Algorithm (17) for the 45 tokens containing cohesive objects (i.e. 39/45=86.7%).

What if we consider Algorithm (17b), which says that if the S occurs with cohesive objects, then use SOV unless two or more of the three factors, SW, TP or ERD, favor OSV. Algorithm (17b) correctly predicts 2 OSV tokens in which both TP and ERD favor OSV (cf. (r) and (t) in Table (85)) that Algorithm (17) incorrectly predicts. The algorithm also incorrectly predicts 5 SOV tokens (cf. (a), (h), (m) and (n) in Table (85)) that are correctly predicted by Algorithm (17). Table (86) shows the result of predictions by Algorithm (17b) for the 45 tokens with cohesive objects.

Table (87) Predictions for Algorithm (17b) for 45 tokens with cohesive objects

Algorithm (17b)	Correct	Incorrect	Total
Predicts SOV	34	4	38
Predicts OSV	2	5	7
Total	36	9	45
Percent	80.0%	20.0%	100%

Table (87) shows that the success rate for Algorithm (17b) is higher than Algorithm (17a), but it is lower than the prediction by Algorithm (17A).

3.7. Demonstratives

In Chapter 2, we have seen the correlation between demonstratives[10] and the initial position. In this section, I will examine whether demonstratives can be a factor for the choice of constituent ordering between SOV and OSV. Table (88) shows the number of tokens that contain demonstrative constituents. One OSV token contains both a demonstrative subject and a demonstrative object. Other tokens contain either a demonstrative subject or a demonstrative object.

Table (88) The number of tokens containing demonstrative constituents

	Only S is demonstrative		Only O is demonstrative		Both S and O are demonstrative		Total	
	#	%	#	%	#	%	#	%
SOV	7	87.5	13	43.3	0	0	20	51.3
OSV	1	12.5	17	56.7	1	100.0	19	48.7
Total	8	100	30	100	1	100	39	100

Table (88) shows a trend that when the subject is demonstrative, it occurs most likely in SOV clauses, and when the object is demonstrative, it is more likely to occur in OSV. This trend reassures a relationship between demonstratives and the clause-initial position.

3.7.1. Demonstrative objects

The number of demonstrative tokens in Table (88) is pulled out to create Table (89) that shows the number of tokens containing a demonstrative object and non-demonstrative object.

---

[10] Demonstratives in this study comprise demonstrative pronouns, nouns modified with demonstrative adjective and noun phrases containing demonstrative adverbs (cf. Section 2.4.6).

Table (89) The number of tokens containing a demonstrative object

Demonstratives	SOV		OSV		Total	
	#	%	#	%	#	%
O = demonstrative	13	41.9	18	58.1	31	100
O ≠ demonstrative	117	86.0	19	14.0	136	100
total	130	77.8	37	22.2	167	100

Table (89) shows that when the object is demonstrative, it is likely to occur in OSV clauses, and that when it is not demonstrative, it occurs most likely in SOV clauses. Based on the trend shown in Tables (88) and the data in (89), I will try to formulate an algorithm using the demonstrative object as a factor.

Algorithm O-demonstrative: If the O is demonstrative, then OSV.

Algorithm O-demonstrative only applies to a set of data that contains a demonstrative object. When we apply Algorithm O-demonstrative to the 31 tokens that contain a demonstrative object, the algorithm correctly predicts the word order for 18 OSV tokens and incorrectly predicts for 13 SOV tokens.

We now integrate Algorithm O-demonstrative with Algorithm (13) in order to see if the algorithm using the demonstrative object as a factor can improve the correct predictions on Algorithm (13). The algorithm we will try is Algorithm (18).

Algorithm (18):

(A) Use OSV if the O is demonstrative.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (18) is formulated as Algorithm O-demonstrative, i.e. (A), overrides Algorithm (13), i.e. (B). When we apply (18A) to the 30 tokens that contain a demonstrative object, it yields 18 correct predictions and 13 incorrect predictions. Now we can apply (18B) to the 136 tokens that do not contain a demonstrative object. Tables (90)-(92) show the interaction of relative ERD, TP and SW for the 137 tokens.

Table (90) Interaction of ERD and TP when the SW of S > the SW of O for the 137 tokens that do not contain a demonstrative object

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	0	3	0.0	2	0	2	0.0	7	0	7	0.0
S=O	1	0	1	0.0	4	0	4	0.0	6	0	6	0.0
S>O	5	0	5	0.0	7	0	7	0.0	18	1	19	5.3

Table (91) Interaction of ERD and TP when the SW of S = the SW of O for the 137 tokens that do not contain a demonstrative object

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	2	0	2	0.0	5	0	5	0.0
S>O	2	0	2	0.0	1	0	1	0.0	9	1	10	10.0

Table (92) Interaction of ERD and TP when the SW of S < the SW of O for the 137 tokens that do not contain a demonstrative object

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	4	0	4	0.0
S=O	4	3	7	42.9	4	5	9	55.6	6	1	7	14.3
S>O	1	2	3	66.7	7	2	9	22.2	16	0	16	0.0

When we apply Algorithm (18) to the 137 tokens shown in Tables (90)-(92), the algorithm produces the result of predictions shown in Table (93).

Table (93) Predictions for (B) Algorithm (18)

Predictions	Correct	Incorrect	Total
Predicts SOV	106	9	115
Predicts OSV	10	11	21
Total	116	20	136
Percent	85.3%	14.7%	100%

Now we merge the result of predictions by (18A), i.e. 18 correct predictions of OSV tokens and 13 incorrect predictions of SOV tokens, with the result by (18B) shown in Table (93). The result is shown in Table (94).

Table (94) Predictions for Algorithm (18)

Algorithm (18)	Correct	Incorrect	Total
Predicts SOV	106	9	115
Predicts OSV	28	24	52
Total	134	33	167
Percent	80.2%	19.8%	100%

The change from Algorithm (13) to Algorithm (18) adds 12 correct predictions of OSV and 12 incorrect predictions of SOV tokens. Table (84) shows that Algorithm (18) produces 80.2% of correct prediction rate, which is comparable to that Algorithm (13). The success rate is lower than 80.8% by Algorithm (14) involving the O-*wa*, 80.8% by Algorithm (17) involving cohesive objects, 82.6% by Algorithm (15) involving the S-*wa* and 84.4% by Algorithm (16) involving the S-*wa* and O-*wa*.

Table (95) shows the relative measurement of the three factors, SW, TP and ERD for the 31 tokens containing demonstrative objects.

Table (95) Relative measurement of the three factors for 31 tokens containing demonstrative objects

	# of tokens	order	SW	TP	ERD
a	5	SOV	S>O	S>O	S>O
b	1	SOV	S>O	S>O	S=O
c	1	SOV	S>O	S>O	S<O
d	3	SOV	S=O	S>O	S>O
e	1	SOV	S=O	S<O	S>O
f	1	SOV	S<O	S>O	S=O
g	1	SOV	S<O	S=O	S<O
h	1	OSV	S>O	S<O	S>O
i	3	OSV	S>O	S=O	S>O
j	3	OSV	S>O	S>O	S>O
k	1	OSV	S>O	S<O	S=O
L	1	OSV	S>O	S>O	S=O
m	4	OSV	S=O	S<O	S>O
n	1	OSV	S=O	S>O	S>O
o	2	OSV	S<O	S<O	S>O
p	1	OSV	S<O	S>O	S>O
q	1	OSV	S<O	S>O	S<O

Algorithm (18) correctly predicts 18 OSV tokens with demonstrative objects and incorrectly predicts 13 SOV tokens with demonstrative objects shown in Table (95). The data in Table (95) shows that one or more of the three factors, SW, TP and ERD, favor SOV in 12 out of 13 SOV tokens with demonstrative objects. Let us propose Algorithm (18a), which says that if the O is marked with a demonstrative, then use OSV unless two or more of the three factors, SW, TP and ERD, favor SOV, in which case we use SOV. Table (96) shows the result of applying Algorithm (18a) to the 31 tokens with demonstrative objects.

Table (96) Predictions for Algorithm (18a) applying to the 31 tokens with demonstrative objects

Algorithm (18a)	Correct	Incorrect	Total
Predicts SOV	7	6	13
Predicts OSV	13	5	18
Total	20	11	31
Percent	64.5%	35.5%	100%

Algorithm (18a) inverts 7 SOV tokens (cf. (a)-(c) in Table (95)) that Algorithm (18) incorrectly predicts into 7 correct predictions of SOV tokens. At the same time, Algorithm (18a) inverts 5 OSV tokens (cf. (j), (L), (q) in Table (95)) that Algorithm (18) correctly predicts into 5 incorrect predictions of OSV tokens. Table (96) shows 64.5% of success rate for Algorithm (18a), which does better than 58.1% for Algorithm (18), i.e. 18/31 tokens = 58.1%). Table (97) gives the result of predictions for Algorithm (18) with Algorithm (18a) when applied to 167 tokens in the data. The numbers in Table (97) are acquired by merging the result shown in Table (96) with that for Algorithm (18B) shown in Table (93).

Table (97) Predictions for Algorithm (18a) applying to 167 tokens of SOV and OSV orders

Algorithm (18a)	Correct	Incorrect	Total
Predicts SOV	113	15	128
Predicts OSV	23	16	39
Total	136	31	167
Percent	81.4%	18.6%	100%

Table (97) shows that Algorithm (18a) makes two more correct predictions than Algorithm (18).

Now, what if we posit Algorithm (18b), which says that if the O is marked with a demonstrative, then use OSV unless one or more of the three factors, SW, TP and ERD, favor SOV, in which case we use SOV. Table (98) shows the result of predictions for Algorithm (18b) when it is applied to the 31 tokens with demonstrative objects.

Table (98) Predictions for Algorithm (18b) applying to the 31 tokens with demonstrative objects

Algorithm (18b)	Correct	Incorrect	Total
Predicts SOV	12	12	24
Predicts OSV	6	1	7
Total	18	13	31
Percent	58.1%	41.9%	100%

Algorithm (18b) inverts 12 correct predictions of OSV by Algorithm (18) into incorrect predictions, and 12 incorrect predictions of SOV by Algorithm (18) into correct predictions (cf. (a)-(d), (f) and (g) in Table (95)). The result indicates that Algorithm (18b) will not improve Algorithm (18) when it is applied to the 167 tokens in the data.

We can try Algorithm (18c), which says that if the O is marked with a demonstrative, and if one or more of the three factors, SW, TP, ERD, favors OSV, then OSV. Table (99) shows the result of predictions for Algorithm (18c) when applied to the 31 tokens containing demonstrative objects.

Table (99) Predictions for Algorithm (18c) applying to the 31 tokens with demonstrative objects

Algorithm (18c)	Correct	Incorrect	Total
Predicts SOV	2	1	3
Predicts OSV	17	11	28
Total	19	12	31
Percent	61.3%	38.7%	100%

Algorithm (18c) inverts 2 incorrect predictions of SOV by Algorithm (18) into correct predictions (cf. (b) and (d) in Table (95)), and one correct predictions of OSV by Algorithm (18) into incorrect predictions (cf. (L) in Table (95)). The result indicates that Algorithm (18c) improves Algorithm (18) by one correct prediction. Algorithm (18c), however, is no better than Algorithm (18a), which does one token better than Algorithm (18c).

The relative measurements in Table (95) show that 11 out of 13 tokens of SOV are the case in which TP favors SOV. Let us try another algorithm based on the TP, i.e. Algorithm (18d), which says that if the O is marked with a demonstrative, then OSV unless TP favors SOV, in which case we use SOV. Table (100) shows the result of predictions for Algorithm (18d).

Table (100) Predictions for Algorithm (18d) applying to the 31 tokens with demonstrative objects

Algorithm (18d)	Correct	Incorrect	Total
Predicts SOV	11	7	18
Predicts OSV	11	2	13
Total	22	9	31
Percent	71.0%	29.0%	100%

Algorithm (18d) inverts 7 OSV tokens that Algorithm (18) correctly predicts into incorrect predictions (cf. (j), (L), (n), (p) and (q) in Table (95)), and 11 SOV tokens that Algorithm (18) incorrectly predicts into correct predictions (cf. (a)-(d) and (f) in Table (95)). The result indicates an improvement of 4 correct predictions from Algorithm (18), and a 71% success rate is the best among the alternative algorithms we have tried in this section, i.e. Algorithms (18a), (18b), (18c) and (18d). When the result in Table (100) is merged with the prediction result for Algorithm (18B), shown in Table (93), we can obtain the success rate for Algorithm (18) with Algorithm (18d) replacing Algorithm (18A), shown in Table (101).

Table (101) Predictions for Algorithm (18d) with Algorithm (18B) when applied to the 167 tokens in the data.

Predictions	Correct	Incorrect	Total
Predicts SOV	117	16	133
Predicts OSV	21	13	34
Total	138	29	167
Percent	82.6%	17.4%	100%

Table (101) shows 82.6% of success rate for Algorithm (18d) with Algorithm (18B). This success rate is as good as that for Algorithm (15) using the S-*wa* as a factor, but lower than 84.4% for Algorithm (16) using the S-*wa* and O-*wa* as factors.

3.7.2. Demonstrative subjects

Table (102) shows the number of SOV and OSV tokens that contain a demonstrative subject and non-demonstrative subject.

Table (102) The number of tokens containing a demonstrative subject

Demonstratives	SOV		OSV		Total	
	#	%	#	%	#	%
S = demonstrative	7	77.8	2	22.2	9	100
S ≠ demonstrative	123	77.8	35	22.2	158	100
total	130	77.8	37	22.2	167	100

Based on the relation between the demonstrative and the clause-initial position, we can try to formulate an algorithm using the demonstrative subject as a factor, Algorithm S-demonstrative.

Algorithm S-demonstrative: If the S is demonstrative, then SOV.

Before trying to apply the algorithm to a set of data that contains a demonstrative subject, I will examine if Algorithm (13) can predict the word order for the 12 tokens that contain a demonstrative subject. Table (103) shows the relative ERD, TP and SW for the 7 SOV tokens and the 2 OSV tokens.

Table (103) The relative measurement of tokens containing a demonstrative subject

(a) 5 SOV tokens: S<O (ERD), S>O (TP) and S>O (SW)
(b) 1 SOV tokens: S<O (ERD), S<O (TP) and S>O (SW)
(c) 1 SOV token: S<O (ERD), S=O (TP) and S<O (SW)
(d) 1 OSV tokens: S=O (ERD), S=O (TP) and S<O (SW)
(e) 1 OSV token: S>O (ERD), S=O (TP) and S>O (SW)

Table (103) shows that Algorithm (13) correctly predicts the word order for the 7 SOV tokens in (a), (b) and (c), and the 1 OSV tokens in (d); the ERD, TP and SW favor SOV in (a), the ERD and SW favor SOV for (b), the ERD favors SOV for (c), and the SW favors OSV for (d). One OSV token in (e) is an incorrect prediction by Algorithm (13). When Algorithm S-demonstrative is applied to the 9 tokens containing a demonstrative subject, it yields 7 correct predictions of SOV tokens that Algorithm (13) correctly predicts, and inverts 1 correct prediction of OSV token by Algorithm (13) into an incorrect prediction. This result indicates that the algorithm using the demonstrative subject does not improve Algorithm (13), and thus the demonstrative subject is dismissed as a factor for determining the constituent ordering of SOV and OSV.

3.8. Focus structure

The fronting of focused constituents in Japanese has been discussed among generative grammarians (e.g. Haig 1996). Based on Dryer's framework of activation, I will examine the distribution of focus-nonfocus in a clause. First, the RD of the subject (S), and that of the open sentence with S as a variable (i.e. the predicate proposition, OV) is measured. Table (104) shows the relative measurement of the RD between S and OV.

Table (104) Relative RD of S and OV

Relative RD	SOV		OSV		Total	
	#	%	#	%	#	%
RD of S > RD of OV	4	80.0	1	20.0	5	100
RD of S = RD of OV	26	55.3	21	44.7	47	100
RD of S < RD of OV	100	87.0	15	13.0	115	100
Total	130	77.8	37	22.2	167	100

Only the difference between RD of S = RD of OV and RD of S < RD of OV is statistically significant ($p=0.000$).

The category when the RD of S > the RD of OV is one where the S is focus and the predicate proposition OV is nonfocus. Similarly, the RD of the object (O) and the open sentence with O as a variable (i.e. the predicate proposition, SV) was measured. Table (105) shows the relative measurement of the RD between O and SV.

Table (105) Relative RD of O and SV

Relative RD	SOV		OSV		Total	
	#	%	#	%	#	%
RD of O > RD of SV	4	66.7	2	33.3	6	100
RD of O = RD of SV	77	81.1	18	18.9	95	100
RD of O < RD of SV	49	74.2	17	25.8	66	100
Total	130	77.8	37	22.2	167	100

The category when the RD of O > the RD of SV is one where the O is focus and the predicate proposition SV is nonfocus. Any of the differenced in Table (105) is not statistically significant.

The number of tokens with focus (S) – nonfocus (OV) as well as focus (O) – nonfocus (SV) was drawn from Tables (104) and (105) in order to determine if the focus status can predict the ordering choice. Table (106) shows the number of the focus on the S and the focus on the O in SOV and OSV clauses. The data is too small to obtain the statistical significance.

Table (106) Focus (S) vs. Focus (O)

Focus	SOV		OSV		Total	
	#	%	#	%	#	%
Focus on the S	4	80.0	1	20.0	5	100
Focus on the O	4	66.7	2	33.3	6	100
Total	8	72.7	3	27.3	11	100

Table (106) shows no relation between the focus element and its position in a clause. The occurrence of SOV is more frequent than OSV whether the S is focus or the O is focus.

Suppose we formulate an algorithm such as “If the focus falls on the O, then OSV”. Before trying to apply such an algorithm to the data, I will examine if Algorithms (13) correctly predicts the word order of tokens with focus structure. Table (107) shows the relative ERD, TP and SW for the 8 tokens with focus on the subject.

Table (107) The relative ERD, TP and SW of tokens with Focus on the subject

(a) 2 SOV: S>O (ERD), S>O (TP) and S=O (SW).
(b) 1 SOV: S>O (ERD), S>O (TP) and S>O (SW).
(c) 1 SOV: S>O (ERD), S=O (TP) and S>O (SW).
(d) 1 OSV: S>O (ERD), S<O (TP) and S<O (SW).

Table (107) shows that Algorithm (13) correctly predicts the word order for the all five tokens; the SW favors SOV for (a), the TP and SW favor SOV for (b), the SW favors SOV for (c), and the ERD, TP and SW favor OSV for (d). Table (108) shows the relative ERD, TP and SW for the six tokens with focus on the object.

Table (108) The relative ERD, TP and SW of tokens with Focus on the object

(a) 2 SOV: S<O (ERD), S=O (TP) and S>O (SW).
(b) 2 SOV: S<O (ERD), S>O (TP) and S<O (SW).
(c) 1 OSV: S<O (ERD), S>O (TP) and S=O (SW).
(d) 1 OSV: S=O (ERD), S>O (TP) and S<O (SW).

Table (108) shows that Algorithm (13) correctly predicts the word order for the 4 SOV tokens; the ERD and SW favor SOV for (a), and the ERD and TP favor SOV for (b). The algorithm, however, does not correctly predicts the word order for the 2 OSV tokens; the ERD and TP favor SOV for (d) and the TP favors SOV for (e).

The result in Table (91) indicates that if I formulate an algorithm such as “If the focus falls on the O, then OSV”, and apply it to a set of data containing the tokens with focus on the O, the algorithm yields two correct predictions of two OSV tokens, and it inverts four correct predictions of SOV by Algorithm (13) into incorrect ones. In other word, the algorithm using the focus structure as a factor does not improve Algorithm (13). As a result, the focus structure is dismissed as a factor for predicting the ordering choice between SOV and OSV.

3.9. Combining algorithms

We have learned that the algorithms using the *wa*-marking on the subject, *wa*-marking of the object, cohesion of objects, or demonstrative object as factor can improve on Algorithm (13) using the ERD, TP and SW as factors. In Section 3.5.3, we learned that Algorithm (16) that uses combined factors of S-*wa* and O-*wa* with the three variables, ERD, TP and SW, has the best prediction rate so far (i.e. 84.4%). In this section, I will try to formulate an algorithm that combines factors and see how it can improve on Algorithm (13).

Table (109) compares the correct prediction rate among the four algorithms, Algorithm O-*wa*, and Algorithm OS-*wa*, Algorithm Cohesion, and Algorithm O-demonstrative.

Algorithm S-*wa*: If the S is marked with *wa*, then SOV.
Algorithm O-*wa*: If the O is marked with *wa*, then OSV.
Algorithm Cohesion If the S occurs with cohesive O, then SOV
Algorithm O-demonstrative: If the O is demonstrative, then OSV.

Table (109) The success rate of Algorithms

Algorithms		Correct Predictions		Incorrect Predictions		Total	
		#	%	#	%	#	%
13A	Algorithm O-wa	9	60.0	6	40.0	15	100
14A	Algorithm S-wa	47	94.0	3	6.0	50	100
17A	Algorithm Cohesion	39	86.7	6	13.3	45	100
18A	Algorithm O-demonstrative	18	58.1	13	41.9	31	100

Table (109) shows that Algorithm S-*wa* and Algorithm Cohesion have a better success rate then Algorithm O-*wa* and Algorithm O-demonstrative. The reason seems to lie on the fact that the former two algorithms are the ones that predict SOV which the default favors, and the latter two are the ones that predict OSV against default favor SOV.

My data shows that 9 out of 13 SOV tokens that Algorithm O-demonstrative does not correctly predict contain a subject marked with *wa*, and 3 out of 6 SOV tokens that Algorithm O-*wa* does not correctly predict contain a *wa*-marked subject. Moreover, 2 out of 3 OSV tokens that Algorithm S-*wa* does not correctly predict contain a demonstrative object as well as 2 out of 6 OSV tokens that Algorithm Cohesion does not correctly predict. The content of predictions by the algorithms is summarized in Tables (110)-(113).

Table (110) Content of correct and incorrect predictions by Algorithm O-*wa*

Algorithm (15A) O-*wa*	Incorrect Predictions	Correct Predictions
	SOV	OSV
	6	9
Breakdown of predictions		1 also predicted by A (18A) O-demonstrative and A (13)
		1 also predicted by A (18A) O-demonstrative and A (17A) Cohesion
		2 also predicted by A (18A) O-demonstrative
		3 also predicted by A (13)
		3 only predicted by O-*wa*
Other features	3 contain S-wa	

Table (111) Content of correct and incorrect predictions by Algorithm S-*wa*

Algorithm (15A) S-*wa*	Correct Predictions	Incorrect Predictions
	SOV	OSV
	47	3
Breakdown of predictions	32 also predicted by A (13)	
	11 also predicted by A (17A) Cohesion and A (13)	
	1 also predicted by A (17A) Cohesion	
	3 only predicted by A (15A) S-*wa*	
Other features	9 contain O-demonstrative	2 contain O-demonstrative
	3 contain O-*wa*	1 contain Cohesive object

Table (112) Content of correct and incorrect predictions by Algorithm O-demonstrative

Algorithm (18A) O-demonstrative	Incorrect Predictions	Correct Predictions
	SOV	OSV
	13	18
Breakdown of predictions		2 also predicted by A (14A) O-*wa*
		5 also predicted by A(13)
		1 also predicted by A (14A) O-*wa* and A(13)
		10 only predicted by O-demo
Other features	6 contain S-wa	
	3 contain Cohesive object	
	1 contain S-wa and Cohesive object	2 contain S-*wa*

Table (113) Content of correct and incorrect predictions by Algorithm Cohesion

Algorithm (17A) Cohesion of object and verb	Correct Predictions	Incorrect Predictions
	SOV	OSV
	39	6
Breakdown of predictions	2 also predicted by A (13)	
	1 also predicted by A (13) and A (15A) S-*wa*	
	11 also predicted by A (15A) S-*wa*	
	25 oly predicted by A (15A) S-*wa*	
Other features	3 contain O-demonstrative	2 contain O-demonstrative

What we are concerned with is, when combining algorithms, how many correct predictions are inverted into incorrect ones, and how many incorrect ones are inverted to correct ones. Now, let us try three different combinations of algorithms based on the information shown in Tables (110)-(113).

S-*wa* and O-demonstrative

(A) Use SOV if the S is marked with *wa*.
Otherwise,
(B) Use OSV if the O is demonstrative.

O-demonstrative and O-*wa*

(A) Use OSV if the O is demonstrative.
Otherwise,
(B) Use OSV if the O is marked with *wa*.

Cohesion and O-demonstrative

(A) Use OSV if the O is demonstrative.
Otherwise,
(B) Use SOV if the S occurs with cohesive objects.

3.9.1. S-*wa* and O-demonstrative

We will try an algorithm using the subject marked with *wa* and the demonstrative object.

Algorithm (19):

(A) (i) Use SOV if the S is marked with *wa*.
Otherwise,
(ii) Use OSV if the O is demonstrative.

Applying Algorithm (19A-i) to a set of data containing a subject marked with *wa* yields 47 correct predictions of SOV and 3 incorrect predictions of OSV. Now we will apply Algorithm (19A-ii) to the 117 tokens (i.e. 167 minus 50 tokens) that do not contain a subject marked with *wa*. However, Algorithm (19A-ii) only applied to the tokens that contain a demonstrative object. Since 9 out of 13 SOV tokens and 2 out of 18 OSV tokens containing a demonstrative object also contain a *wa*-marked subject, the number of tokens that contain a demonstrative object among the 117 tokens amounts to 20 tokens (i.e. 4 SOV tokens and 16 OSV tokens). Applying Algorithm (19A-ii) to the 20 tokens yields 16 correct predictions of OSV and 4 incorrect predictions of SOV. Algorithm (19A), when it is applied to the 70 tokens (i.e. 50 plus 20 tokens), makes 63 correct predictions and 7 incorrect predictions. The success rate for Algorithm (19A) is 90%. Table (114) shows the result of predictions for Algorithm (19A).

Table (114) Predictions for Algorithm (19A)

Predictions	Correct	Incorrect	Total
A (19A-i) predicts SOV	47	3	50
A (19A-ii) predicts OSV	16	4	20
Total	63	7	70
Percent	90.0%	10.0%	100%

What if we reverse the application order of Algorithms S-*wa* and O-demonstrative. Would the result of predictions be different? Let us try Algorithm (20) which applies Algorithm O-demonstrative first and then Algorithm S-*wa* to a set of data.

Algorithm (20):

(A) (i) Use OSV if the O is demonstrative.
Otherwise,
(ii) Use SOV if the S is marked with *wa*.

Applying Algorithm O-demonstrative to a set of data containing a demonstrative object yields 18 correct predictions of OSV and 13 incorrect predictions of SOV. Now we will apply Algorithm S-*wa* to the 136 tokens (167 minus 31 tokens) that do not contain a demonstrative object. Among the 136 tokens, there are 40 tokens (39 SOV tokens and 1 OSV tokens) that contain a *wa*-marked subject. Applying Algorithm S-*wa* to the 136 tokens yields 39 correct predictions of SOV and 1 incorrect prediction of OSV. Algorithm (20A) is applied to a total of 71 tokens and produce 57 correct predictions (i.e. 13 OSV and 39 SOV) and 14 incorrect predictions (i.e. 13 SOV and 1 OSV). The success rate for Algorithm (20A) is 80.3%. The results indicate that Algorithm (19A) has a better success rate than Algorithm (20A).

Now we can formulate Algorithm (19) that integrates Algorithm (19A) with Algorithm (13).

Algorithm (19):

(A) (i) Use SOV if the S is marked with *wa*.
Otherwise,
(ii) Use OSV if the O is demonstrative.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (19) consists of three parts, Algorithm S-*wa*, i.e. (A-i), Algorithm O-demonstrative, i.e. (A-ii), and Algorithm (13), i.e. (B).

Let us apply Algorithm (13) to the 97 tokens that contain neither a subject marked with *wa* nor a demonstrative object (167 tokens minus 70 tokens). Tables (115)-(117) show the interaction of ERD, TP and SW for the 97 tokens.

Table (115) Interaction of ERD and TP when the SW of S > the SW of O for the 97 tokens that contain neither a S-wa nor an O-demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	3	0	3	0.0	2	1	3	33.3	1	0	1	0.0
S=O	1	0	1	0.0	4	0	4	0.0	2	0	2	0.0
S>O	5	0	5	0.0	5	0	5	0.0	14	1	15	6.7

Table (116) Interaction of ERD and TP when the SW of S = the SW of O for the 97 tokens that contain neither a S-wa nor an O-demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	1	0	1	0.0	2	0	2	0.0
S>O	2	0	2	0.0	1	0	1	0.0	5	1	6	16.7

Table (117) Interaction of ERD and TP when the SW of S < the SW of O for the 97 tokens that contain neither a S-wa nor an O-demonstrative

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	4	0	4	0.0
S=O	2	3	5	60.0	3	5	8	62.5	4	1	5	20.0
S>O	0	2	2	100.0	6	1	7	14.3	9	1	10	10.0

Applying Algorithm (13) to the 97 tokens presented in Tables (115)-(117) yields the result of predictions shown in Table (118).

Table (118) Predictions for Algorithm (13) when applied to the 97 tokens that contain neither a S-wa nor an O-demonstrative

Algorithm (19B)	Correct	Incorrect	Total
Predicts SOV	71	8	79
Predicts OSV	10	8	18
Total	81	16	97
Percent	83.5%	16.5%	100%

When we combine the results of prediction by Algorithm (19A) shown in Table (114) and by Algorithm (19B) shown in Table (118), we can obtain the prediction rate for Algorithm (19). The result is shown in Table (119).

Table (11) Predictions for Algorithm (19)

Algorithm (19)	Correct	Incorrect	Total
Predicts SOV	118	11	129
Predicts OSV	26	12	38
Total	144	23	167
Percent	86.2%	13.8%	100%

Table (118) shows that algorithm (19) correctly predicts 144 tokens out of 167 tokens in the data with 86.2% of success rate. This success rate is higher than 80.2% by Algorithms (13) and (18) that is Algorithm O-demonstrative with Algorithm (13), and 82.6% by Algorithm (15) that is Algorithm S-*wa* with Algorithm (13).

3.9.2. O-demonstrative and O-*wa*

In this section, I will try an algorithm using the demonstrative object and the object marked with *wa*. The algorithm we will try is Algorithm (21).

Algorithm (21)
(A) (i) Use OSV if the O is demonstrative.
Otherwise,
(ii) Use OSV if the O is marked with *wa*.

Applying Algorithm (21A-i) to the 31 tokens containing a demonstrative object yields 18 correct predictions of OSV and 13 incorrect tokens of SOV. It is reminded that algorithm (21A-ii) only applies to the tokens that contain a *wa*-marked object. Since 3 out of 9 OSV tokens that contain a *wa*-marked object also contain a demonstrative object, among the 136 tokens that do not contain a demonstrative object, there are 12 tokens (i.e. 6 SOV and 6 OSV tokens) containing a *wa*-marked object. Now, we will apply Algorithm (21A-ii) to the 12 tokens. Applying Algorithm O-*wa* to the 12 tokens yields 6 correct predictions of OSV and 6 incorrect predictions of SOV. Algorithms (21A), when they are applied to the 43 tokens (31 tokens plus 12 tokens), produce 24 correct predictions and 18 incorrect predictions. Table (120) shows the result of prediction for Algorithm (21A).

Table (120) Predictions for Algorithm (21A)

Predictions	Correct	Incorrect	Total
A (21A-i) predicts OSV	18	13	31
A (21A-ii) predicts OSV	6	6	12
Total	24	19	43
Percent	55.8%	44.2%	100%

Let us try Algorithm (22A) which reverses the order of applying Algorithms O-demonstrative and O-*wa* from Algorithm (21A)

Algorithm (22)
(A) (i) Use OSV if the O is marked with *wa*.
Otherwise,
(ii) Use OSV if the O is demonstrative.

Applying Algorithm (22A-i) to a set of data containing a *wa*-marked object yields 9 correct predictions of OSV and 6 incorrect predictions of SOV. Now we have 152 tokens that do not contain a *wa*-marked object. Algorithm (22A-ii) only applies to the tokens that contain a demonstrative object. Among the 31 tokens (13 SOV and 18 OSV tokens) containing a demonstrative object, there are 3 OSV tokens that also contain a *wa*-marked object. That is, among 152 tokens that do not contain a *wa*-marked object, there are 28 tokens (13 SOV and 15 OSV tokens) containing a demonstrative object. Applying Algorithm (21A-ii) to the 28 tokens yields 15 correct predictions of OSV and 13 incorrect predictions of SOV. Table (120) shows the result of predictions for Algorithm (22A).

Table (121) Predictions for Algorithm (22A)

Predictions	Correct	Incorrect	Total
A (22A-i) predicts OSV	9	6	15
A (22A-ii) predicts OSV	15	13	28
Total	24	19	43
Percent	55.8%	44.2%	100%

Table (121) shows that Algorithm (21A) and (22A) have the same success rate for the present data.

We will now posit Algorithm (21).

Algorithm (21):

(A) (i) Use OSV if the O is demonstrative.
Otherwise,
(ii) Use OSV if the O is marked with *wa*.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (21) consists of three parts, Algorithm O-demonstrative, i.e. (21A-i), Algorithm O-*wa*, i.e. (21A-ii), and Algorithm (13), i.e. (21B).

The present data contains 124 tokens (i.e. 167 minus 43 tokens) that contain neither a demonstrative object nor a *wa*-marked object. Tables (121)-(123) show the interaction of ERD, TP and SW for the 124 tokens.

Table (121) Interaction of ERD and TP when the SW of S > the SW of O for the 124 tokens that contain neither a O-wa nor an O-demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	3	0	3	0.0	2	1	3	33.3	6	0	6	0.0
S=O	1	0	1	0.0	4	0	4	0.0	6	0	6	0.0
S>O	3	0	3	0.0	7	0	7	0.0	18	0	18	0.0

Table (122) Interaction of ERD and TP when the SW of S = the SW of O for the 124 tokens that contain neither a O-wa nor an O-demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %	SOV #	OSV #	Total #	OSV %
S<O	1	1	2	50.0	0	0	0	n/a	2	1	3	33.3
S=O	1	0	1	0.0	2	0	2	0.0	5	0	5	0.0
S>O	2	0	2	0.0	1	0	1	0.0	8	0	8	0.0

Table (123) Interaction of ERD and TP when the SW of S < the SW of O for the 124 tokens that contain neither a O-wa nor an O-demonstrative

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	4	0	4	0.0
S=O	4	2	7	57.1	2	3	5	60.0	6	0	6	20.0
S>O	1	2	3	66.7	7	2	9	22.2	14	0	14	0.0

When we apply Algorithm (13) to the 124 tokens shown in Tables (121)-(123), the result of predictions shown in Table (124) is yielded.

Table (124) Predictions for Algorithm (13) when applied to the 124 tokens that contain neither a O-wa nor an O-demonstrative

Prediction	Correct	Incorrect	Total
Predicts SOV	100	6	106
Predicts OSV	7	11	18
Total	107	17	124
Percent	86.3%	13.7%	100%

Combining the results of prediction by Algorithm (21A) shown in Table () and by Algorithm (21B) in Table (124) gives us the result of predictions for Algorithm (21). The result is shown in Table (125).

Table (125) Predictions for Algorithm (21)

Algorithm (20)	Correct	Incorrect	Total
Predicts SOV	118	19	137
Predicts OSV	13	17	30
Total	131	36	167
Percent	78.4%	21.6%	100%

Table (125) shows a 78.4% success rate for Algorithm (21). This success rate is the lowest among algorithms we have tried so far.

3.9.3. Cohesive objects and O-demonstrative

In this section, I will try an algorithm using the cohesion of object and verb and the demonstrative object. Table (113) indicates that 2 tokens of SOV as well as 2 OSV tokens contain an object that is demonstrative and cohesive.

Algorithm (23)

(1) (i) Use SOV if the S occurs with a cohesive object.
Otherwise,
(ii) Use OSV if the O is demonstrative.

When we apply Algorithm (23A-i) to a set of date that contain a subject occurring with a cohesive object, it yields 39 correct predictions of SOV and 6 incorrect predictions of OSV. Among 31 tokens containing a demonstrative object, 3 out of 10 SOV tokens and 2 out of 18 OSV tokens contain a cohesive object. Thus, when we apply Algorithm (23A-ii) to a set of data containing a demonstrative object, it yields 6 correct predictions of OSV and 10 incorrect predictions SOV. Table (126) shows the result of predictions when Algorithm (13A) is applied to the 71 tokens containing an object that are either cohesive or demonstrative (49 SOV and 22 OSV).

Table (126) Predictions for Algorithm (23A)

Predictions	Correct	Incorrect	Total
A (23A-i) predicts SOV	39	6	45
A (23A-ii) predicts OSV	16	10	26
Total	55	16	71
Percent	77.5%	22.5%	100%

We will also try an algorithm that applies Algorithm O-demonstrative before applying Algorithm Cohesion.

Algorithm (24)

(A) (i) Use OSV if the O is demonstrative
Otherwise,
(ii) Use SOV if the S occurs with a cohesive object.

Table (127) shows that result of predictions for Algorithm (24A).

Table (127) Predictions for Algorithm (24A)

Predictions	Correct	Incorrect	Total
A (24A-i) predicts OSV	18	13	31
A (24A-ii) predicts SOV	36	4	40
Total	54	17	71
Percent	76.1%	23.9%	100%

Table (127) shows that Algorithm (23A) has a better prediction rate by one token than Algorithm (24A).

Let us formulate an algorithm combining Algorithm (23A) and Algorithm (13).

Algorithm (23)

Algorithm (23):

(A) (i) Use SOV if the S occurs with a cohesive object.
Otherwise,
(ii) Use OSV if the O is demonstrative.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Algorithm (23) consists of three parts, Algorithm Cohesion, i.e. (A-i), Algorithm O-demonstrative, i.e. (A-ii), and Algorithm (13) i.e. (B).

We now apply Algorithm (23B) to the 96 tokens containing an object that are neither cohesive nor demonstrative (167 tokens minus 71 tokens). Tables (128)-(130) show the interaction of ERD, TP and SW for the 96 tokens.

Table (128) Interaction of ERD and TP when the SW of S > the SW of O for the 96 tokens containing an object that is neither cohesive nor demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	2	0	2	0.0	2	1	3	33.3	4	0	4	0.0
S=O	1	0	1	0.0	3	0	3	0.0	4	0	4	0.0
S>O	2	0	2	0.0	3	0	3	0.0	11	1	12	8.3

Table (129) Interaction of ERD and TP when the SW of S = the SW of O for the 96 tokens containing an object that is neither cohesive nor demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	0	1	0.0	0	0	0	n/a	2	0	2	0.0
S=O	1	0	1	0.0	0	0	0	n/a	3	0	3	0.0
S>O	1	0	1	0.0	1	0	1	0.0	7	1	8	12.5

Table (130) Interaction of ERD and TP when the SW of S < the SW of O for the 96 tokens containing an object that is neither cohesive nor demonstrative

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	2	3	66.7	0	0	0	n/a	3	0	3	0.0
S=O	2	3	5	60.0	4	4	8	50.0	4	1	5	20.0
S>O	1	2	3	66.7	5	1	6	16.7	13	0	13	0.0

Table (131) shows the result of predictions for Algorithm (23B) when applied to the 96 tokens.

Table (131) Predictions for Algorithm (23B)

Prediction	Correct	Incorrect	Total
Predicts SOV	72	7	79
Predicts OSV	9	8	17
Total	81	15	96
Percent	84.4%	15.6%	100%

When we combine the results shown in Table (126) and (131), we can obtain the results of predictions for Algorithm (23). Table (132) shows the results.

Table (132) Predictions for Algorithm (23)

Algorithm (23)	Correct	Incorrect	Total
Predicts SOV	111	13	124
Predicts OSV	25	18	43
Total	136	31	167
Percent	81.4%	18.6%	100%

The success rate of 81.4% is the fourth highest after Algorithm (19) involving S-*wa* and O-demonstrative, Algorithm (16) involving S-*wa* and O-*wa*, and Algorithm (15) involving S-*wa*.

3.9.4. Combining algorithms[11]

In this section, I will try to combine the four algorithms using factors S-*wa*, O-*wa*, O-demonstrative and Cohesive objects. In Sections 3.9.1 to 3.9.3, we learned that combining an algorithm predicting SOV with an algorithm predicting OSV has a better result than combining two algorithms predicting the same order. We also learned that applying an algorithm that affects a bigger portion of the data before an algorithm that applies to a smaller portion of the data produces a better result than applying an algorithm that applies to a smaller portion of the data first.

Based on the above information, I will combine the four algorithms in the order of (i) S-*wa*, (ii) Cohesive object, (iii) O-demonstrative and (iv) O-*wa*. The algorithm we will try is Algorithm (25A).

[11] Researchers using the framework of Optimality Theory to account for scrambling (e.g. Choi 1999, Bluntner et al. ms.) take a similar approach as my algorithms. They posist competing constraints that are violable, and constraints that are ranked higher overrule low-ranked constraints. The outcomes of the structured constraints are well-formed linguistic forms. Their approach may be as expressive as mine. The crucial component in a framework is factors or constraints in their term. In this study, factors are identified from the patterns that language samples demonstrate as opposed to feature-based constraints.

Algorithm (25A):

(i) Use SOV if the S is marked with *wa*.
Otherwise,
(ii) Use SOV if the S occurs with a cohesive object.
Otherwise,
(iii) Use OSV if the O is demonstrative.
Otherwise,
(iv) Use OSV if the O is marked with *wa*.

Table (132) shows the application of subalgorithms (i) to (iv) and the result of predictions.

Table (133) Predictions for Algorithm (25A)

Order of application	Algorithms	Correct	Incorrect	Total
1	S-*wa* predicts SOV	47	3	50
2	Cohesion predicts SOV	27	5	32
3	O-demonstrative predicts OSV	14	2	16
4	O-*wa* predicts OSV	6	3	9
Total		94	13	107
Percent		87.9%	12.1%	100%

Table (133) shows an 87.9% of success rate for Algorithm (25A).

Table (134) Predictions for Algorithm (26A)

Order of application	Algorithms	Correct	Incorrect	Total
1	S-*wa* predicts SOV	47	3	50
2	O-*wa* predicts OSV	9	3	12
3	Cohesion predicts SOV	27	4	31
4	O-demonstrative predicts OSV	12	2	14
Total		95	12	107
Percent		88.8%	11.2%	100%

In what follows, I will try different algorithms by altering the order of applying subalgorithms (i)-(iv) in Algorithm (25A). Tables (134)-(137) show the results of predictions for the variations of Algorithm (25A) in terms of application order.

Table (135) Predictions for Algorithm (27A)

Order of application	Algorithms	Correct	Incorrect	Total
1	S-*wa* predicts SOV	47	3	50
2	O-*wa* predicts OSV	9	3	12
3	O-demonstrative predicts OSV	13	4	17
4	Cohesion predicts SOV	25	3	28
Total		94	13	107
Percent		87.9%	12.1%	100%

Table (136) Predictions for Algorithm (28A)

Order of application	Algorithms	Correct	Incorrect	Total
1	Cohesion predicts SOV	39	6	45
2	S-*wa* predicts SOV	35	2	37
3	O-demonstrative predicts OSV	14	2	16
4	O-*wa* predicts OSV	6	3	9
Total		94	13	107
Percent		87.9%	12.1%	100%

Table (137) Predictions for Algorithm (29A)

Order of application	Algorithms	Correct	Incorrect	Total
1	S-*wa* predicts SOV	47	3	50
2	O-demonstrative predicts OSV	16	4	20
3	O-*wa* predicts OSV	6	3	9
4	Cohesion predicts SOV	25	3	28
Total		94	13	107
Percent		87.9%	12.1%	100%

Tables (134)-(137) show that (26A) has the correct highest prediction rate among the four subalgorithms using the same factors but with different application order. The difference in correct predictions between (26A) and the other three subalgorithms was by one token. This result suggests that insofar as the four factors are concerned, the order of subalgorithms does not matter much.

Now, we will formulate an algorithm by combining Algorithm (26A) and Algorithm (13). The algorithm we will try is Algorithm (26).

Algorithm (26):

(A) (i) Use SOV if the S is marked with *wa*.
Otherwise,
(ii) Use OSV if the O is marked with *wa*.
Otherwise,
(iii) Use SOV if the S occurs with a cohesive object.
Otherwise,
(iv) Use OSV if the O is demonstrative.
Otherwise,
(B) Use SOV if
(i) ERD of S < ERD of O or
(ii) TP of S > TP of O or
(iii) SW of S > SW of O,
or (iv)
ERD of S = ERD of O and
TP of S = TP of O and
SW of S = SW of O.
Otherwise, use OSV.

Tables (138)-(139) show the interaction of relative ERD, TP and SW for the 60 tokens that contain neither the S-*wa*, O-*wa*, O-demonstrative nor cohesive objects (i.e. 167 minus 107 tokens).

Table (138) Interaction of ERD and TP when the SW of S > the SW of O for the 60 tokens that contain neither a S-*wa* , an O-*wa* , an O-demonstrative nor a cohesive object.

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	2	0	2	0.0	2	1	3	33.3	0	0	0	n/a
S=O	1	0	1	0.0	3	0	3	0.0	0	0	0	n/a
S>O	2	0	2	0.0	3	0	3	0.0	9	0	9	0.0

Table (139) Interaction of ERD and TP when the SW of S > the SW of O for the 60 tokens that contain neither a S-*wa* , an O-*wa* , an O-demonstrative nor a cohesive object.

	S>O (ERD)				S=O (ERD)				S<O (ERD)			
TP	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	0	0	0	n/a	0	0	0	n/a	2	0	2	0.0
S=O	1	0	1	0.0	0	0	0	n/a	0	0	0	n/a
S>O	1	0	1	0.0	1	0	1	0.0	4	0	4	0.0

Table (140) Interaction of ERD and TP when the SW of S > the SW of O for the 60 tokens that contain neither a S-*wa* , an O-*wa* , an O-demonstrative nor a cohesive object.

TP	S>O (ERD)				S=O (ERD)				S<O (ERD)			
	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV	SOV	OSV	Total	OSV
	#	#	#	%	#	#	#	%	#	#	#	%
S<O	1	1	2	50.0	0	0	0	n/a	3	0	3	0.0
S=O	1	2	3	66.7	1	2	3	66.7	3	0	3	0.0
S>O	0	2	2	100.0	4	1	5	20.0	7	0	7	0.0

When we apply Algorithm (26B) to the 60 tokens, it yields 52 correct predictions and 8 incorrect predictions, as shown in Table (141).

Table (141) Predictions for Algorithm (13) when applied to the 60 tokens that contain neither a S-wa, an O-wa, an O-demonstrative nor a cohesive object

Prediction	Correct	Incorrect	Total
Predicts SOV	47	4	51
Predicts OSV	5	4	9
Total	52	8	60
Percent	86.7%	13.3%	100%

Combining the results shown in Tables (134) and (141) produces the result of predictions for Algorithm (26).

Table (142) Predictions for Algorithm (26)

Algorithm (26)	Correct	Incorrect	Total
Predicts SOV	121	11	132
Predicts OSV	26	9	35
Total	147	20	167
Percent	88.0%	12.0%	100%

Table (142) shows an 88% of success rate for Algorithm (26). This percentage is the highest in all algorithms we have tried in this chapter.

3.10. Incorrect predictions

Table (143) shows the features of subjects and objects in tokens whose word orders were incorrectly predicted by Algorithm (26). The cases where the factor ERD, TP or SW favors the word order of the token are indicated in bold and italics.

Table (143) Twenty incorrect predictions produced by Algorithm(26)

	Algorithm	Predicts	Order of Token	ERD		TP		SW		Other features	
				O	S	O	S	O	S	O	S
1	S-wa	SOV	OSV	***1***	***F***	1	5	2	2	demonstrative	1sg
2	S-wa	SOV	OSV	20+	3	3	6	***8***	***2***	demonstrative	1sg
3	S-wa	SOV	OSV	2	2	0	5	***24***	***2***		1sg
4	Cohesion	SOV	OSV	***1***	***6***	1	2	3	4		
5	Cohesion	SOV	OSV	***4***	***20+***	***2***	***1***	5	5		
6	Cohesion	SOV	OSV	1	1	1	1	***2***	***1***		quantifier
7	Cohesion	SOV	OSV	F	1	***1***	***0***	2	2	marked with *mo*	quantifier
8	ERD, TP and SW	SOV	OSV	F	F	***3***	***2***	2	3		
9	ERD, TP and SW	SOV	OSV	***1***	***8***	1	5	***6***	***1***		
10	ERD, TP and SW	SOV	OSV	***1***	***F***	0	6	***6***	***4***		
11	ERD, TP and SW	SOV	OSV	***1***	***8***	0	1	***21***	***3***		
	Algorithm	Predicts	Order of Token	ERD		TP		SW		Other features	
				S	O	S	O	S	O	S	O
12	O-wa	OSV	SOV	F	F	0	0	3	8		
13	O-wa	OSV	SOV	F	F	0	0	2	3	quantifier	
14	O-wa	OSV	SOV	20+	F	***1***	***0***	1	12	1sg	
15	O-demonstrative	OSV	SOV	20+	F	***7***	***1***	***2***	***1***		
16	O-demonstrative	OSV	SOV	6	1	***6***	***0***	2	2		
17	ERD, TP and SW	OSV	SOV	20+	5	1	1	2	2		
18	ERD, TP and SW	OSV	SOV	20+	2	1	2	2	10	marked with *mo* , 1sg	
19	ERD, TP and SW	OSV	SOV	3	1	1	1	1	8	1sg	
20	ERD, TP and SW	OSV	SOV	F	F	1	1	7	9		

1sg=1st singular pronoun

Table (143) shows that while one or two of the three factors of ERD, TP and SW favor all eleven OSV clauses, there are only three SOV clauses whose order is favored by one or two of the three factors. That is, non-default orders are motivated by more factors than the default order. This result suggests that the effect of the default order is taking place. Tokens #1 and #2 present a case where the factor demonstrative (favoring OSV) and the factor particle *wa* (favoring SOV) are in conflict. In these tokens, the factor O-demonstrative won over the factor particle *wa*. In addition, the object in Token #1 has a lower RD than the subject, and the object in Token #2 is syntactically longer than the subject is. One of the reasons that subjects in Tokens #1, #2 and #3 have a higher TP is that the referents of these subjects are the speaker, which is prone to be thematic in discourse. Tokens #3 and #8 present a case where the factors TP and SW are in conflict. In Token #3, SW won over TP, and in #8, TP won over SW. The higher TP for the subject in Token #3 is attributed to the referent of the subject being the speaker. Token #8 suggests that the difference in syntactic unit counts by one is not significant compared to the difference in cataphoric persistence. In Tokens #4 and #5, the initial constituent, the object, is more activated than the subject is.

Token #6 involves a lexically primed entity and an anaphoric noun phrase.

Token # 6

tigau	koto	min'na	siteru	wake-de
different	matter	everybody	do.PPG	reason-COP

'So, each of them is doing different things.'

In the immediately preceding clause, there is an expression, *hanarebanare na no* (scattered PrNom FP) '(they are) scattered'. The referent of the object is lexically primed by 'scattered' while the antecedent of the subject is an elliptical argument. In this token, a lexically primed entity comes before an entity that is not lexically primed. The tendency for lexically primed entities to be introduced earlier in a sentence is reported in Arnold et al. (2000).

Tokens #7 and #18 suggest a possible motivation factored by the particle *mo*.

Token #7

okkake	mo	san'nin	de	yattari	nanka
being.groupie	also	3.people	by	do.Inexhaustive	such.as

siteru-n-desu	kedo
do.PPG-NMLZ-COP	though

'The three people are also doing things like being groupie.'

Token #18

atasi	mo	X[12]	tte	iu	no	mimasita	mono
1sg	also	X	QT	say	PRO	see.PST	indeed

'I also saw so-called 'X', indeed.'

The use of particle *mo* presupposes the presence of an entity or a group of entities that are comparable to the *mo*-marked entity. The presupposition implied by the use of *mo* is some other things that these three people often do together (in token #7) or some other people who saw something called 'X' (in Token #18). The speaker believes that the entity introduced by the use of *mo* becomes activated at the moment of utterance in the hearer's mind. We discussed this notion of metapresupposition in Chapter 1. Although we did not find a positional preference of the particle *mo* in the present data, the metapresupposition can be argued as a factor here.

Another Token that contains the subject with the first singular pronoun is Token #19.

Token #19

watasi	an'mari	K-san	ga	hohoho	tte	nokezotte	warau	no
1sg	not.much	K-Ms.	NOM	hohoho	QT	bend.back	laugh	NMLZ

mita-koto-nai-n-desu	kedo
saw.PST-matter-NOM-NEG-NMLZ-COP	though

'I haven't seen much Ms. K bend back and chuckle like that.'

The subject 'I' is the speaker, and Ms. K in the object noun phrase is the hearer. That is, the referent of the subject is more activated than that of the object. The relative activation and the 'defaultness' seem to be relevant for this token.

Tokens #12 and #20 suggest that the focus status might be a relevant factor.

---

[12] 'X' contains seven syntactic units.

Token #12

iron'na	kata	ga	motiron	hahaoya	ga	sinde	kanasii	tte
various	people.HON	NOM	of.course	mother	NOM	die.TE	sad	QT

koto	wa	kaiterassyaru	kedo
matter	TOP	write.HON.PPG	though

'Of course, everybody writes on how sad they were after their mother passed away.'

Since the conversation started on how wonderfully the hearer described in some writing his feeling when his mother passed away in the 26th clauses from the current clause in the preceding discourse, the referent of the object has been a topic of conversation in the recent discourse. That is, the open proposition (X write about his mother) is active at the moment of utterance, while the variable X is nonactive. In Token #12, the variable X is the object 'various people'. This active and nonactive distribution indicates the nonactive being the narrow focus. Although the narrow focus in terms of active and nonactive distribution was not proved to be a factor for constituent ordering in the present data (cf. Chapter three), the focus status of the initial constituent might be relevant in this token.

Token #20

zimusyo	no	syatyoo-san	ka	nanka	ga
office	GEN	president-Mr	whether	something	NOM

moo	intai-simasu	tte	iu	yoo-na	koto	o	ossyatta	no	kasira
already	retirement-do	QT	say	like-PrNOM	thing	ACC	say.HON.PST	NMLZ	I.wonder

'I wonder if the president of the office or someone said things like you would retire soon.'

There is an expression 'I quit once and for all.' in the preceding discourse. Although there is no referential expression for *intai-simasu* 'to retire', which is the content of the object noun phrase, in the recent discourse, we can argue that in the speaker's mind, the open proposition with the variable X (X said that the hearer would retire) was active, or at least more activated than (X) at the time of utterance. This active and nonactive distribution can indicate the focus status of the initial constituent.

Similarly, Token #17 suggests a case of double focus structure.

Token #17

demo	otoo-sama	sukosi	o-kane	okutte-kudasatta-n-desu	tte	ne
but	father-HON	a.little	POL-money	send.TE-give.PST-NMLZ-COP	QT	FP

'But your father sent you some money, didn't he?'

The topic has been around the hearer's struggle for money when she just started going to an acting school in the city. The conjunction *demo* at the clause-initial position indicates the speaker's intention to introduce a new topic, which is the hearer's father, *otoosama*. In the preceding discourse, the hearer talks about her relatives sending her care packages. For Token #17, we can consider a case of double focus structure, X1 do Y1 and X2 do Y2, such as in *You wash dishes, and I clean the table*.

Token #13 is a case where the initial constituent is 'hail'.

Token #13

mina-sama	o-namae	wa	goran'ninatta-koto-ga-aru-kamosirenai	kedo
everybody.HON	HON-name	TOP	see.HON.PST-matter-NOM-exist-maybe	though

'Ladies and Gentleman, you may have seen his name,'

The subject *minasama* 'everybody' refers to the audience who is watching this talk show. The object *onamae* 'name' refers to the name of the hearer (=the guest of the day) who was introduced in the previous two clauses. The use of *wa*-marking on the object implies the presence of a contrastive entity in the speaker's mind at the time of utterance, which is the hearer's 'face' as opposed to her 'name' as in *You might have heard about him, but* *you might not have seen him*. This token is uttered at the very beginning of the conversation (i.e. the third clause after the talk show started), and *minasama* 'everybody' is functioning as an attention getter, as translated in English 'Ladies and Gentlemen'.

Token #14 presents a case where an activated entity and a contrastive entity compete for the initial position.

Token #14

atasi	panda	ga	bisuketto	moratta	toki	ni
1sg	panda	NOM	biscuit	receive.PST	time	TEMP

ahiru	ni	yatta	tte	hanasi	wa	kiita-koto-ga-arimasu	kedo	ne
duck	to	give.PST	QT	story	TOP	hear.PST-matter-NOM-exist	though	FP

'I heard a story that when a panda got a piece of biscuit, the panda gave it to a duck.'

The referent of the subject is the speaker herself. It is obvious that the referent is already activated in the speaker and hearer's minds despite the fact that there is no antecedent for the subject. The use of the particle *wa* implies the contrast between the story that the hearer introduced in the preceding discourse and the one that the speaker is giving in this token. The hearer's story is about weird animal behavior such as a dachshund falling in love with a hamster, a duck falling in love with a dog, or a young yellow tail falling in love with a killer whale or a dolphin. In the fifth clause in the preceding discourse from the current clause, the speaker responded saying *sore mo taihen'na hanasi desu ne* (it also horrible story COP FP) 'It's also a horrible story, isn't it.' Unlike noun phrases marked by *wa or mo*, zero-marked noun phrases do not imply the presence of an entity or a group of entities in the speaker's mind with respect to the entity that the speaker currently talks about at the time of utterance. The 1st singular person pronoun with zero-particle gives a simple neutral introduction of the actor of the verb. Between an activated entity (the subject) and a nonactive contrastive entity (the object), an activated entity comes before a nonactive contrastive entity in this particular token. Alternatively, we may simply argue that the 'defaultness' is relevant here.

3.11. Summary

Tables (144) and (145) show the percentage of occurrence for SOV and OSV orders when the factor favors. Algorithms based upon the factor particle *wa*, demonstrative and cohesive objects are applicable to a limited set of data. Therefore, the overall percentage was not calculated for these algorithms.

Table (144) Comparison of success rates: single factor

Algorithm based on a single factor	Correct Prediction					
	SOV		OSV		Overall	
	#	%	#	%	#	%
RD	105	80.8	15	40.5	120	71.9
ERD	118	90.8	16	43.2	134	80.2
TP	109	83.8	12	32.4	121	72.5
SW	85	65.4	18	48.6	103	61.7
wa-marked O	0	0.0	9	60.0	n/a	
wa-marked S	47	94.0	0	0.0	n/a	
wa-marked O and S	47	83.9	9	16.1	n/a	
Demonstrative O	0	0.0	18	58.1	n/a	
Cohesive Object	39	86.7	0	0.0	n/a	
Overall occurrence	130	77.8	37	22.2	167	100

Table (145) Comparison of success rates: combined factors

Algorithm based on combined factors		Correct Prediction					
		SOV		OSV		Overall	
		#	%	#	%	#	%
A5	RD and TP	118	90.8	14	37.8	132	79.0
A7	RD and SW	113	86.9	13	35.1	126	75.4
A9	TP and SW	106	81.5	18	48.6	124	74.3
A11	RD, TP and SW	121	93.1	11	29.7	132	79.0
A12	ERD, TP and SW	118	90.8	16	43.2	134	80.2
A14	O-wa, ERD, TP and SW	114	87.7	21	56.8	135	80.8
A15	S-wa, ERD, TP and SW	122	93.8	16	43.2	138	82.6
A16	S-wa, O-wa, ERD, TP, and SW	121	93.1	20	54.1	141	84.4
A18	Demonstrative-O, ERD, TP and SW	106	81.5	28	75.7	134	80.2
A17	Cohesive-O, ERD, TP and SW	121	93.1	14	37.8	135	80.8
A19	S-wa, Demonstrative-O, ERD, TP and SW	118	90.8	26	70.3	144	86.2
A21	O-wa, Demonstrative-O, ERD, TP and SW	118	90.8	13	35.1	131	78.4
A23	Demonstrative-O, Cohesive-O, ERD, TP and SW	111	85.4	25	67.6	136	81.4
A26	S-wa, O-wa, Cohesive-O, Demonstrative-O, ERD, TP and SW	121	93.1	26	70.3	147	88.0
Overall Occurrence in the data		130	77.8	37	22.2	167	100

Algorithms based on the single factor of referential distance (RD), Extended referential distance (ERD), or topic persistence (TP) predicted the occurrence of SOV and OSV, while the algorithm based on syntactic weight (SW) only predicted the occurrence of OSV better than the overall percentage of occurrences. Algorithms based on two factors, RD and TP, RD and SW, and TP and SW made better predictions than the algorithms with the single factor of RD, TP or SW. The algorithm based on ERD made better predictions than the algorithm based on RD, proving that ERD is more relevant to determining the word order, particularly to predicting the occurrence of OSV order.

Algorithm (13), based on ERD. TP and SW:

Use SOV if

(i)	ERD	of	S	<	ERD	of	O	or
(ii)	TP	of	S	>	TP	of	O	or
(iii)	SW	of	S	>	SW	of	O,	
Or (iv)								
	ERD	of	S	=	ERD	of	O	and
	TP	of	S	=	TP	of	O	and
	SW	of	S	=	SW	of	O.	

Otherwise, use OSV.

yielded an 80.2% (134/167 tokens) success rate, which is the highest overall success rate among the algorithms based on the factor RD, ERD, TP and SW.

The algorithm using factors such as the demonstrative object, the particle *wa*, or the cohesion of object and verb, when they are integrated with Algorithm (13), improved the success rate on Algorithm (13): 80.8% for Algorithm (14) using the object marked with *wa*, and for Algorithm (17) using cohesive objects with Algorithm (13), 82.6% for Algorithm (15) using the *wa*-marked subject, 84.4% for Algorithm (16) using the *wa*-marking on the subject and the object with Algorithm (13), 86.2% for Algorithm (19) using the *wa*-marking on the subject and the demonstrative object with Algorithm (13), and 81.4% for Algorithm (23) using the cohesive object and the demonstrative object with Algorithm (13).

When we combined all factors to formulate Algorithm (26):

Algorithm (26)

(A) (i) Use SOV if the S is marked with *wa*.
Otherwise,
(ii) Use OSV if the O is marked with *wa*.
Otherwise,
(iii) Use SOV if the S occurs with a cohesive object.
Otherwise,
(iv) Use OSV if the O is demonstrative.

Otherwise,
(B) Use SOV if

(i)	ERD	of	S	<	ERD	of	O	or
(ii)	TP	of	S	>	TP	of	O	or
(iii)	SW	of	S	>	SW	of	O,	

Or (iv)

ERD	of	S	=	ERD	of	O	and
TP	of	S	=	TP	of	O	and
SW	of	S	=	SW	of	O.	

Otherwise, use OSV.

the highest success rate 88% was acquired.

The highest success rate for predicting the occurrence of OSV order was 75.7% for Algorithm (18) using demonstrative objects and Algorithm (13), and the highest percentage for predicting the occurrence of SOV order was 93.8%, tie for Algorithm (15) using S-*wa*, ERD, TP and SW. The factors demonstrative object and SW are stronger motivation for OSV order, while they are not for SOV order.

The fact that algorithms based on combined factors have a higher success rate than those based upon the single factor indicates that factors are complementarily distributed in that when one factor is not relevant, another factor is. One of the interests of this thesis was the relative strength among the three factors, RD, TP and SW. RD was overall the strongest, followed by TP, and the by SW. The fact that the number of cases where the three factors are in conflict was very small suggests that they are complementarily working together rather than acting as single factors.

Chapter 4

Constituent ordering in the ≥3NPs construction

In Chapter 3, we compared SOV and OSV orders. Algorithm (13) that uses the ERD (extended RD), TP, and SW as factors predicted the ordering choice between SOV and OSV correctly in 80.2% of the data. As well as Algorithm (13), a number of secondary algorithms using the particle *wa*, demonstrative and cohesive object as factors succeeded in predicting the ordering in another 7.8% of the data. Algorithm (26) that combines all the factors attained a success rate of 88%.

In comparing SOV and OSV, two issues, (1) which precedes the other and (2) which comes first in the clause, are logically equivalent. These issues, however, are important for this chapter since XSOV places S before O but does not place S in the clause-initial position. In this chapter, I will address two issues: (1) whether the principles governing order are crucially a matter of what precedes what or what comes first; and (2) whether the principle governing order in the 2NPs construction is valid with the ≥3NPs construction.

Tables (1)-(3) display the number of tokens according to the types of constituent ordering and the kinds of other phrases (X)[1]: Table (1) for the 3NPs construction, Table (2) for the 4NPs construction and Table (3) for the 5NPs construction. There are tokens containing two of the same type of NPs that may or may not be coreferential. For

[1] See Section 2.4.4. in Chapter 2 for the examples of different types of word orders with different kinds of Xs.

example, there are two instances of OSOV in the 3NPs construction that contains two objects, and in both instances, the two objects are not coreferential. In the instance of SLLPOV in the 5NPs construction, the two locatives are coreferential. Clause (1) is an example of two objects that are not coreferential, and clause (2) is that of two locatives that are coreferential. The two objects in (1) and the two locatives in (2) are underlined.

(1)

ano	hito	no	o-sibai	o	boku	wa
that	person	GEN	POL-play	ACC	1sg.male	TOP

urakata	o	tetudatta
backstage.work	ACC	assist.PST

'I assisted that person with backstage work.'

(2)

boku	wa	soko	de	zyazu	konsaato	de
1sg.male	TOP	there	LOC	jazz	concert	LOC

syooappu-suru	no	ni	isu	motte-iku
show.up-do	NMLZ	Purpose	chair	take-go

'There, in the jazz concert, I bring a chair with me for a show.'

Table (1) The number of tokens with the 3NPs construction according to the types of word orders and the kinds of X

3NPs	Subject Initial		X Initial		Object Initial		Total
Type of X	SOXV	SXOV	XSOV	XOSV	OSXV	OXSV	
	#	#	#	#	#	#	#
Dative (D)	SODV	SDOV					
	3	2					5
Locative (L)	SOLV	SLOV	LSOV	LOSV	OSLV		
	2	6	7	2	4		21
Instrumental (I)	SOIV	SIOV				OISV	
	2	4				1	7
Comitative (Com)		SComOV	ComSOV				
		2	1				3
Quotative (Q)	SOQV	SQOV	QSOV		OSQV		
	5	1	2		3		11
Temporal (T)		STOV	TSOV	TOSV	OSTV	OTSV	
		13	13	1	2	1	30
Purpose (P)			PSOV				
			1				1
Direct Object (O)					OSOV		
					2		2
Complement (C)	SOCV				OSCV		
	1				2		3
Topical (To)			ToSOV				
			4				4
Adverbial (A)		SAOV	ASOV		OSAV		
		4	1		1		6
Sub-total	13	32	29	3	14	2	93
Total	45		32		16		93

First, Table (1) shows that S-initial is most frequent (48.4%, i.e. 45/93 tokens) followed by X-initial (34.4%, i.e. 32/93 tokens), and O-initial is least frequent (17.2%, i.e. 16/93 tokens). The table also shows three preferences: SO > OS by 74 to 19, SX > XS by 59 to 34, and XO > OX by 64 to 29.

Table (2) The number of tokens with the 4NPs construction according to the type of word orders and the kinds of X

4NPs	S-initial				X-initial		O-initial		Total
Type of X	SXOXV	#	SXXOV	#	XSXOV	#	OXOSV	#	#
L					LSLOV	2			2
T			STTOV	1	TSTOV	1			2
L & T			SLTOV	1					1
T & L			STLOV	1					1
Com & L			SCom LOV	1					1
T & Com			STComOV	1					1
T & Ins	STOInsV	1	STInsOV	2					3
A					ASAOV	1			1
Ins							OInsOSV	1	1
A & Ins			SAInsOV	1					1
A & Q			SAQOV	1					1
A & D			SADOV	1					1
Sub-total		1		10		4		1	16
Total	11				4		1		16

L=locative, T=temporal, Com=comitative, Ins=instrumental, A=Adverbial, Q=quotative, D=dative

Table (3) The number of tokens with the % NPs construction according to the type of word orders and the kinds of X

5NPS	S-initial		X-initial						Total
Type of X	SXXXOV	#	XXSXOV	#	XOSOXV	#	XOSSSV	#	#
To							ToSSSV	1	1
T & L	STTLOV	1							1
T & Q					TOSOQV	1			1
L & P	SLLPOV	1							1
T & L & D			TLSDOV	1					1
Sub-total		2		1		1		1	5
Total	2		3						5

To=topical, T=temporal, L=locative, Q=quotative, D= dative

Similar to Table (1), Tables (2) and (3) shows that S-initial prevails the most, then X-initial, and the occurrence of O-initial is least frequent. The tables also show the three preferences: SO > OS by 19 to 3 (XOSOXV token is counted for SO and OS), SX > XS by 19 to 8 (XSXOV, XXSXOV and XOSOXV tokens are counted for SX and XS), and XO > OX by 21 to 3 (SXOXV, OXOSV and XOSOXV tokens are counted for XO and OX).

In what follows, I will first examine the factors governing the ordering choice of XSOV and XOSV orders, and next I will study the factors we identified in Chapter 2 is working for the choice of constituent ordering in the ≥3NPs construction.

4.1. What precedes what vs. what comes first: XSOV vs. XOSV

In this section, I will examine the property of the X in XSOV and XOSV clauses in terms of six factors RD, TP, SW, particle *wa*, demonstrative and cohesive object. The following is the definition of six factors that we will examine.

(1)	RD:	The RD of X is lower than the RD of S and O.
(2)	TP:	The TP of X is higher than the TP of S and O.
(3)	SW:	The SW of X is higher than the SW of S and O.
(4)	Particle *wa*:	The X is marked by *wa* while the S and O are not marked by *wa*.
(5)	Demonstrative:	The X is demonstrative while the S and O are not demonstrative.
(6)	Cohesive objects:	The X occurs with a cohesive object.

When the RD of X is lower than the RD of S and O, the factor RD favors X-initial ordering over S-initial and O-initial. When the TP of X is higher than the TP of S and O, the factor TP favors X-initial. When the SW of X is higher than the SW of S and O, the factor SW favors X-initial. When the X is marked by *wa* while the S and O are not, the factor particle *wa* favors X-initial. When the X is demonstrative while the S and O are not, the factor demonstrative favors X-initial. When the X occurs with a cohesive object, the factor Cohesion favors X-initial. If the X in both XSOV and XOSV is proved to be favored by the factors defined above, we can assume that those factors governing the choice of SOV vs. OSV are now playing a role in predicting what the initial constituent would be in the 3NPs construction, and thus those factors are more relevant to what comes first than what precedes what. Table (4) shows the factors that motivated the occurrence of XSOV and XOSV. In Table (4), for example, when RD is a sole factor for

X-initial ordering, there is one token of XSOV; when both factors RD and demonstrative are the motivation of X-initial ordering and nothing else, there is one token of XOSV; and when nothing else than all three factors TP, SW and particle *wa* favor X-initial, there is one token of XSOV.

Table (4) Factors that the X in XSOV and XOSV demonstrate.

Factors	X in XSOV		X in XOSV		Total	
	#	%	#	%	#	%
Single factor						
RD	1	100			1	100
TP	1	100			1	100
SW	2	100			2	100
Particle *wa*	1	50	1	50	2	100
Demonstrative	2	100			2	100
Cohesive objects	3	100			3	100
Two factors						
RD and SW	1	100			1	100
RD and Demonstrative			1	100	1	100
RD and Cohesive objects	1	100			1	100
TP and SW	1	100			1	100
SW and Demonstrative	2	100			2	100
SW and Cohesive objects	4	100			4	100
Demonstrative and Cohesive objects	2	100			2	100
Three factors						
RD, Particle *wa* and Demonstrative	1	100			1	100
RD, TP and Demonstrative	1	100			1	100
RD, TP and Particle *wa*			1	100	1	100
TP, SW and Particle *wa*	1	100			1	100
RD, Demonstrative and Cohesive objects	1	100			1	100
SW, Demonstrative and Cohesive objects	1	100			1	100
Four factors						
RD, SW, Demonstrative and Cohesive objects	1	100			1	100
None	3	100			3	100
Total	29	90.6	3	9.4	32	100

The numbers in Table (4) are too small to be statistically significant. Table (4) shows that in 89.7% (26/29 tokens) of XSOV and 100% of XOSV, X-initial ordering is favored by the factors governing the ordering choice between SOV and OSV. The result suggests that the factors that place the S before the O in SOV and place the O before the S in OSV also work for placing the X before the S and O in XSOV and XOSV clauses. Following this result, I will look at tokens in the ≥3NPs construction to determine whether the occurrence of initial constituents in the ≥3NPs construction is motivated by the factors we identified in Chapter 3.

4.2. S-initial, O-initial, and X-initial Orders

In this section I will study how the factors RD, ERD, TP, SW, particle *wa*, demonstrative and cohesion of object and verb, can work for choosing the order of Subject-initial, Object-initial, and X-initial in ≥3NPs constructions containing the subject (S), the direct object (O) and other phrases (X). Table (5) gives the number of S-initial, O-initial and X-initial tokens. The figures were taken from Tables (1)-(3). A hundred fourteen tokens in the data will be examined in this section.

Table (5) The number of tokens with ≥3NPs constructions

# of NPs	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
3NPs	45	48.4	16	17.2	32	34.4	93	100
4NPs	11	68.8	1	6.3	4	25.0	16	100
5NPs	2	40.0	0	0.0	3	60.0	5	100
Total	58	50.9	17	14.9	39	34.2	114	100

4.2.1. Single variable

4.2.1.1. Referential distance (RD)

In this section, I will examine if RD has any relevance to constituent ordering in S-initial, X-initial and O-initial clauses with the ≥3NPs construction. Table (6) shows the number of S-initial, O-initial and X-initial tokens under the seven conditions defined according to the relative RD of the constituents. The seven conditions are as follows:

(1) The RD of S is lower than the RD of O and X
(2) The RD of O is lower than the RD of S and X
(3) The RD of X is lower than the RD of S and O
(4) The RD of S and O is the same, but lower than the RD of X

(5) The RD of S and X is the same, but lower than the RD of O
(6) The RD of O and X is the same, but lower than the RD of S
(7) The RD of all three is the same

Table (6) The number of tokens under the seven conditions of relative RD

Seven Conditions of Relative RD	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
1 RD of S < RD of O and X	36	63.2	3	5.3	18	31.6	57	100
2 RD of O < RD of S and X	6	35.3	9	52.9	2	11.8	17	100
3 RD of X < RD of S and O	5	31.3	1	6.3	10	62.5	16	100
4 (RD of S = RD of O) < RD of X	2	33.3	2	33.3	2	33.3	6	100
5 (RD of S = RD of X) < RD of O	2	50.0	0	0.0	2	50.0	4	100
6 (RD of O = RD of X)< RD of S	0	0.0	1	100.0	0	0	1	100
7 RD of S = RD of O = RD of X	7	53.8	1	7.7	5	38.5	13	100
Total	58	50.9	17	14.9	39	34.2	114	100

Table (6) shows that the occurrence of S-initial is highest when the RD of S is lower than the RD of O and X, i.e. the condition (1), the occurrence of O-initial is highest when the RD of O is lower than the RD of S and X, i.e. the condition (2), and the occurrence of X-initial is highest when the RD of X is lower than the RD of S and O, i.e. the condition (3); 63.2% for S-initial, 52.9% for O-initial and 62.5% for X-initial. Moreover, the occurrence of S-initial under the condition (1) is higher than the occurrence of non-S-initial (63.2% for S-initial vs. 36.9% for non-S-initial), the occurrence of O-initial under the condition (2) is higher than the occurrence of non-O-initial (52.9% for O-initial and 47.1% for non-O-initial), and the occurrence of X-initial under the condition (3) is higher than that of non-X-initial (62.5% for X-initial vs. 37.5% for non-initial). These results indicate that the relative RD is relevant to the ordering choice among S-initial, O-initial or X-initial. The table also shows that when the RD of all three is the same, the percentage of occurrence is highest with S-initial and lowest with O-initial (53.8% for S-initial, 38.5% for X-initial and 7.7% for O-initial).

Furthermore, it is clear from Table (6) that the RD is more relevant to predicting S-initial and O-initial than X-initial. For example, 62.1% of initial subjects (i.e. 6/ 58 tokens) are ones where the RD of the S is lowest, and 52.9% of initial objects (i.e. 9/17 tokens) are ones where the RD of the O is lowest, while only 25.6% of initial Xs (i.e. 10/39) are ones where the RD of the X is lowest. One reason that we can consider is that the initial Xs are expected to be temporal or locative expressions which are scene-setting topics by specifying the time and location of what the speaker is about to talk about, and thus they are expected to be first mentions or not to have a referent in the recent discourse. In fact, 28.2% of X-initial clauses (i.e. 11/39 tokens) are locative-initial, and 43.6% of X-initial clauses (i.e. 17/39 tokens) are temporal-initial. We will further discuss the scene-setting constituents later in this chapter.

4.2.1.2. Extended referential distance (ERD)

In Chapter 3, we learned that Algorithm (13) using the ERD that measures the RD of the initial constituent of the subject or the object produces a better correct prediction rate than Algorithm (1b) using the RD of the subject or the object as a factor. Table (7) shows the number of S-initial, O-initial and X-initial tokens under the seven conditions defined according to the relative ERD of the S, O and X. The seven conditions are as follows:

(1) The ERD of S is lower than the ERD of O and X
(2) The ERD of O is lower than the ERD of S and X
(3) The ERD of X is lower than the ERD of S and O
(4) The ERD of S and O is the same, but lower than the ERD of X
(5) The ERD of S and X is the same, but lower than the ERD of O
(6) The ERD of O and X is the same, but lower than the ERD of S
(7) The ERD of all three is the same

Table (7) The number of tokens under the seven conditions of relative ERD

Seven Conditions of relative ERD		S-initial		O-initial		X-initial		Total	
		#	%	#	%	#	%	#	%
1	ERD of S < ERD of O and X	30	62.5	2	4.2	16	33.3	48	100
2	ERD of O < ERD of S and X	6	27.3	12	54.5	4	18.2	22	100
3	ERD of X < ERD of S and O	9	50.0	0	0.0	9	50.0	18	100
4	(ERD of S = ERD of O) < ERD of X	1	20.0	2	40.0	2	40.0	5	100
5	(ERD of S = ERD of X) < ERD of O	2	40.0	0	0.0	3	60.0	5	100
6	(ERD of O = ERD of X) < ERD of S	2	50.0	1	25.0	1	25.0	4	100
7	ERD of S = ERD of O = ERD of X	8	66.7	0	0.0	4	33.3	12	100
Total		58	50.9	17	14.9	39	34.2	114	100

Table (7) shows that the percentage of occurrence of S-initial when the ERD of S is lower than the ERD of O and X is slightly lower than the percentage when the RD of S is lower than the ERD of O and X (62.5% with the ERD and 63.2% with the RD). In contrast, the percentage of occurrence of O-initial when the ERD of O is lower than the ERD of S and X is slightly higher than the percentage when the RD of O is lower than the RD of S and X (54.5% with the ERD and 52.9% with the RD). The table also shows that the occurrence of X-initial when the RD of X is lower than the RD of S and O is higher than when we measure the ERD of the constituents (50% with the ERD and 62.5% with the RD).

Moreover, it is clear from Table (7) that the ERD is most relevant to choosing O-initial than S-initial or X-initial. For example while 70.6% of initial objects (i.e. 12/17 tokens) are ones where the ERD of the O is lowest, 51.7% of initial subjects (i.e. 30 out of 58) are ones where the ERD of the S is lowest, and only 23.2% of initial Xs (i.e. 9/39 tokens) are ones where the ERD of the X is lowest. This result parallels the result by the RD, which indicates that the RD is not a determining factor for X-initial order.

4.2.1.3. Topic persistence (TP)

In this section, I will examine if the relative TP is a relevant factor for the choice of S-initial, O-initial and X-initial in the ≥3NPs construction. Table (8) shows the number of S-initial, O-initial and X-initial tokens under the seven conditions defined by the relative TP of the constituents.

(1) The TP of S is higher than the TP of O and X
(2) The TP of O is higher than the TP of S and X
(3) The TP of X is higher than the TP of S and O
(4) The TP of S and O is the same, but higher than the TP of X
(5) The TP of S and X is the same, but higher than the TP of O
(6) The TP of O and X is the same, but higher than the TP of S
(7) The TP of all three is the same

Table (8) The number of tokens under the seven conditions of relative TP

Seven Conditions of relative TP	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
1 TP of S > TP of O and X	29	50.0	9	15.5	20	34.5	58	100
2 TP of O > TP of S and X	11	52.4	4	19.0	6	28.6	21	100
3 TP of X > TP of S and O	2	25.0	1	12.5	5	62.5	8	100
4 (TP of S = TP of O) > TP of X	2	40.0	2	40.0	1	20.0	5	100
5 (TP of S = TP of X) > TP of O	2	100.0	0	0.0	0	0.0	2	100
6 (TP of O = TP of X) > TP of S	1	100.0	0	0.0	0	0.0	1	100
7 TP of S = TP of O = TP of X	11	57.9	1	5.3	7	36.8	19	100
Total	58	50.9	17	14.9	39	34.2	114	100

Table (8) shows that the proportion of occurrences when the TP of S is higher than the TP of O and X (50% for S-initial, 15.5% for O-initial and 34.5% for X-initial), and when the TP of O is higher than the TP of S and X (52.5% for S-initial, 19% for O-initial and 28.6% for X-initial) is not much different from the proportion of overall occurrences (50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial). These results indicate that TP is not relevant in determining S-initial and O-initial orders. When the TP of X is higher than the TP of S and O, however, the percentage of occurrence is highest with X-initial (62.5%), which is much higher than 34.2%, overall occurrence of X-initial. This result indicates that TP might be an indicator for choosing X-initial.

4.2.1.4. Syntactic weight (SW)

In this section, I will examine if the relative SW is relevant to determining S-initial, O-initial or X-initial order in the ≥3NPs construction. Table (9) shows the number of tokens for S-initial, O-initial and X-initial under the seven conditions defined by the relative SW of the constituents.

(1) The SW of S is higher than the SW of O and X
(2) The SW of O is higher than the SW of S and X
(3) The SW of X is higher than the SW of S and O
(4) The SW of S and O is the same, but higher than the SW of X
(5) The SW of S and X is the same, but higher than the SW of O
(6) The SW of O and X is the same, but higher than the SW of S
(7) The SW of all three is the same

Table (9) The number of tokens under the seven conditions of relative SW

Seven Conditions of relative SW	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
1 SW of S > SW of O and X	8	42.1	3	15.8	8	42.1	19	100
2 SW of O > SW of S and X	20	57.1	7	20.0	8	22.9	35	100
3 SW of X > SW of S and O	21	52.5	4	10.0	15	37.5	40	100
4 (SW of S = SW of O) > SW of X	0	0.0	1	100.0	0	0.0	1	100
5 (SW of S = SW of X) > SW of O	5	55.6	0	0.0	4	44.4	9	100
6 (SW of O = SW of X) > SW of S	0	0.0	0	0.0	2	100.0	2	100
7 SW of S = SW of O = SW of X	4	50.0	2	25.0	2	25.0	8	100
Total	58	50.9	17	14.9	39	34.2	114	100

Table (9) shows that when the SW of S is higher than the SW of O and X, the occurrence of S-initial is no more frequent than the overall percentage of 50.9%; and when the SW of X is higher than the SW of S and O, the occurrence of S-initial is higher than that of X-initial (52.5% for S-initial and 37.5% for X-initial). While the occurrence of O-initial is lowest when the SW of O is higher than the SW of S and X, the percentage of occurrence (20%) is higher than the overall occurrence of O-initial (14.9%). Crucially, the percentage of occurrences, 42.1% for S-initial, 20% for O-initial and 37.5% for X-initial when the SW favors S-initial, O-initial or X-initial is little different from the overall percentages, 50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial. These results indicate that there is little evidence for SW being a relevant factor for ordering.

4.2.1.5. RD, ERD, TP, and SW

Table (10) summarizes the percentages of occurrences of S-initial, O-initial and X-initial when each factor RD, ERD, TP and SW favors S-initial, O-initial and X-initial. The numbers were pulled out from Tables (6)-(9).

Table (10) Comparison of factors

Factor	S-initial	O-initial	X-initial
	%	%	%
RD	63.2	52.9	62.5
ERD	62.5	54.5	50.0
TP	50.0	19.0	62.5
SW	42.1	20.0	37.5
Overall Occurrence	50.9	14.9	34.2

Table (10) shows that the percentage of occurrence of S-initial when RD or ERD favors S-initial is greater than the overall percentage of S-initial, 50.9%, while it is not greater than 50.9% when TP or SW favors S-initial. Similarly, the percentage for O-initial is higher than the overall percentage of O-initial, 14.9% when RD or ERD favors O-initial. For X-initial, the percentage of occurrence is greater than 34.2%, overall percentage of X-initial when RD, ERD or TP favors X-initial. From Table (10) we can conclude that RD (or ERD) is relevant to preferring S-initial and O-initial, but TP and SW are not, and that RD (or ERD) and TP are relevant to preferring X-initial, but SW is not.

4.2.1.6. Particle *wa*

In Chapter 2, we learned that constituents marked with the particle *wa* tend to appear in the clause-initial position. In Chapter 3, Algorithm (16) using the particle *wa*, when it is incorporated into Algorithm (13) based on the relative ERD, TP and SW, predicted the ordering choice of SOV vs. OSV for 84.4% of the data. In this section, I will examine if the particle *wa* is a relevant factor for choosing S-initial, O-initial and X-initial orders. In the present data, there are 42 tokens that contain *wa*-marked constituents. Table (11) shows the number of tokens and the percentages of occurrence for each order.

Table (11) The number of tokens containing *wa* -marked constituents

Wa -marking	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is marked with *wa*	18	69.2	4	15.4	4	15.4	26	100
O is marked with *wa*	3	50.0	3	50.0	0	0.0	6	100
X is marked with *wa*	1	16.7	0	0.0	5	83.3	6	100
S and O are marked with *wa*	1	50.0	1	50.0	0	0.0	2	100
S and X are marked with *wa*	0	0.0	0	0.0	1	100	1	100
O and X are marked with *wa*	0	0.0	0	0.0	1	100	1	100
None of S, O or X is marked with *wa*	35	48.6	9	12.5	28	38.9	72	100
Total	58	50.9	17	14.9	39	34.2	114	100

Table (11) shows that when the subject is marked with *wa*, the occurrence of S-initial is most frequent (69.2% for S-initial, 15.4% for O-initial and for X-initial). The percentage, 69.2% is much higher than the overall occurrence of S-initial, i.e. 50.9%. The table also shows that when the X is marked with *wa*, the occurrence of X-initial is the most frequent (83.3% for X-initial, 16.7% for S-initial, and 0% for O-initial), and 83.3% is much greater than 34.2%, which is the overall percentage of X-initial. Although the occurrence of O-initial is not highest (50% for S-initial and O-initial) when the object is marked with *wa*, the percentage, 50% is much greater than the overall occurrence of O-initial, i.e. 14.9%. These results suggest that marking with the particle *wa* is a relevant factor for preferring S-initial, O-initial and X-initial orders.

4.2.1.7. Demonstratives

In Chapter 2, we learned that demonstrative constituents tend to appear in the clause-initial position, and in Chapter 3, Algorithm (18) using demonstrative objects as a factor, when it was incorporated into Algorithm (13) based on ERD, TP and SW, predicted the choice between SOV and OSV for 80.2% of the data. In this section, I will study if demonstrative is a relevant factor for choosing S-initial, O-initial and X-initial orders. In the present data, there are 50 tokens containing demonstrative constituents. Table (12) shows the number of tokens and the percentage of occurrence for each order.

Table (12) The number of tokens containing demonstrative constituents

Demonstrative	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is demonstrative	8	88.9	1	11.1	0	0.0	9	100
O is demonstrative	2	18.2	9	81.8	0	0.0	11	100
X is demonstrative	8	38.1	0	0.0	13	61.9	21	100
S and O are demonstrative	0	n/a	0	n/a	0	n/a	0	n/a
S and X are demonstrative	1	50.0	0	0.0	1	50.0	2	100
O and X are demonstrative	2	28.6	2	28.6	3	42.9	7	100
None of constituents is demonstrative	37	57.8	5	7.8	22	34.4	64	100
Total	58	50.9	17	14.9	39	34.2	114	100

Table (12) shows that when the subject is demonstrative, the percentage of occurrence is highest with S-initial (88.9% for S-initial vs. 11.1% for O-initial), when the object is demonstrative, the percentage of occurrence is highest with O-initial (81.8% for O-initial, 18.2% for S-initial), and when the X is demonstrative, the percentage of occurrence is highest with X-initial (61.9% for X-initial and 38.1% for S-initial). These percentages, 88.9% for S-initial, 81.8% for O-initial, and 61.9% for X-initial are much greater than the overall occurrence of each order (50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial). This result suggests that the factor demonstrative play a role in the choice of S-initial, O-initial and X-initial orders.

4.2.1.8. Cohesive objects

In Chapter 3, Algorithm (17) using the cohesion of object and verb as a factor predicted the choice of SOV vs. OSV at a success rate of 80.8%. The reasoning for using the cohesion of object and verb for the algorithm is that the cohesion of the object, which is the object in light verb construction or idiomatic expressions, to the verb tends to place the object adjacent to the verb, resulting in SOV order. In the 2NPs construction, SOV is a logical consequence of the algorithm using cohesive objects since the choice is between SOV and OSV. In the ≥3NPs construction, however, this logic does not work. In the 3NPs construction, for example, there are two choices, XSOV and SXOV orders when the cohesion of object and verb anchors the object in the immediate preverbal position. As going through the tokens in the ≥3NPs construction, I noticed that the majority of cohesive objects appear immediately preverbally. Table (13) shows the number of tokens containing a cohesive object in the ≥3NPs construction.

Table (13) The number of tokens containing a cohesive object

	Cohesive objects	with cohesive objects		without cohesive objects		Total	
		#	%	#	%	#	%
Object = immediately preverbal	SXXOV	6	60.0	4	40.0	10	100
	SXOV	16	50.0	16	50.0	32	100
	XSXOV	2	50.0	2	50.0	4	100
	XSOV	12	41.4	17	58.6	29	100
	SXXXOV	0	0.0	2	100.0	2	100
	OSOV	0	0.0	2	100	2	100
	XXSXOV	0	0.0	1	100	1	100
	sub-total	*36*	*45.0*	*44*	*55.0*	*80*	*100*
Object ≠ immediately preverbal	OSXV	2	16.7	10	83.3	12	100
	SOXV	1	7.7	12	92.3	13	100
	XOSV	0	0.0	3	100.0	3	100
	OXSV	0	0.0	2	100.0	2	100
	SXOXV	0	0.0	1	100.0	1	100
	OXOSV	0	0.0	1	100.0	1	100
	XOSOXV	0	0.0	1	100.0	1	100
	XOSSSV	0	0.0	1	100.0	1	100
	Total	39	34.2	75	65.8	114	100

Table (13) shows that 92.3% (i.e. 36/39 tokens) of tokens containing a cohesive object are ones in which the object appears in the immediately preverbal position. The table also shows that among the 80 tokens in which the object appears in the immediately preverbal position, 55% of tokens do not contain cohesive objects. This result suggests that when the object in the ≥3NPs construction clauses is cohesive, it is most likely to appear in the immediately preverbal position.

4.2.1.9. Locatives

In Section 4.2.1.1., we learned that the RD is irrelevant to the choice of X-initial. In this section, we will examine if the semantic type of the X has any relevance in determining where the X appears in a clause. Table (14) shows the position of locatives and their semantic types.

Table (14) Type and Position of Locatives

	Position of locative	At-locative		To- or From locative		Total
		#	%	#	%	#
Locative = initial	LSOV	6	*25.0*	1	12.5	7
	LOSV	2	*8.3*	0	0.0	2
	LSLOV	2	*8.3*	0	0.0	2
	sub-total	*10*	*41.7*	*1*	*9.1*	*11*
Locative ≠ initial	SOLV	0	*0.0*	2	100.0	2
	SLOV	4	*16.7*	2	33.3	6
	SLTOV	1	*4.2*	0	0.0	1
	STLOV	1	*4.2*	0	0.0	1
	SComLOV	1	*4.2*	0	0.0	1
	STTLOV	0	*0.0*	1	100.0	1
	S**L**LPOV	1	*4.2*	0	0.0	1
	SL**L**POV	1	*4.2*	0	0.0	1
	TLSDOV	1	*4.2*	0	0.0	1
	LS**L**OV	1	*4.2*	1	50.0	2
	OSLV	3	*12.5*	1	25.0	4
	sub-total	*14*	*58.3*	*7*	*33.3*	*21*
	Total	24	*100.0*	8	25.0	32

At-locatives are often scene-setting elements, and to- or from-locatives are not. The numbers in Table (14) were summarized in Table (15) in order to determine if there is a relation between the scene-setting locatives and clause-initial position.

Table (15) Locatives in initial and non-initial position

Type of locative	Locative = initial		Locative ≠ initial		Total	
	#	%	#	%	#	%
At-locative	10	41.7	14	58.3	24	100
To- or from-locative	1	12.5	7	87.5	8	100
Total	11	34.4	21	65.6	32	100

Table (15) shows that all locatives appear more commonly in non-initial position. The table also shows that almost half the at-locatives appear in the initial position, while a relatively greater number of to/from-locatives are non-initial. The result indicates that locatives occur in the initial position more frequently when they are at-locatives than when they are to/from-locatives.

4.2.1.10. Temporals

In the present data, there are three types of temporal constituents, time, duration and frequency. ‘Time’ is a scene-setting element while ‘duration’ and ‘frequency’, which are often secondary information subordinate to ‘time’ temporal, are not. Table (16) shows the type and position of temporal constituents.

Table (16) Type and Position of Temporals

	Position of temporals	Time #	Time %	Duration or Frequency #	Duration or Frequency %	Total #
Temporal = initial	TSOV	12	92.3	1	7.7	13
	TSTOV	1	100.0	0	0.0	1
	TLSDOV	1	100.0	0	0.0	1
	TOSV	1	100.0	0	0.0	1
	TOSOQV	1	100.0	0	0.0	1
	sub-total	*16*	*94.1*	*1*	*5.9*	*17*
Temporal ≠ initial	STOV	13	100.0	0	0.0	13
	SLTOV	0	0.0	1	100.0	1
	STLOV	0	0.0	1	100.0	1
	S**T**TOV	1	100.0	0	0.0	1
	ST**T**OV	0	0.0	1	100.0	1
	STComOV	1	100.0	0	0.0	1
	STInsOV	1	50.0	1	50.0	2
	STOInsV	1	100.0	0	0.0	1
	S**T**TLOV	1	100.0	0	0.0	1
	ST**T**LOV	0	0.0	1	100.0	1
	TS**T**OV	0	0.0	1	100.0	1
	OSTV	2	100.0	0	0.0	2
	OTSV	1	100.0	0	0.0	1
	sub-total	*21*	*77.8*	*6*	*22.2*	*27*
	Total	37	84.1	7	15.9	44

There are two S-initial tokens and one X-initial token that contain two temporal constituents. In Table (16), I coded the type and position for each instance of temporal constituents. Temporal constituents in those tokens, TSTOV, STTOV and STTLOV, are indicated in bold face. When there are two temporal constituents in a clause, my data shows that the first one is a scene-setting 'time' element, and the second one indicates either the duration or frequency of the event/action. The figures in Table (16) were summarized in Table (17), which shows the relation between the type of temporal and the initial position.

Table (17) The number of tokens containing tempral in the clause-initial position

Type of temporal	Temporal = initial #	Temporal = initial %	Temporal ≠ initial #	Temporal ≠ initial %	Total #	Total %
Time	16	43.2	21	56.8	37	100
Duration or Frequency	1	14.3	6	85.7	7	100
Total	17	38.6	27	61.4	44	100

Table (17) shows that all temporals occur more commonly in non-initial position. While almost half of time expressions appear in the initial position, relatively few duration- or frequency-temporals do so. This result suggests that temporals occur in the initial position more frequently when they are scene-setting elements indicating 'time' than when they are non-scene-setting elements indicating 'duration' or 'frequency'.

4.2.1.11. Topical elements

Topical element, or topics, typically "set a spatial, temporal, or individual framework within which the main predication holds" (Chafe 1976:50), and they are constituents that are not subcategorized by the verb[2]. In Japanese, they are often marked with particle *wa* or *mo*[3]. Clause (3) is an example of clauses containing topical elements.

(3)	jurasikku	paaku	tte	no	wa	watasi	hon	o	yom-imasita
	Jurassic	park	QT	GEN	TOP	1sg	book	ACC	read-PST
	'Jurassic Park (topic), I read the book.'								

Topical elements are by definition other phrases than the subject and object. There are five tokens containing a topical element in the present data of 114 tokens. In all instances, the topical elements are clause-initial.

---

[2] See Li (1976) and Chafe (1976) for discussion on linguistic properties of 'topics'.

[3] *Wa* and *mo* are adverbial particles that belong to the same class (*kakari-zyosi*) in traditional Japanese grammar. Both *wa* and *mo* are set-anaphoric. *Wa* indicates an elimination of an entity from the set, and *mo* indicates an addition of an entity to the set.

4.2.2. Two variables

4.2.2.1. RD and TP

Tables (18) and (19) show the interaction of the seven conditions defined in Sections 4.2.1.1. and 4.2.1.3. according to the relative RD and TP of constituents. The seven conditions are repeated here.

<u>Seven Conditions for RD</u>
(1) The RD of S is lower than the RD of O and X
(2) The RD of O is lower than the RD of S and X
(3) The RD of X is lower than the RD of S and O
(4) The RD of S and O is the same, but lower than the RD of X
(5) The RD of S and X is the same, but lower than the RD of O
(6) The RD of O and X is the same, but lower than the RD of S
(7) The RD of all three is the same

<u>Seven Conditions for TP</u>
(1) The TP of S is higher than the TP of O and X
(2) The TP of O is higher than the TP of S and X
(3) The TP of X is higher than the TP of S and O
(4) The TP of S and O is the same, but higher than the TP of X
(5) The TP of S and X is the same, but higher than the TP of O
(6) The TP of O and X is the same, but higher than the TP of S
(7) The TP of all three is the same

Table (18) shows the number of tokens, and Table (19) the percentage of occurrence compared to the other two orders under the same condition. In both tables, “S” stands for S-initial, “O” for O-initial and “X” for X-initial orders. In Table (19), the percentages that are greater than the overall percentages of each order (i.e. 50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial) are indicated in italics.

Table (18) Interaction of RD and TP under the 7 conditions: the number of tokens

		Conditions of RD																						
		1			2			3			4			5			6			7			Total		
	Order	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of TP	1	22	3	12	1	4	0	1	1	4	2	0	1	1	0	1	0	1	1	2	0	2	29	9	21
	2	5	0	2	4	2	0	0	0	1	0	2	0	0	0	1	0	0	0	2	0	2	11	4	6
	3	2	0	1	0	1	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2	1	5
	4	1	0	1	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1
	5	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
	6	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	7	5	0	2	1	0	0	1	0	3	0	0	1	1	0	0	0	0	0	3	1	1	11	1	7
Total		36	3	18	6	9	2	5	1	10	2	2	2	2	0	2	0	1	1	7	1	5	58	17	40

Table (19) Interaction of RD and TP under the 7 conditions in percentages

		Conditions of RD																				
		1			2			3			4			5			6			7		
	Order	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of TP	1	*59*	8	32	20	*80*	0	17	*17*	*67*	*67*	0	33	50	0	*50*	0	*50*	*50*	50	0	*50*
	2	*71*	0	29	*67*	*33*	0	0	0	*100*	0	*100*	0	0	0	*100*	0	0	0	50	0	*50*
	3	*67*	0	33	0	*33*	*67*	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0
	4	50	0	*50*	0	*100*	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	*100*	0	0	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	*71*	0	29	*100*	0	0	25	0	*75*	0	0	*100*	*100*	0	0	0	0	0	*60*	*20*	20

Table (19) shows that under the condition of RD = (1) and TP = (1), i.e. both RD and TP favor S-initial, the percentage of S-initial is higher than the overall percentage (59.5% > 50.9%). The table also shows that under the condition of RD = (2) and TP = (2), i.e. both RD and TP favor O-initial, the percentage of O-initial is higher than the overall percentage (33.3% > 14.9%) although O-initial is a minority (2 out of 6 tokens). Moreover, under the condition of RD = (3) and TP = (3), i.e. both RD and TP favor X-initial, the percentage of X-initial is greater than the overall percentage (100% > 34.2%).

Table (20) shows the number of tokens and the percentage of occurrence under certain conditions defined according to the interaction of the seven conditions for RD and TP. The numbers were taken from Table (18). In Table (20), the percentages of occurrence as opposed to the other two orders that are higher than the overall percentages of each order (i.e. 50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial) are indicated in italics.

Table (20) The occurrence of S-initial, O-initial and X-initial clauses under the conditions according to the interaction of the seven conditions for RD and TP

Condition	RD & TP	S-initial		O-initial		X-initial		Total	
		#	%	#	%	#	%	#	%
A	RD=(1) while TP≠(2) or (3), or TP=(1) while RD≠(2) or (3)	34	*59.6*	4	7.0	19	33.3	57	100
B	RD=(2) while TP≠(1) or (2), or TP=(2) while RD≠(1) or (3)	7	43.8	6	*37.5*	3	18.8	16	100
C	RD=(3) while TP=(1) or (2), or TP=(3) while RD≠(1) or (2)	4	44.4	0	0.0	5	*55.6*	9	100
D	RD=(1) while TP=(2)	5	*71.4*	0	0.0	2	28.6	7	100
E	RD=(1) while TP=(3)	2	*66.7*	0	0.0	1	33.3	3	100
F	RD=(2) while TP=(1)	1	20.0	4	*80.0*	0	0.0	5	100
G	RD=(2) while TP=(3)	0	0.0	1	*33.3*	2	*66.7*	3	100
H	RD=(3) while TP=(1)	1	16.7	1	*16.7*	4	*66.7*	6	100
I	RD=(3) while TP=(2)	0	0.0	0	0.0	1	*100.0*	1	100
J	Other cells	4	*57.1*	1	14.3	2	28.6	7	100
	Total	58	50.9	17	14.9	39	34.2	114	100

Table (20) shows that when the condition of RD is (1) while the condition of TP is not (2) or (3), or when the condition of TP is (1) while the condition of RD is not (2) or (3), i.e. Condition A where either RD or TP favors S-initial without conflicting with other conditions, the percentage of S-initial is higher than the overall percentage (59.6% > 50.9%). The table also shows that the condition of RD is (2) while the condition of TP is not (1) or (3), or when the condition of TP is (2) while the condition of RD is not (1) or (3), i.e. Condition B where either RD or TP favors O-initial without conflicting with other conditions, the percentage of O-initial is higher than the overall percentage (37.5% > 14.3%). Likewise, when the condition of RD is (3) while the condition of TP is not (1) or (2), or when the condition of TP is (3) while the condition of RD is not (1) or (2), i.e. Condition C where either RD or TP favors X-initial without conflicting with other conditions, the percentage of X-initial is greater than the overall percentage (55.6% >

34.2%). However, 59.6% for S-initial under Condition A and 37.5% for O-initial under Condition B are smaller than the percentages of occurrences for S-initial and O-initial when RD favors these orders (63.2% for S-initial and 52.9% for O-initial, cf. Table (6) in Section 4.2.1.1.). This result indicates that the factor TP does not improve on RD. Similarly, 55.6% for X-initial under Condition C is not higher than the percentage of occurrence for X-initial when RD or TP favors X-initial (62.5% with RD or TP, cf. Table (6) in Section 4.2.1.1. and Table (8) in Section 4.2.1.3.).

Moreover, when RD favors S-initial while TP favors O-initial (Condition D) or X-initial (Condition E), the percentages of S-initial is higher than the overall percentage (71.4% and 66.7% > 50/9%). Similarly, when RD favors O-initial while TP favors S-initial (Condition F) or X-initial (Condition G), the percentages of O-initial are higher than when it is overall (80% and 33.3% > 14.9%), and when RD favors X-initial while TP favors S-initial (Condition H) or O-initial (Condition I), the percentages of X-initial are greater than the overall percentage (66.7% and 100% > 34.2%).

In contrast, when TP favors S-initial while RD favors O-initial (Condition F) or X-initial (Condition H), the percentages of S-initial is lower than the overall percentage (20% and 16.7% > 50.9%), and when TP favors O-initial while RD favors S-initial (Condition E) or X-initial (Condition I), the percentages of O-initial is lower than the overall percentage of O-initial (0% > 14.9%). For X-initial, when TP favors X-initial while RD favors O-initial (Condition G), the percentage of X-initial is higher than the overall percentage (66.7% > 34.3%), whereas the percentage is lower than the overall percentage of occurrence under Condition E when TP favors X-initial while RD favors S-initial (33.3% < 34.2%). This result suggests that RD favoring S-initial is stronger than TP favoring X-initial.

4.2.2.2. RD and SW

Tables (21) and (22) show the interaction of the seven conditions defined in Sections 4.2.1.1. and 4.2.1.4. according to the relative RD and SW of constituents.

Seven Conditions for RD

(1) The RD of S is lower than the RD of O and X
(2) The RD of O is lower than the RD of S and X
(3) The RD of X is lower than the RD of S and O
(4) The RD of S and O is the same, but lower than the RD of X
(5) The RD of S and X is the same, but lower than the RD of O
(6) The RD of O and X is the same, but lower than the RD of S
(7) The RD of all three is the same

Seven Conditions for SW

(1) The SW of S is higher than the SW of O and X
(2) The SW of O is higher than the SW of S and X
(3) The SW of X is higher than the SW of S and O
(4) The SW of S and O is the same, but higher than the SW of X
(5) The SW of S and X is the same, but higher than the SW of O
(6) The SW of O and X is the same, but higher than the SW of S
(7) The SW of all three is the same

Table (21) gives the number of tokens, and Table (22) the percentage of occurrence compared to the other two orders. In Table (22), the percentages that are greater than the overall percentages of each order (i.e. 50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial), are indicated in italics.

Table (21) Interaction of RD and SW under the 7 conditions: the number of tokens

		Conditions of RD																						
		1			2			3			4			5			6			7			Total		
	Order	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of SW	1	3	0	1	2	2	0	1	0	4	0	0	0	0	0	0	0	1	0	2	0	3	8	3	8
	2	17	3	4	0	2	0	2	1	1	0	1	1	0	0	1	0	0	0	1	0	1	20	7	8
	3	12	0	11	1	2	1	1	0	1	2	1	1	2	0	0	0	0	0	3	1	1	21	4	15
	4	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
	5	2	0	0	1	0	1	1	0	2	0	0	0	0	0	1	0	0	0	1	0	0	5	0	4
	6	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
	7	2	0	1	2	2	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4	2	2
Total		36	3	18	6	9	2	5	1	10	2	2	2	2	0	2	0	1	0	7	1	5	58	17	39

Table (22) Interaction of RD and SW under the 7 conditions in percentages

		Conditions of RD																				
		1			2			3			4			5			6			7		
	Order	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of SW	1	*75*	0	25	50	*50*	0	20	0	*80*	0	0	0	0	0	0	0	*100*	0	40	0	*60*
	2	*71*	12	17	0	*100*	0	50	*25*	25	0	*50*	*50*	0	0	*100*	0	0	0	50	0	*50*
	3	*52*	0	*48*	25	*50*	25	50	0	*50*	50	*25*	25	*100*	0	0	0	0	0	*60*	*20*	20
	4	0	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	*100*	0	0	50	0	*50*	33	0	*67*	0	0	0	0	0	*100*	0	0	0	*100*	0	0
	6	0	0	*100*	0	0	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0
	7	*67*	0	33	50	*50*	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0

Table (22) shows that when the condition of RD is (1) and the condition of SW is (1), i.e. both RD and SW favor S-initial, the percentage of S-initial is higher than the overall percentage (75% > 50.9%). The table also shows that when the condition of RD is (2) and the condition of SW is (2), i.e. both RD and SW favor O-initial, the percentage of O-initial is higher than the overall percentage (100% > 14.9%). Similarly, when the condition of RD is (3) and the condition of SW is (3), i.e. both RD and SW favor X-initial, the percentage of X-initial is higher than the overall percentage (50% > 34.2%). Table (23) shows the number of tokens and the percentage of occurrence of S-initial, O-initial and X-initial clauses under the defined conditions according to the interaction of conditions for RD and SW. In Table (23), the percentages of each order compared to the other two orders that are higher than when they are overall are indicated in italics.

Table (23) The occurrence of S-initial, O-initial and X-initial clauses under the conditions based on the interaction of the seven conditions for RD and SW

Condition	RD & SW	S-initial		O-initial		X-initial		Total	
		#	%	#	%	#	%	#	%
A	RD=(1) while SW≠(2) or (3), or SW=(1) while RD≠(2) or (3)	9	*56.3*	1	6.3	6	*37.5*	16	100
B	RD=(2) while SW≠(1) or (2), or SW=(2) while RD≠(1) or (3)	4	28.6	6	*42.9*	4	28.6	14	100
C	RD=(3) while SW=(1) or (2), or SW=(3) while RD≠(1) or (2)	9	50.0	2	11.1	7	*38.9*	18	100
D	RD=(1) while SW=(2)	17	*70.8*	3	12.5	4	16.7	24	100
E	RD=(1) while SW=(3)	12	*52.2*	0	0.0	11	*47.8*	23	100
F	RD=(2) while SW=(1)	2	50.0	2	*50.0*	0	0.0	4	100
G	RD=(2) while SW=(3)	1	25.0	2	*50.0*	1	25.0	4	100
H	RD=(3) while SW=(1)	1	20.0	0	0.0	4	*80.0*	5	100
I	RD=(3) while SW=(2)	2	50.0	1	*25.0*	1	25.0	4	100
J	Other cells	1	50.0	0	0.0	1	*50.0*	2	100
	Total	58	50.9	17	14.9	39	34.2	114	100

Table (23) shows that when the condition of RD is (1) while the condition of SW is not (2) or (3), or when the condition of SW is (1) while the condition of RD is not (2) or (3), i.e. Condition A where either RD or SW favors S-initial without conflicting with other conditions, the percentage of S-initial is higher than the overall percentage (56.3% > 50.9%). The table also shows that when the condition of RD is (2) while the condition of SW is not (1) or (3), or when the condition of SW is (2) while the condition of RD is not (2) or (3), i.e. Condition B where either RD or SW favors O-initial without conflicting other conditions, the percentage of O-initial is higher than the overall percentage (42.9% > 14.9%). When the condition of RD is (3) while the condition of SW is not (1) or (2), or when the condition of SW is (3) while the condition of RD is not (1) or (2), i.e. Condition C where either RD or SW favors X-initial without conflicting with other conditions, the percentage of X-initial is slightly higher than the overall percentage (38.9% > 34.2%). However, 56.3% for S-initial under Condition A, 42.9% for O-initial under Condition B and 38.9% for X-initial under Condition C are lower than the percentages of occurrences for S-initial, O-initial and X-initial when RD favors these orders (63.2% for S-initial, 52.9% for O-initial and 62.5% for X-initial, cf. Table (6) in Section 4.2.1.1.). This result indicates that the factor SW does not improve on RD.

Moreover, when RD favors S-initial while SW favors O-initial (Condition D) or X-initial (Condition E), the percentages of S-initial are higher than the overall percentage (70.8% and 52.2% > 50.9%). In contrast, when SW favors S-initial while RD favors O-initial (Condition F) or X-initial (Condition H), the percentages of S-initial are lower than the overall percentage (50% and 20% < 50.9%). This result suggests that RD is a stronger factor than SW for preferring S-initial order.

When RD favors O-initial while SW favors S-initial (Condition F) or X-initial (Condition G), the percentages of O-initial are higher than the overall percentage (50% and 50% > 14.9%). When SW favors O-initial while RD favors S-initial (Condition D), however, the percentage of O-initial is lower than the overall percentage (12.5% < 14.9%), whereas when SW favors O-initial while RD favors X-initial (Condition I), the percentage of X-initial is higher than the overall percentage (25% > 14.9%). This result suggests that RD favoring O-initial is stronger than SW favoring the other two orders, and that RD favoring S-initial is stronger than SW favoring O-initial while SW favoring O-initial is stronger than RD favoring X-initial.

Furthermore, when RD favors X-initial while SW favors S-initial (Condition H), the percentage of X-initial is higher than the overall (80% > 34.2%). When RD favors X-initial while SW favors O-initial (Condition I), however, the percentage of X-initial is lower than the overall (25% < 34.2%). This result suggests that RD favoring X is stronger than SW favoring S-initial but not stronger than SW favoring O-initial. When SW favors

X-initial while RD favors S-initial (Condition E), the percentage of X-initial is higher than the overall percentage (47.8% > 34.2%). When SW favors X-initial while RD favors O-initial (Condition G), however, the percentages of X-initial are lower than the overall percentage (25% < 34.2%). This result suggests that SW favoring X-initial is stranger than RD favoring S-initial but not stronger than RD favoring O-initial.

4.2.2.3. TP and SW

Tables (24) and (25) show the interaction of the seven conditions defined in Sections 4.2.1.3. and 4.2.1.4. according to the relative TP and SW of constituents.

<u>Seven Conditions for TP</u>
(1) The TP of S is higher than the TP of O and X
(2) The TP of O is higher than the TP of S and X
(3) The TP of X is higher than the TP of S and O
(4) The TP of S and O is the same, but higher than the TP of X
(5) The TP of S and X is the same, but higher than the TP of O
(6) The TP of O and X is the same, but higher than the TP of S
(7) The TP of all three is the same

<u>Seven Conditions for SW</u>
(1) The SW of S is higher than the SW of O and X
(2) The SW of O is higher than the SW of S and X
(3) The SW of X is higher than the SW of S and O
(4) The SW of S and O is the same, but higher than the SW of X
(5) The SW of S and X is the same, but higher than the SW of O
(6) The SW of O and X is the same, but higher than the SW of S
(7) The SW of all three is the same

Table (24) gives the number of tokens with S-initial, O-initial and X-initial orders, and Table (25) the percentage of occurrence of each order compared to the other two orders. In Table (25), the percentages that are higher than the overall percentage of occurrence (50.9% for S-initial, 24.9% for O-initial and 34.2% for X-initial) are indicated in italics.

Table (24) Interaction of TP and SW under the 7 conditions: the number of tokens

		Conditions of SW																					Total		
		1			2			3			4			5			6			7					
Order		S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of TP	1	2	1	1	12	5	3	13	2	12	0	0	0	1	0	2	0	0	1	1	1	1	29	9	20
	2	3	1	2	1	1	3	4	1	0	0	0	0	2	0	1	0	0	0	1	1	0	11	4	6
	3	1	1	1	0	0	0	0	0	2	0	1	0	0	0	1	0	0	0	1	0	1	2	2	5
	4	0	0	0	2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1
	5	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	2	0	0
	6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	7	2	0	4	3	0	2	4	1	0	0	0	0	1	0	0	0	0	1	1	0	0	11	1	7
Total		8	3	8	20	7	8	21	4	15	0	1	0	5	0	4	0	0	2	4	2	2	58	17	39

Table (25) Interaction of RD and TP under the 7 conditions in percentages

		Conditions of SW																				
		1			2			3			4			5			6			7		
Order		S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X	S	O	X
Conditions of TP	1	50	*25*	25	*60*	*25*	15	48	7	*44*	0	0	0	33	0	*67*	0	0	*100*	33	33	33
	2	50	*17*	33	20	*20*	*60*	*80*	*20*	0	0	0	0	*67*	0	33	0	0	0	50	*50*	0
	3	33	*33*	33	0	0	0	0	0	*100*	0	*100*	0	0	0	*100*	0	0	0	50	0	*50*
	4	0	0	0	*67*	*33*	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	*100*	0	0	0	0	0	0	0	0	*100*	0	0	0	0	0	0	0	0
	6	0	0	0	*100*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	33	0	*67*	*60*	0	*40*	*80*	*20*	0	0	0	0	*100*	0	0	0	0	*100*	*100*	0	0

Table (25) shows that when the conditions of TP and SW are both (2), i.e. both TP and SW favor O-initial, the percentage of O-initial is higher than the overall percentage (20% > 14.9%). The table also shows that when the conditions of TP and SW are both (3), i.e. both TP and SW favor X-initial, the percentage of X-initial is greater than the overall percentage (100% > 34.2%). When both TP and SW favor S-initial, however, the percentage of S-initial is lower than the overall percentage (50% < 50.9%).

Table (26) provides the number of tokens and the percentage of occurrence of S-initial, O-initial and X-initial orders under certain conditions defined according to the interaction of conditions for TP and SW. The numbers were taken from Table (24). The percentages that are higher than the overall percentage (i.e. 50.9% for S-initial, 14.9% for O-initial and 34.2% for X-initial) are indicated in italics.

Table (26) The occurrence of S-initial, O-initial and X-initial clauses under the conditions based on the interaction of the seven conditions for TP and SW

Condition	TP & SW	S-initial		O-initial		X-initial		Total	
		#	%	#	%	#	%	#	%
A	TP=(1) while SW≠(2) or (3), or SW=(1) while TP≠(2) or (3)	6	35.3	2	11.8	9	*52.9*	17	100
B	TP=(2) while SW≠(1) or (2), or SW=(2) while TP≠(1) or (3)	11	*55.0*	3	*15.0*	6	30.0	20	100
C	TP=(3) while SW=(1) or (2), or SW=(3) while TP≠(1) or (2)	5	41.7	2	*16.7*	5	*41.7*	12	100
D	TP=(1) while SW=(2)	3	50.0	1	*16.7*	2	33.3	6	100
E	TP=(1) while SW=(3)	1	33.3	1	*33.3*	1	33.3	3	100
F	TP=(2) while SW=(1)	12	*60.0*	5	*25.0*	3	15.0	20	100
G	TP=(2) while SW=(3)	0	n/a	0	n/a	0	n/a	0	n/a
H	TP=(3) while SW=(1)	13	48.1	2	7.4	12	*44.4*	27	100
I	TP=(3) while SW=(2)	4	*80.0*	1	*20.0*	0	0.0	5	100
J	Other cells	3	*75.0*	0	0.0	1	25.0	4	100
	Total	58	50.9	17	14.9	39	34.2	114	100

The conditions where the percentage of S-initial is higher than the overall percentage are Condition F when SW favors S-initial while TP favoring O-initial (60% > 50.9%), Condition I when SW favors O-initial while TP favoring X-initial (80% > 50.9%), and Condition B when either TP or SW favors O-initial. The conditions where the percentage of O-initial is higher than the overall percentage are Condition F when TP favors O-initial while SW favoring S-initial (25% > 14.9%), Condition I when SW favors O-initial while TP favoring X-initial (20% > 14.9%), and Condition E when TP favors S-initial while SW favoring X-initial (33.3% > 14.9%). The conditions where the percentage of X-initial is higher than when it is overall are Condition C when either TP or SW favors X-initial (41.7% > 34.2%), Condition H when TP favors X-initial while SW favoring S-initial (44.4% > 34.2%), and Condition A when either TP or SW favors S-initial. These results indicate little evidence for any interaction of conditions for TP and SW as to factoring the choice of S-initial, O-initial and X-initial.

4.2.3. Algorithms

Based on the patterns between factors and the ordering choice, which we observed in Sections 4.2.1. and 4.2.2., we will now posit algorithms to predict the ordering choice of S-initial vs. O-initial vs. X-initial.

4.2.3.1. Algorithm-*wa*

In Section 4.2.1.6., we compared the percentage of occurrences of S-initial, O-initial and X-initial clauses when (1) S is marked with *wa*, (ii) O is marked with *wa*, (iii) X is marked with *wa*, (iv) S and O are marked with *wa*, (v) S and X are marked with *wa*, and (vi) O and X are marked with *wa*. Based on the result for this comparison, I will posit an algorithm using the particle *wa* as a factor.

Algorithm-*wa*
(i) If the S is marked with *wa*, use S-initial,
(ii) If the O is marked with *wa*, use O-initial,
(iii) If the X is marked with *wa*, use X-initial,
(iv) If the S and O are marked with *wa*, use S-initial,
(v) If the S and X are marked with *wa*, use S-initial, or
(vi) If the O and X are marked with *wa*, use X-initial.

The fact that S-initial is the default order is taken into account for positing (iv) and (v) in Algorithm-*wa* since there was no significant difference in token counts when both S and O are marked with *wa* and when both S and X are marked with *wa* (cf. Table (11) in Section 4.1.2.6.).

Let us apply Algorithm-*wa* to the 42 tokens that contain *wa*-marked constituents. The result of predictions for Algorithm-*wa* is shown in Table (27) and summarized in (28).

Table (27) Predictions for Algorithm-wa

Algorithm-wa	Correct Predictions		Incorrect Predictions				Total
S is marked with *wa*	S-initial	18	O-initlal	4	X-initial	4	26
O is marked with *wa*	O-initial	3	S-initial	3	X-initial	0	6
X is marked with *wa*	X-initial	5	S-initial	1	O-initial	0	6
S and O are marked with wa	S-initial	1	O-initial	1	X-initial	0	2
S and X are marked with wa	S-initial	0	O-initial	0	X-initial	1	1
O and X are marked with wa	X-initial	1	S-initial	0	O-initial	0	1
Total	28		14				42
Percent	66.7%		33.3%				100%

Table (28) Summary of predictions for Algorithm-wa

Algorithm-wa	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	19	82.6	3	37.5	6	54.5	28	66.7
Incorrect predictions	4	17.4	5	62.5	5	45.5	14	33.3
Total	23	100	8	100	11	100	42	100

Table (28) shows that Algorithm-*wa* yielded a 66.7% success rate. The percentage of correct predictions indicates that the algorithm is more relevant to predicting S-initial and X-initial orders than O-initial.

4.2.3.2. Algorithm-demonstrative

In Section 4.2.1.7., we examined the percentage of occurrences under the conditions when (i) S is demonstrative, (ii) O is demonstrative, (iii) X is demonstrative, (iv) S and O are demonstrative, (v) S and X are demonstrative, and (vi) O and X are demonstrative. Based on the comparative percentage of occurrences we obtained, I will try Algorithm-demonstrative, using demonstrative as a factor.

Algorithm-demonstrative
(i) If the S is demonstrative, use S-initial,
(ii) If the O is demonstrative, use O-initial,
(iii) If the X is demonstrative, use X-initial,
(iv) If the S and O are demonstrative, use S-initial,
(v) If the S and X are demonstrative, use S-initial, or
(vi) If the O and X are demonstrative, use X-initial.

In positing (iv) and (v) in Algorithm-demonstrative, S-initial being the default order was taken into account since token counts under the condition when both S and O are demonstrative and when both S and X are demonstrative did not show a significant difference (cf. Table (12) in Section 4.2.1.7.). Algorithm-demonstrative applies a set of data that contains demonstrative constituents. Tables (29) and (30) show the result of predictions for Algorithm-demonstrative when it is applied to the 50 tokens containing demonstrative constituents.

Table (29) Predictions for Algorithm-Demonstrative

A-Demonstrative	Correct Predictions		Incorrect Predictions				Total
S=demo	S-initial	8	O-initlal	1	X-initial	0	9
O=demo	O-initial	9	S-initial	2	X-initial	0	11
X=demo	X-initial	13	S-initial	8	O-initial	0	21
S and O = demo	S-initial	0	S-initial	0	X-initial	0	0
S and X = demo	S-initial	1	O-initial	0	X-initial	1	2
O and X = demo	X-initial	3	S-initial	2	O-initial	2	7
Total	34		16				50
Percent	68.0%		32.0%				100%

Table (30) Summary of predictions for Algorithm-Demonstrative

Algorithm-Demonstrative	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	9	42.9	9	75.0	16	94.1	34	68.0
Incorrect predictions	12	57.1	3	25.0	1	5.9	16	32.0
Total	21	100	12	100	17	100	50	100

Table (30) shows a 68% success rate for Algorithm-demonstrative. The correct prediction rate indicates that the algorithm is relevant to predicting X-initial and O-initial orders, but not S-initial. This result conforms to the result in Chapter 3 that demonstrative subject is not an influential factor upon the ordering choice between SOV and OSV orders.

4.2.3.3. Algorithm-locative

In Section 4.2.1.9., we discussed that at-locatives are more likely to occur in the initial position than to/from locatives. In this section, we will try Algorithm-locative, based on the semantic type of locatives.

Algorithm-locative: If there is an at-locative, then use X-initial.

Table (31) shows the number of tokens containing at-locatives in S-initial, O-initial and X-initial clauses and the percentages of occurrence for each order.

Table (31) The number of tokens containing at-locatives in the ≥3NPs construction

Type of locative	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
At-locatives	9	37.5	3	12.5	12	50.0	24	100

When we apply Algorithm-locative to the 24 tokens, it yields 50% of success rate, which is slightly lower than the overall percentage of occurrence for S-initial (i.e. 50.9%), but higher than the overall occurrence of X-initial (i.e. 34.2%).

4.2.3.4. Algorithm-temporal

In Section 4.2.1.10., the data indicated that time-temporals are more likely to appear in the initial position than duration- or frequency-temporals. Let us posit an algorithm based on the semantic type of temporals.

Algorithm-temporal: If there is a time expression, then use X-initial.

Table (32) shows the number of tokens that contain time expressions and the percentage of occurrence in S-initial, O-initial and X-initial clauses.

Table (32) The number of tokens containing time expressions in the ≥3NPs construction

Type of temporal	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Time-temporal	18	48.6	3	8.1	16	43.2	37	100

When we apply Algorithm-temporal to the 37 tokens containing time expressions, the success rate of predictions is 43.2%, which is the percentage of X-initial. This success rate is lower than the overall percentage of occurrence of S-initial (i.e. 50.9%), but higher than the overall occurrence of X-initial (i.e. 34.2%).

4.2.3.5. Algorithm-RD

The data in Sections 4.2.1.1. shows that RD is a relevant factor to favoring S-initial, O-initial and X-initial orders. In this section, I will try to formulate an algorithm using RD as a factor.

Algorithm-RD
(a) If the RD of one of the S, O and X is lower than the RD of the other two, then the constituent with the lowest RD comes first.
Or
(b) If the RD of the O and X is the same, but lower than the RD of S, then use O-initial.
Otherwise,
(c) Use S-initial.

Algorithm-RD (b) was formulated so that O-initial wins over S-initial because under the 6th condition in Table (6) in Section 4.2.1.1., (RD of O = RD of X) < RD of S, the percentage of O-initial is higher than the overall percentage of O-initial (i.e. 14.9%) while the percentage of X-initial is not. Algorithm-RD (c) takes into account the fact that S-initial order is the default order and that there is not significant different in token counts under the 4th and 5th conditions in Table (6). Applying Algorithm-RD to the 114 tokens in the data yields 58.8% of success rate of predictions. The result is shown in Tables (33) and (34).

Table (33) Breakdown of predictions for Algorithm-RD

	Algorithm-RD	(i)		(a)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	36	65.5	0	0.0	11	100.0	47	70.1
	O-initial	9	16.4	1	100.0	0	0.0	10	14.9
	X-initial	10	18.2	0	0.0	0	0.0	10	14.9
	Total	55	100	1	100	11	100	67	100
Incorrect prediction	S-initial	11	31.4	0	0.0	n/a	n/a	11	23.4
	O-initial	4	11.4	n/a	n/a	3	25.0	7	14.9
	X-initial	20	57.1	0	0.0	9	75.0	29	61.7
	Total	35	100	0	0	12	100	47	100

Table (34) Predictions for Algorithm-RD

Algorithm-RD	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	47	81.0	10	58.8	10	25.6	67	58.8
Incorrect predictions	11	19.0	7	41.2	29	74.4	47	41.2
Total	58	100	17	100	39	100	114	100

Tables (33) and (34) indicate that Algorithm-RD yields a better result for predicting S-initial and O-initial than X-initial. The correct prediction rate for O-initial is no more than the overall occurrence of O-initial (14.9%), and that for X-initial is lower than the overall occurrence of X-initial (i.e. 34.2%).

4.2.3.6. Algorithm-ERD

The study in Section 4.2.1.2. indicated a relation between a lower ERD and the initial position. It also showed that ERD is particularly relevant to predicting O-initial. In this section, I will try to formulate an algorithm using ERD as a factor.

Algorithm-ERD
(a) If the ERD of one of the S, O and X is lower than the ERD of the other two, then the constituent with the lowest ERD comes first.
Or
(b) If the ERD of the O and X is the same, but lower than the ERD of S, then use O-initial.
Otherwise,
(c) Use S-initial.

Algorithm-ERD (a) was formulated based on the relation between the initial position and a lower ERD. Algorithm-ERD (b) is based on the result from Table (7) in Section 4.2.1.2. that under the 6th condition, (ERD of O = ERD of X) < ERD of S, O-initial is only order which showed a higher occurrence rate than the overall occurrence (25% vs. 14.9%). Algorithm-ERD (c) takes into account S-order being the default order. Applying Algorithm-ERD to the 114 tokens in the data produces the result shown in Tables (35) and (36).

Table (35) Breakdown of predictions for Algorithm-ERD

	Algorithm-ERD	(a)		(b)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	30	58.8	0	0.0	11	100.0	41	65.1
	O-initial	12	23.5	1	100.0	0	0.0	13	20.6
	X-initial	9	17.6	0	0.0	0	0.0	9	14.3
	Total	51	100	1	100	11	100	63	100
Incorrect prediction	S-initial	15	40.5	2	0.0	n/a	n/a	17	33.3
	O-initial	2	5.4	n/a	n/a	2	18.2	4	7.8
	X-initial	20	54.1	1	0.0	9	81.8	30	58.8
	Total	37	100	3	0	11	100	51	100

Table (36) Predictions for Algorithm-ERD

Algorithm-RD	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	41	70.7	13	76.5	9	23.1	63	55.3
Incorrect predictions	17	29.3	4	23.5	30	76.9	51	44.7
Total	58	100	17	100	39	100	114	100

Table (36) shows that the overall correct prediction rate for Algorithm-ERD is lower than that for Algorithm-RD (55.3 % vs. 58.5 %). While Algorithm-RD does not predict O-initial no more than the overall occurrence of O-initial (cf. Table (33) in Section 4.2.3.5.), Algorithm-ERD predicts more. (14.9% for Algorithm-RD vs. 20.6% for Algorithm-ERD). On the other hand, Algorithm-RD predicts S-initial better than Algorithm-ERD (70.1% for Algorithm-RD vs. 65.1% for Algorithm-ERD).

4.2.3.7. Algorithm-TP

The study in Section 4.2.1.3. indicates that the factor TP is little relevant to preferring S-initial and O-initial orders, but it might relevant to choosing X-initial. I will try to formulate an algorithm here, using TP as a factor.

Algorithm-TP
(a) If the TP of one of the S, O and X is higher than the TP of the other two, then the constituent with the highest TP comes first.
Otherwise,
(b) Use S-initial

Algorithm-TP was formulated based on a relation between a higher TP and the initial position, which is indicated in Table (8) in Section 4.2.1.3. The fact that S-initial is the default order was taken into account as well. Applying Algorithm-TP to the 114 tokens in the data yields the result of predictions shown in Tables (37) and (38).

Table (37) Breakdown of predictions for Algorithm-TP

	Algorithm-TP	(a)		(b)		Total	
		#	%	#	%	#	%
Correct prediction	S-initial	29	76.3	16	100.0	45	83.3
	O-initial	4	10.5	0	0.0	4	7.4
	X-initial	5	13.2	0	0.0	5	9.3
	Total	38	100	16	100	54	100
Incorrect prediction	S-initial	13	26.5	n/a	0.0	13	21.7
	O-initial	10	20.4	3	n/a	13	21.7
	X-initial	26	53.1	8	0.0	34	56.7
	Total	49	100	11	0	60	100

Table (38) Predictions for Algorithm-TP

Algorithm-TP	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	45	77.6	4	23.5	5	12.8	54	47.4
Incorrect predictions	13	22.4	13	76.5	34	87.2	60	52.6
Total	58	100	17	100	39	100	114	100

The success rate of predictions for Algorithm-TP, 47.4% is lower than the overall percentage of S-initial occurrence, i.e. 50.9%. While Algorithm-TP does not correctly predict O-initial and X-initial, the correct prediction rate for S-initial is noticeably higher than predicting the other two orders.

4.2.3.8. Algorithm-SW

The study in Section 4.2.1.4. indicates that the factor SW is little relevant to predicting S-initial, O-initial and X-initial ordering. The present data also shows that 42.1% of objects in O-initial tokens (7/17 tokens) and 38.5% of Xs in X-initial tokens (15/39 tokens) have a higher SW than other phrases in the same clause, while only 13.8% of subjects in S-initial (8/58 tokens) have a higher SW than the object and X in the same clause. In this section, I will try to formulate an algorithm using SW as a factor.

Algorithm-SW
(a) If the SW of one of the S, O and X is higher than the SW of the other two, then the constituent with the highest TP comes first.
Or
(b) If the SW of the O and X is the same, but higher than the SW of S, then use X-initial.
Otherwise,
(c) Use S-initial.

Algorithm-SW (c) was formulated based on the fact that S-initial is the default order. Applying Algorithm-SW to the 114 tokens in the data yields the result of predictions shown in Tables (39) and (40).

Table (39) Breakdown of predictions for Algorithm-SW

	Algorithm-SW	(a)		(b)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	8	26.7	0	0.0	9	100.0	17	41.5
	O-initial	7	23.3	0	0.0	0	0.0	7	17.1
	X-initial	15	50.0	2	100.0	0	0.0	17	41.5
	Total	30	100	2	100	9	100	41	100
Incorrect prediction	S-initial	41	64.1	0	0.0	n/a	n/a	41	56.2
	O-initial	7	10.9	0	n/a	3	33.3	10	13.7
	X-initial	16	25.0	0	0.0	6	66.7	22	30.1
	Total	64	100	0	0	9	100	73	100

Table (40) Predictions for Algorithm-SW

Algorithm-SW	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	17	29.3	7	41.2	17	43.6	41	36.0
Incorrect predictions	41	70.7	10	58.8	22	56.4	73	64.0
Total	58	100	17	100	39	100	114	100

The overall success rate for Algorithm-SW is 36%, which is lower than the overall occurrence of S-initial (50.9%). While the correct prediction rate for S-initial is lower than its overall occurrence (41.5% vs. 50.9%), the correct prediction rates for O-initial and for X-initial are slightly higher than their overall occurrences (17.1% vs. 14.9% for O-initial, and 41.5% vs. 34.2% for X-initial). Moreover, Algorithm-SW predicts O-initial and X-initial better than Algorithm-TP. For example, Algorithm-SW correctly predicts O-initial for 41.2% of O-initial tokens, while Algorithm-TP predicts O-initial for 23.5% of O-initial tokens. Likewise, Algorithm-SW predicts X-initial order for 43.6% of X-initial tokens while the prediction rate of Algorithm-TP for X-initial is 12.8%.

4.2.4. Building algorithms

In this section, I will seek an algorithm to predict the ordering choice of S-initial, O-initial and X-initial with the best prediction rate by combining Algorithms based on a single factor that were formulated in Section 4.2.3.

4.2.4.1. Algorithm-demonstrative and Algorithm-*wa*

In Chapter 3, we saw a tendency that applying an algorithm that is applicable to a larger set of data before applying an algorithm that is applicable to a smaller set of data yields a better result of predictions than applying the algorithms in a reversed order. In the present data, there are 50 tokens containing demonstrative constituents and 42 tokens containing *wa*-marked constituents. Therefore, Algorithm (30) is laid out in a way that Algorithm-demonstrative outranks Algorithm-*wa*.

Algorithm (30)
(a) If there are demonstrative constituents, and
- (i) If the S is demonstrative, use S-initial,
- (ii) If the O is demonstrative, use O-initial,
- (iii) If the X is demonstrative, use X-initial,
- (iv) If the S and O are demonstrative, use S-initial,
- (v) If the S and X are demonstrative, use S-initial, or
- (vi) If the O and X are demonstrative, use X-initial.

Otherwise,
(b) If there are *wa*-marked constituents, and
- (i) If the S is marked with *wa*, use S-initial,
- (ii) If the O is marked with *wa*, use O-initial,
- (iii) If the X is marked with *wa*, use X-initial,
- (iv) If the S and X are marked with *wa*, use S-initial,
- (v) If the S and O are marked with *wa*, use S-initial, or
- (vi) If the O and X are marked with *wa*, use X-initial.

Now we will try to apply Algorithm (30b) to the 64 tokens that do not contain demonstrative constituents (i.e. 114 tokens minus 50 tokens). Among the 64 tokens, there are 30 tokens that contain *wa*-marked constituents (i.e. 42 tokens containing *wa*-marked constituents minus 12 tokens (3 S-initial, 5 O-initial and 4 X-initial) containing constituents that are demonstrative and marked with *wa* at the same time. Therefore, Algorithm (30b) only applies to the 30 tokens. Table (41) shows the number of *wa*-marked constituents in the 64 tokens.

Table (41) The number of tokens containing *wa* -marked constituents in the 64 tokens that do not contain demonstrative constituents

Wa -marking	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is marked with *wa*	16	80.0	1	5.0	3	15.0	20	100
O is marked with *wa*	3	60.0	2	40.0	0	0.0	5	100
X is marked with *wa*	0	0.0	0	0.0	2	0.0	2	n/a
S and O are marked with *wa*	1	100.0	0	0.0	0	0.0	1	100
S and X are marked with *wa*	0	0.0	0	0.0	1	0	1	n/a
O and X are marked with *wa*	0	0.0	0	0.0	1	0	1	n/a
None of S, O or X is marked with *wa*	17	50.0	2	5.9	15	44.1	34	100
Total	37	57.8	5	7.8	22	34.4	64	100

Algorithm (30b) is applied to the 30 tokens that contain *wa*-marked constituents, and it correctly predicts the ordering for 22 tokens. The result of predictions for Algorithm (30b) is shown in Tables (42) and (43).

Table (42) Predictions for Algorithm (30b)

Algorithm (30b)	Correct Predictions		Incorrect Predictions				Total
S is marked with *wa*	S-initial	16	O-initlal	1	X-initial	3	20
O is marked with *wa*	O-initial	2	S-initial	3	X-initial	0	5
X is marked with *wa*	X-initial	2	S-initial	0	O-initial	0	2
S and O are marked with wa	S-initial	1	O-initial	0	X-initial	0	1
S and X are marked with wa	S-initial	0	O-initial	0	X-initial	1	1
O and X are marked with wa	X-initial	1	S-initial	0	O-initial	0	1
Total	22		8				30
Percent	73.3%		26.7%				100%

When we combine the results of predictions in Table (30) and (42), we can obtain the result of predictions for Algorithm (30a-b). The result is shown in Table (43).

Table (43) Predictions for Algorithm (30a-b)

Algorithm (30a-b)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	26	63.4	11	73.3	19	79.2	56	70.0
Incorrect predictions	15	36.6	4	26.7	5	20.8	24	30.0
Total	41	100	15	100	24	100	80	100
# of remaining tokens	17	50.0	2	5.9	15	44.1	34	100

Table (43) shows that Algorithm (30a-b) was applied to the 80 tokens that contain *wa*-marked or demonstrative constituents, and yielded a 70% success rate. We still have 34 tokens that none of the algorithms have applied to. I will apply Algorithm-RD to the remaining 34 tokens. Algorithm-RD is integrated with Algorithm (30a,b) here as subalgorithm (c).

Algorithm (30)
(a) If there are demonstrative constituents, and
  (i) If the S is demonstrative, use S-initial,
  (ii) If the O is demonstrative, use O-initial,
  (iii) If the X is demonstrative, use X-initial,
  (iv) If the S and O are demonstrative, use S-initial,
  (v) If the S and X are demonstrative, use S-initial, or
  (vi) If the O and X are demonstrative, use X-initial.
Otherwise,
(b) If there are *wa*-marked constituents, and
  (i) If the S is marked with *wa*, use S-initial,
  (ii) If the O is marked with *wa*, use O-initial,
  (iii) If the X is marked with *wa*, use X-initial,
  (iv) If the S and X are marked with *wa*, use S-initial,
  (v) If the S and O are marked with *wa*, use S-initial, or
  (vi) If the O and X are marked with *wa*, use X-initial.
Otherwise,
(c) (i) If the RD of one of the S, O and X is lower than the RD of the other two, then the constituent with the lowest RD comes first.
  Or
  (ii) If the RD of the O and X is the same, but lower than the RD of S, then use O-initial.
  Otherwise,
  (iii) Use S-initial.

Table (44) shows the result of predictions for Algorithm (30c), when applied to the 34 tokens.

Table (44) Predictions for Algorithm (30c)

Algorithm (30c)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	13	76.5	1	50.0	2	13.3	16	47.1
Incorrect predictions	4	23.5	1	50.0	13	86.7	18	52.9
Total	17	100	2	100	15	100	34	100

Let us combine the results in Tables (43) and (44) to yield the result of predictions for Algorithm (30). Tables (45) and (46) show the result.

Table (45) Breakdown of predictions for Algorithm (30)

	Algorithm (30)	(a)		(b)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	9	26.5	17	77.3	13	81.3	39	54.2
	O-initial	9	26.5	2	9.1	1	6.3	12	16.7
	X-initial	16	47.1	3	13.6	2	12.5	21	29.2
	Total	34	100	22	100	16	100	72	100
Incorrect prediction	S-initial	12	75.0	3	37.5	4	22.2	19	45.2
	O-initial	3	18.8	1	12.5	1	5.6	5	11.9
	X-initial	1	6.3	4	50.0	13	72.2	18	42.9
	Total	16	100	8	100	18	100	42	100

Table (46) Predictions for Algorithm (30)

Algorithm (30)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	39	67.2	12	70.6	21	53.8	72	63.2
Incorrect predictions	19	32.8	5	29.4	18	46.2	42	36.8
Total	58	100	17	100	39	100	114	100

The overall success rate for Algorithm (30) is 63.2%. Table (45) shows that while the correct prediction rate for S-initial and O-initial is slightly higher than their overall occurrences (54.2% vs. 50.9% for S-initial and 16.7% vs. 14.9% for O-initial), the correct prediction rate for X-initial is lower than its overall occurrence (29.2% vs. 34.2%).

4.2.4.2. Algorithm-*wa* and Algorithm-demonstrative

Now, I will try Algorithm (31) that is also a combination of Algorithm-*wa* Algorithm-demonstrative, and Algorithm-RD, but applies Algorithm-*wa* before Algorithm-demonstrative.

Algorithm (31)

(a) If there are *wa*-marked constituents, and
- (i) If the S is marked with *wa*, use S-initial,
- (ii) If the O is marked with *wa*, use O-initial,
- (iii) If the X is marked with *wa*, use X-initial,
- (iv) If the S and X are marked with *wa*, use S-initial,
- (v) If the S and O are marked with *wa*, use S-initial, or
- (vi) If the O and X are marked with *wa*, use X-initial.

Otherwise,
(b) If there are demonstrative constituents, and
- (i) If the S is demonstrative, use S-initial,
- (ii) If the O is demonstrative, use O-initial,
- (iii) If the X is demonstrative, use X-initial,
- (iv) If the S and O are demonstrative, use S-initial,
- (v) If the S and X are demonstrative, use S-initial, or
- (vi) If the O and X are demonstrative, use X-initial.

Otherwise,
(c) (i) If the RD of one of the S, O and X is lower than the RD of the other two, then the constituent with the lowest RD comes first.
Or
(ii) If the RD of the O and X is the same, but lower than the RD of S, then use O-initial.
Otherwise,
(iii) Use S-initial.

Among the 114 tokens in the data, there are 42 tokens that contain *wa*-marked constituents. Algorithm-*wa*, i.e. Algorithm (31a), was applied to the 42 tokens, and correctly predicted the ordering for 28 tokens (i.e. 66.7% success rate). Among the 72 tokens that do not contain *wa*-marked constituents, there are 38 tokens (18 S-initial, 7 O-initial and 13 X-initial) containing demonstrative constituents. Table (47) shows the position and number of demonstrative constituents in the 38 tokens.

Table (47) The number of tokens containing demonstrative constituents in the 72 tokens that do not contain wa-marked constituents

Demonstrative	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is demonstrative	7	87.5	1	12.5	0	0.0	8	100
O is demonstrative	2	25.0	6	75.0	0	0.0	8	100
X is demonstrative	6	35.3	0	0.0	11	64.7	17	100
S and O are demonstrative	0	n/a	0	n/a	0	n/a	0	n/a
S and X are demonstrative	1	100.0	0	0.0	0	0.0	1	100
O and X are demonstrative	2	50.0	0	0.0	2	50.0	4	100
None of constituents is demonstrative	17	50.0	2	5.9	15	44.1	34	100
Total	35	48.6	9	12.5	28	38.9	72	100

Tables (48) and (49) show the result of predictions for Algorithm (31b) when applied to the 38 tokens in the data.

Table (48) Predictions for Algorithm (31b)

Demonstrative	Correct Predictions		Incorrect Predictions				Total
S=demo	S-initial	7	O-initltal	1	X-initial	0	8
O=demo	O-initial	6	S-initial	2	X-initial	0	8
X=demo	X-initial	11	S-initial	6	O-initial	0	17
S and O = demo	S-initial	0	S-initial	0	X-initial	0	0
S and X = demo	S-initial	1	O-initial	0	X-initial	0	1
O and X = demo	X-initial	2	S-initial	2	O-initial	0	4
Total	27		11				38
Percent	71.1%		28.9%				100%

Table (49) Predictions for Algorithm (31b)

Algorithm (31b)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	8	44.4	6	85.7	13	100.0	27	71.1
Incorrect predictions	10	55.6	1	14.3	0	0.0	11	28.9
Total	18	100	7	100	13	100	38	100

When we combine the results in Table (28) and in Table (49), we can obtain the result of predictions for Algorithm (31a,b). The result is shown in Table (50).

Table (53) Predictions for Algorithm (31a.b)

Algorithm (31a,b)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	27	65.9	9	60.0	19	79.2	55	68.8
Incorrect predictions	14	34.1	6	40.0	5	20.8	25	31.3
Total	41	100	15	100	24	100	80	100
# of remaining tokens	17	50.0	2	5.9	15	44.1	34	100

Let us apply now Algorithm (31c) to the remaining 34 tokens that contain neither *wa*-marked constituents nor demonstratives. The result of predictions for Algorithm (31c) is the same as the result for Algorithm (30c), which is shown in Table (44) in Section 4.2.4.1. Thus, we can obtain the result of predictions for Algorithm (31) by combining the results from Tables (44) and (53). Tables (54) and (55) show the result.

Table (54) Breakdown of predictions for Algorithm (31)

	Algorithm (31)	(a)		(b)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	19	67.9	8	29.6	13	81.3	40	56.3
	O-initial	3	10.7	6	22.2	1	6.3	10	14.1
	X-initial	6	21.4	13	48.1	2	12.5	21	29.6
	Total	28	100	27	100	16	100	71	100
Incorrect prediction	S-initial	4	28.6	10	90.9	4	22.2	18	41.9
	O-initial	5	35.7	1	9.1	1	5.6	7	16.3
	X-initial	5	35.7	0	0.0	13	72.2	18	41.9
	Total	14	100	11	100	18	100	43	100

Table (55) Predictions for Algorithm (31)

Algorithm (31)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	40	69.0	10	58.8	21	53.8	71	62.3
Incorrect predictions	18	31.0	7	41.2	18	46.2	43	37.7
Total	58	100	17	100	39	100	114	100

The overall success rate for Algorithm (31) is 62.3%, which is higher than Algorithm (30) by one token. The correct prediction rates for S-initial, O-initial and X-initial are comparable between Algorithms (30) and (31) (cf. Table (45) in Section 4.2.4.1.).

4.2.4.3. Algorithms-locative, -temporal and –topic (Algorithm-X)

In this section, I will try an algorithm based on the combined factors of locatives, time expressions, and topical elements.

Algorithm-X:
If there is an at-locative, time expression or topical element, then use X-initial.

Table (56) shows the number of tokens containing at-locatives, time expressions and topical elements.

Table (56) The number of constituents containing at-locatives, time expressions and topical elements

containing	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
At-locative	7	36.8	3	15.8	9	47.4	19	100
Time-temporal	18	50.0	3	8.3	15	41.7	36	100
Topical	0	0.0	0	0.0	5	100.0	5	100
Time-temporal & at-locative	0	0.0	0	0.0	1	100.0	1	100
Two at-locatives	1	0.0	0	0.0	1	100.0	2	100
Total	26	41.3	6	9.5	31	49.2	63	100
None of above	32	62.7	11	21.6	8	15.7	51	100

The result of predictions for Algorithm-X when it is applied to 63 tokens containing at-locatives, time expressions and topic elements is shown Table (57).

Table (57) Predictions for Algorithm-X

Algorithm-X	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	0	0.0	0	0.0	31	100.0	31	49.2
Incorrect predictions	26	100.0	6	100.0	0	0.0	32	50.8
Total	26	100	6	100	31	100	63	100

Table (57) shows that while Algorithm-X does not predict S-initial or O-initial orders. the algorithm predicts 31 tokens of X-initial orders, which is 79.5% of X-initial tokens (i.e. 31/39 tokens).

4.2.4.4. Algorithms-X, -demonstrative and -*wa*

In this section, I will try to formulate an algorithm based on Algorithms-demonstrative, -*wa*, -locative, -temporal and -topic. The algorithm I will try is Algorithm (32) wherein Algorithm-X is applied before Algorithm-demonstrative, followed by Algorithm-*wa*.

Algorithm (32)
(a) If there is an at-locative, time expression or topical element, then use X-initial.
Otherwise,
(b) If there are *wa*-marked constituents, and
 (i) If the S is marked with *wa*, use S-initial,
 (ii) If the O is marked with *wa*, use O-initial,
 (iii) If the X is marked with *wa*, use X-initial,
 (iv) If the S and X are marked with *wa*, use S-initial,
 (v) If the S and O are marked with *wa*, use S-initial, or
 (vi) If the O and X are marked with *wa*, use X-initial.
Otherwise,
(c) If there are demonstrative constituents, and
 (i) If the S is demonstrative, use S-initial,
 (ii) If the O is demonstrative, use O-initial,
 (iii) If the X is demonstrative, use X-initial,
 (iv) If the S and O are demonstrative, use S-initial,
 (v) If the S and X are demonstrative, use S-initial, or
 (vi) If the O and X are demonstrative, use X-initial.

Algorithm (32a) was applied to the 63 tokens containing at-locatives, temporal expressions and topical elements, and yielded a 49.2% success rate (cf. Section 4.2.4.3.). Table (58) shows the number of tokens containing demonstrative constituents in the 51 tokens that contain neither at-locatives, time expressions nor topical elements.

Table (58) The number of tokens containing demonstrative constituents in the 51 tokens that contain neither at-locatives, time expressions nor topical elements

Demonstrative	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is demonstrative	4	80.0	1	20.0	0	0.0	5	100
O is demonstrative	0	0.0	6	100.0	0	0.0	6	100
X is demonstrative	1	20.0	0	0.0	4	80.0	5	100
S and O are demonstrative	0	n/a	0	n/a	0	n/a	0	n/a
S and X are demonstrative	1	100.0	0	0.0	0	0.0	1	100
O and X are demonstrative	2	66.7	1	33.3	0	0.0	3	100
None of constituents is demonstrative	24	77.4	3	9.7	4	12.9	31	100
Total	32	62.7	11	21.6	8	15.7	51	100

Table (58) shows that in the 51 tokens, there are 20 tokens containing demonstrative constituents. When Algorithm (32b) is applied to the 20 tokens containing demonstrative constituents, it yields 15 correct predictions (five S-initial, six O-initial and four X-initial), and 5 incorrect predictions (three S-initial and two O-initial). Tables (59) and (60) show the result of predictions for Algorithm (32a-b).

Table (59) Breakdown of predictions for Algorithm (32a-b)

	Algorithm (32a-b)	(a)		(b)		Total	
		#	%	#	%	#	%
Correct prediction	S-initial	0	0.0	5	33.3	5	10.9
	O-initial	0	0.0	6	40.0	6	13.0
	X-initial	31	100.0	4	26.7	35	76.1
	Total	31	100	15	100	46	100
Incorrect prediction	S-initial	26	81.3	3	60.0	29	78.4
	O-initial	6	18.8	2	40.0	8	21.6
	X-initial	0	0.0	0	0.0	0	0.0
	Total	32	100	5	100	37	100

Table (60) Predictions for Algorithm (32a-b)

Algorithm (32a-b)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	5	14.7	6	42.9	35	100.0	46	55.4
Incorrect predictions	29	85.3	8	57.1	0	0.0	37	44.6
Total	34	100	14	100	35	100	83	100
# of remaining tokens in the data	24	77.4	3	9.7	4	12.903	31	100

Now, we will apply Algorithm (32c) to the remaining 31 tokens in the data. Table (61) shows the number of *wa*-marked constituents in the 31 tokens.

Table (61) The number of tokens containing *wa* -marked constituents in the 31 tokens that contain neither demonstrative constituents, at-locataives, time expressions nor topical elements

Wa -marking	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
S is marked with *wa*	10	90.9	1	9.1	0	0.0	11	100
O is marked with *wa*	3	75.0	1	25.0	0	0.0	4	100
X is marked with *wa*	0	0.0	0	0.0	0	0.0	0	n/a
S and O are marked with *wa*	1	100.0	0	0.0	0	0.0	1	100
S and X are marked with *wa*	0	0.0	0	0.0	0	0	0	n/a
O and X are marked with *wa*	0	0.0	0	0.0	0	0	0	n/a
None of S, O or X is marked with *wa*	10	66.7	1	6.7	4	26.7	15	100
Total	24	77.4	3	9.7	4	12.9	31	100

Table (61) indicates that Algorithm (32c) will correctly predict for 12 tokens (eleven S-initial and one O-initial), and incorrectly predicts 4 tokens (three S-initial and one O-initial). Tables (62) and (63) shows the result of predictions for Algorithm (32a-c).

Table (62) Breakdown of predictions for Algorithm (32a-c)

	Algorithm (30a-c)	(a)		(b)		(c)		Total	
		#	%	#	%	#	%	#	%
Correct prediction	S-initial	0	0.0	5	33.3	11	91.7	16	27.6
	O-initial	0	0.0	6	40.0	1	8.3	7	12.1
	X-initial	31	100.0	4	26.7	0	0.0	35	60.3
	Total	31	100	15	100	12	100	58	100
Incorrect prediction	S-initial	26	81.3	3	60.0	3	75.0	32	78.0
	O-initial	6	18.8	2	40.0	1	25.0	9	22.0
	X-initial	0	0.0	0	0.0	0	0.0	0	0.0
	Total	32	100	5	100	4	100	41	100

Table (63) Predictions for Algorithm (32a-c)

Algorithm (32a-c)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	16	33.3	7	43.8	35	100.0	58	58.6
Incorrect predictions	32	66.7	9	56.3	0	0.0	41	41.4
Total	48	100	16	100	35	100	99	100
# of remaining tokens in the data	10	66.7	1	6.7	4	26.7	15	100

Examining the remaining 15 tokens reveals that the factor RD is an effective predictor for constituent ordering for the 15 tokens. Based on this observation, I will add another subalgorithm (d) to Algorithm (32).

Algorithm (32')
(a) If there is an at-locative, time expression or topical element, then use X-initial.
Otherwise,
(b) If there are *wa*-marked constituents, and
  (i) If the S is marked with *wa*, use S-initial,
  (ii) If the O is marked with *wa*, use O-initial,
  (iii) If the X is marked with *wa*, use X-initial,
  (iv) If the S and X are marked with *wa*, use S-initial,
  (v) If the S and O are marked with *wa*, use S-initial, or
  (vi) If the O and X are marked with *wa*, use X-initial.
Otherwise,
(c) If there are demonstrative constituents, and
  (i) If the S is demonstrative, use S-initial,
  (ii) If the O is demonstrative, use O-initial,
  (iii) If the X is demonstrative, use X-initial,
  (iv) If the S and O are demonstrative, use S-initial,
  (v) If the S and X are demonstrative, use S-initial, or
  (vi) If the O and X are demonstrative, use X-initial.

Otherwise,
(d) (i) If the RD of one of the S, O and X is lower than the RD of the other two, then the constituent with the lowest RD comes first.
Or
(ii) If the RD of the O and X is the same, but lower than the RD of S, then use O-initial.
Otherwise,
(iii) Use S-initial.

Table (64) shows the result of prediction for Algorithm (32'd), when applied to the 15 tokens that contain neither demonstratives, *wa*-marked constituents, at-locatives, time expressions nor topical elements.

Table (64) Predictions for Algorithm (32'd)

Algorithm (32'd)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	7	70.0	1	100.0	1	25.0	9	60.0
Incorrect predictions	3	30.0	0	0.0	3	75.0	6	40.0
Total	10	100	1	100	4	100	15	100

When we combine the results in Tables (63) and (64), we can obtain the result of predictions for Algorithm (32'). Tables (65) and (66) show the result.

Table (65) Breakdown of predictions for Algorithm (32')

	Algorithm (32')	(a)		(b)		(c)		(d)		Total	
		#	%	#	%	#	%	#	%	#	%
Correct prediction	S-initial	0	0.0	5	33.3	11	91.7	7	77.8	23	34.3
	O-initial	0	0.0	6	40.0	1	8.3	1	11.1	8	11.9
	X-initial	31	100	4	26.7	0	0.0	1	11.1	36	53.7
	Total	31	100	15	100	12	100	9	100	67	100
Incorrect prediction	S-initial	26	81.3	3	60.0	3	75.0	3	50.0	35	74.5
	O-initial	6	18.8	2	40.0	1	25.0	0	0.0	9	19.1
	X-initial	0	0.0	0	0.0	0	0.0	3	50.0	3	6.4
	Total	32	100	5	100	4	100	6	100	47	100

Table (66) Predictions for Algorithm (32')

Algorithm (32')	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	23	39.7	8	47.1	36	92.3	67	58.8
Incorrect predictions	35	60.3	9	52.9	3	7.7	47	41.2
Total	58	100	17	100	39	100	114	100

Table (66) shows a 58.8% success rate for Algorithm (32'), which is lower than Algorithms (30) and (31). While Algorithm (32') predicts X-initial better than the other two orders, it does noticeably poorly for predicting S-initial and O-initial.

4.2.4.5. Algorithms-demonstrative, -*wa*, -temporal and -topic

Tables (67) and (68) show the number of tokens wherein the initial constituent is either *wa*-marked or demonstratives for S-initial, O-initial and X-initial clauses that contain either at-locatives or time expressions.

Table (67) The properties of initial constituents in clauses containing at-locatives: # of tokens

Properties	Subject	Object	X	Total
marked with wa	4	0	0	4
demonstrative	0	2	6	8
marked with wa and demonstrative	1	1	1	3
none of above	3	0	4	7
Total	8	3	11	22

Table (68) The properties of initial constituents in clauses containing time expressions: # of tokens

Properties	Subject	Object	X	Total
marked with wa	3	1	2	6
demonstrative	3	1	4	8
marked with wa and demonstrative	0	0	0	0
none of above	12	1	10	23
Total	18	3	16	37

Tables (67) and (68) show that the proportion of the number of tokens wherein the initial constituent is either *wa*-marked or demonstrative to the number of tokens wherein the initial constituent is neither *wa*-marked nor demonstrative is higher in clauses containing at-locatives than those containing time expressions. This result suggests that even if I leave out the factor at-locatives in Algorithm-X, algorithms based on *wa* and demonstrative will correctly predict constituent ordering for tokens containing at-locative constituents. Based on this observation, I will try an algorithm without Algorithm-locatives. The algorithm I will try is Algorithm (33).

Algorithm (33)
(a) If there are demonstrative constituents, and
  (i) If the S is demonstrative, use S-initial,
  (ii) If the O is demonstrative, use O-initial,
  (iii) If the X is demonstrative, use X-initial,
  (iv) If the S and O are demonstrative, use S-initial,
  (v) If the S and X are demonstrative, use S-initial, or
  (vi) If the O and X are demonstrative, use X-initial.
Otherwise,
(b) If there are *wa*-marked constituents, and
  (i) If the S is marked with *wa*, use S-initial,
  (ii) If the O is marked with *wa*, use O-initial,
  (iii) If the X is marked with *wa*, use X-initial,
  (iv) If the S and X are marked with *wa*, use S-initial,
  (v) If the S and O are marked with *wa*, use S-initial, or
  (vi) If the O and X are marked with *wa*, use X-initial.
Otherwise,
(c) If there is an at-locative, time expression or topical element, then use X-initial.

The subalgorithms (a) and (b) in Algorithm (33) are equal to Algorithm (30), which consists of Algorithm-demonstrative and Algorithm-*wa*. Algorithm (30) was applied to the 80 tokens that contain demonstrative and *wa*-marked constituents, and yielded 70% of correct prediction rate (cf. Tables (42) and (43) in Section 4.2.4.1.). Table (69) shows the result of predictions for Algorithm (30).

Table (69) Predictions for Algorithm (30)

Algorithm (30)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	26	63.4	11	73.3	19	79.2	56	70.0
Incorrect predictions	15	36.6	4	26.7	5	20.8	24	30.0
Total	41	100	15	100	24	100	80	100
# of remaining tokens	17	50.0	2	5.9	15	44.1	34	100

Now, we will apply the subalgorithm (c) to the 34 tokens that contain neither demonstratives nor *wa*-marked constituents (i.e. 114 tokens minus 80 tokens).

Table (70) shows that 19 of the 34 tokens contain either an at-locative, time expression or topical element.

Table (70) The number of tokens containing at-locative, time expressions and topical elements in the 34 tokens that contain neither demonstrative or wa-marked constituents

containing	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
At-locative	3	60.0	0	0.0	2	40.0	5	100
Time expression	4	33.3	1	8.3	7	58.3	12	100
Topical	0	0.0	0	0.0	1	100.0	1	100
Time expression and at-locative	0	0.0	0	0.0	1	100.0	1	100
None of above	10	66.7	1	6.7	4	26.7	15	100
Total	17	50.0	2	5.9	15	44.1	34	100

Considering the number and position of at-locatives, at-locative does not seem to be a good factor for predicting constituent ordering; accordingly, I will modify the subalgorithm (c) in Algorithm (33).

Algorithm (34)
(a) If there are demonstrative constituents, and
 (i) If the S is demonstrative, use S-initial,
 (ii) If the O is demonstrative, use O-initial,
 (iii) If the X is demonstrative, use X-initial,
 (iv) If the S and O are demonstrative, use S-initial,
 (v) If the S and X are demonstrative, use S-initial, or
 (vi) If the O and X are demonstrative, use X-initial.
Otherwise,
(b) If there are *wa*-marked constituents, and
 (i) If the S is marked with *wa*, use S-initial,
 (ii) If the O is marked with *wa*, use O-initial,
 (iii) If the X is marked with *wa*, use X-initial,
 (iv) If the S and X are marked with *wa*, use S-initial,
 (v) If the S and O are marked with *wa*, use S-initial, or
 (vi) If the O and X are marked with *wa*, use X-initial.
Otherwise,
(c) If there is a time expression or topical element, then use X-initial.

The subalgorithm (34c), when applied to the 19 tokens that contain time expressions and topical elements, it correctly predicts 9 tokens of X-initial order. The result is shown in Table (71).

Table (71) Predictions for Algorithm (34c)

Algorithm (34c)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	0	0.0	0	0.0	9	100.0	9	64.3
Incorrect predictions	4	100.0	1	100.0	0	0.0	5	35.7
Total	4	100	1	100	9	100	14	100

When we combine the results in Tables (69) and (71), we can obtain the result of predictions for Algorithm (34a,b,c). Table (72) shows the result.

Table (72) Predictions for Algorithm (34a,b,c)

Algorithm (34a,b,c)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	26	57.8	11	68.8	28	84.8	65	69.1
Incorrect predictions	19	42.2	5	31.3	5	15.2	29	30.9
Total	45	100	16	100	33	100	94	100
# of remaining tokens	13	65.0	1	0.0	6	30.0	20	95

There are still 20 tokens that none of the algorithms have applied to. I will apply Algorithm-RD to the 20 tokens.

Algorithm (34)

(a) If there are demonstrative constituents, and
  - (i) If the S is demonstrative, use S-initial,
  - (ii) If the O is demonstrative, use O-initial,
  - (iii) If the X is demonstrative, use X-initial,
  - (iv) If the S and O are demonstrative, use S-initial,
  - (v) If the S and X are demonstrative, use S-initial, or
  - (vi) If the O and X are demonstrative, use X-initial.

Otherwise,

(b) If there are *wa*-marked constituents, and
  - (i) If the S is marked with *wa*, use S-initial,
  - (ii) If the O is marked with *wa*, use O-initial,
  - (iii) If the X is marked with *wa*, use X-initial,
  - (iv) If the S and X are marked with *wa*, use S-initial,
  - (v) If the S and O are marked with *wa*, use S-initial, or
  - (vi) If the O and X are marked with *wa*, use X-initial.

Otherwise,

(c) If there is a time expression or topical element, then use X-initial.

Otherwise,
(d) (i) If the RD of one of the S, O and X is lower than the RD of the other two, then the constituent with the lowest RD comes first.
Or
(ii) If the RD of the O and X is the same, but lower than the RD of S, then use O-initial.
Otherwise,
(iii) Use S-initial.

Table (73) shows the result of predictions for Algorithm (34d) when applied to the remaining 20 tokens.

Table (73) Predictions for Algorithm (34d)

Algorithm (34d)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	9	69.2	1	100.0	1	16.7	11	55.0
Incorrect predictions	4	30.8	0	0.0	5	83.3	9	45.0
Total	13	100	1	100	6	100	20	100

When we combine the result in Tables (72) and (73), we can obtain the result of predictions for Algorithm (34). Tables (74) and (75) shows the result.

Table (74) Breakdown of predictions for Algorithm (34)

	Algorithm (34)	(a)		(b)		(c')		(d)		Total	
		#	%	#	%	#	%	#	%	#	%
Correct prediction	S-initial	9	26.5	17	77.3	0	0.0	9	81.8	35	46.1
	O-initial	9	26.5	2	9.1	0	0.0	1	9.1	12	15.8
	X-initial	16	47	3	13.6	9	100.0	1	9.1	29	38.2
	Total	34	100	22	100	9	100	11	100	76	100
Incorrect prediction	S-initial	12	75.0	3	37.5	4	80.0	4	44.4	23	60.5
	O-initial	3	18.8	1	12.5	1	20.0	0	0.0	5	13.2
	X-initial	1	6.3	4	50.0	0	0.0	5	55.6	10	26.3
	Total	16	100	8	100	5	100	9	100	38	100

Table (75) Predictions for Algorithm (34)

Algorithm (34)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	35	60.3	12	70.6	29	74.4	76	66.7
Incorrect predictions	23	39.7	5	29.4	10	25.6	38	33.3
Total	58	100	17	100	39	100	114	100

Table (75) shows a 66.7% success rate, which is the highest among Algorithms we have tried (63.2% for Algorithm (30), 62.3% for Algorithm (31), and 58.8% for Algorithm (32)).

4.2.4.6. Final algorithm

Algorithm (34) contains subalgorithm (c) that states,

(c) If there is a time expression or topical element, then use X-initial.

The subalgorithm (34c) yielded a 64.3% correct prediction rate, when it was applied to the 19 tokens that contain time expressions and topical elements. I will try a modified version of the subalgorithm (34c), the subalgorithm (c'')

(c'') If there is a time expression or topical element, and
(i) if the RD of S = 1, then use S-initial
Otherwise,
(ii) use X-initial

I will apply the subalgorithm (c'') to the 19 tokens containing time expressions and topical elements. The result is shown in Table (76). For comparison, I will display the result of the subalgorithm (c') in Table (77).

Table (76) Predictions for the subalgorithm (c")

subalgorithm (c")	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	1	25.0	0	0.0	6	66.7	7	50.0
Incorrect predictions	3	75.0	1	100.0	3	33.3	7	50.0
Total	4	100	1	100	9	100	14	100

Table (77) Predictions for the sublgorithm (34c)

subalgorithm (34c)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	0	0.0	0	0.0	9	100.0	9	64.3
Incorrect predictions	4	100.0	1	100.0	0	0.0	5	35.7
Total	4	100	1	100	9	100	14	100

Despite my attempt to find a better algorithm than Algorithm (34), the result of predictions for the subalgorithm (c'') indicates that the subalgorithm (c'') is not a promising alternative.

In conclusion, Algorithm (34) produced the highest overall prediction rate. I will repeat here Algorithm (34), and Tables (74) and (75) that show the results of predictions.

Algorithm (34)

(a) If there are demonstrative constituents, and
   (i) If the S is demonstrative, use S-initial,
   (ii) If the O is demonstrative, use O-initial,
   (iii) If the X is demonstrative, use X-initial,
   (iv) If the S and O are demonstrative, use S-initial,
   (v) If the S and X are demonstrative, use S-initial, or
   (vi) If the O and X are demonstrative, use X-initial.

Otherwise,

(b) If there are *wa*-marked constituents, and
   (i) If the S is marked with *wa*, use S-initial,
   (ii) If the O is marked with *wa*, use O-initial,
   (iii) If the X is marked with *wa*, use X-initial,
   (iv) If the S and X are marked with *wa*, use S-initial,
   (v) If the S and O are marked with *wa*, use S-initial, or
   (vi) If the O and X are marked with *wa*, use X-initial.

Otherwise,

(c) If there is a time expression or topical element,
   then use X-initial.

Otherwise,

(d) (i) If the RD of one of the S, O and X is lower than the RD of the other two,
   then the constituent with the lowest RD comes first.
   Or
   (ii) If the RD of the O and X is the same, but lower than the RD of S,
   then use O-initial.
   Otherwise,
   (iii) Use S-initial.

Table (74) Breakdown of predictions for Algorithm (34)

	Algorithm (34)	(a)		(b)		(c')		(d)		Total	
		#	%	#	%	#	%	#	%	#	%
Correct prediction	S-initial	9	26.5	17	77.3	0	0.0	9	81.8	35	46.1
	O-initial	9	26.5	2	9.1	0	0.0	1	9.1	12	15.8
	X-initial	16	47	3	13.6	9	100.0	1	9.1	29	38.2
	Total	34	100	22	100	9	100	11	100	76	100
Incorrect prediction	S-initial	12	75.0	3	37.5	4	80.0	4	44.4	23	60.5
	O-initial	3	18.8	1	12.5	1	20.0	0	0.0	5	13.2
	X-initial	1	6.3	4	50.0	0	0.0	5	55.6	10	26.3
	Total	16	100	8	100	5	100	9	100	38	100

Table (75) Predictions for Algorithm (34)

Algorithm (34)	S-initial		O-initial		X-initial		Total	
	#	%	#	%	#	%	#	%
Correct predictions	35	60.3	12	70.6	29	74.4	76	66.7
Incorrect predictions	23	39.7	5	29.4	10	25.6	38	33.3
Total	58	100	17	100	39	100	114	100

Half of incorrect predictions for S-initial and O-initial are the outcome of Algorithm=Demonstrative (i.e. Algorithm (33a)). These tokens contain either more than one demonstrative or demonstratives in non-initial position. The majority of initial constituents in these tokens have a lower RD than the other two constituents. The incorrect predictions for X-initial that Algorithm-*wa* (i.e. Algorithm (33b)) produced are the clauses containing *wa*-marked subjects, although Xs in these clauses are either scene-setting or topical elements. All Xs in the clauses that Algorithm-RD (i.e. Algorithm (33d) did not correctly predict are syntactically heavier than the other two constituents in the same clause. The success rate for Algorithm (33), which involves more then three constituents in a clause, is lower than that for Algorithm (26) which correctly predicted the choice between SOV and OSV orders for 88% of the data. The reason lies in the complexity in formulating algorithms, which reflects the complexity of sentence production or processing in human minds.

4.3. More on What precedes what vs. what comes first

Yamashita and Chang (2001) and Yamashita (2002) report a tendency for syntactically heavy constituents to shift clause-internally in front of syntactically light constituents (i.e. scrambling two clause-internal entities). For example, Yamashita and Chang demonstrate experimentally that Japanese speakers tend to shift accusative objects in front of dative objects when direct objects are long than when they are short, and Yomashita presents her finding based on her corpus analysis that in both “canonical” order and “internally scrambled sentences”, heavy constituents tend to precede light ones.

At the beginning of Chapter 4, I argued that those factors identified in Chapter 3 that predicted the ordering choice of SOV vs. OSV are more relevant to determine what comes first than what precedes what. Through the study in this chapter, we learned that factors that apply to a set of data such as the particle *wa*, demonstratives, scene-setting locatives and temporal expressions and topical elements, are strongly the properties of clause-initial constituents. At the same time, we observed that initial constituents tend to be the ones where the RD is lowest, where the TP is highest, or where the SW is highest.

In what follows, I will examine the order of OX vs. XO in SOXV and SXOV clauses to see to what extent the factors RD, TP and SW are relevant to the ordering of clause-internal constituents.

4.3.1. OX vs. XO in SXOV and SOXV

In this section, I will look at the OX and XO orders in SOXV and SXOV clauses. Table (78) shows our data for this.

Table (78) The number of tokens : SOXV vs. SXOV

3NPs	Subject Initial		Total
Type of X	SOXV	SXOV	
	#	#	#
Dative (D)	SODV	SDOV	
	3	2	5
Locative (L)	SOLV	SLOV	
	2	6	8
Instrumental (I)	SOIV	SIOV	
	2	4	6
Quotative (Q)	SOQV	SQOV	
	5	1	6
Total	12	13	25

In both SOXV and SXOV, *wa*-marked constituents are all objects (one in SQOV, one in SOLV, and two in SOQV), and demonstrative constituents are either subjects or Xs. Cohesive objects occur only in SXOV clauses. Tables (79)-(81) show the interaction of relative RD, TP and SW of the X and O in SXOV and SOXV clauses.

Table (79) Relative RD of X and O in SXOV and SOXV orders

RD	SXOV		SOXV		Total	
	#	%	#	%	#	%
X>O	3	75.0	1	25.0	4	100
X=O	7	50.0	7	50.0	14	100
X<O	3	42.9	4	57.1	7	100
Total	13	52.0	12	48.0	25	100

Table (80) Relative TP of X and O in SXOV and SOXV orders

TP	SXOV		SOXV		Total	
	#	%	#	%	#	%
X>O	2	66.7	1	33.3	3	100
X=O	8	53.3	7	46.7	15	100
X<O	3	42.9	4	57.1	7	100
Total	13	52.0	12	48.0	25	100

Table (81) Relative SW of X and O in SXOV and SOXV orders

SW	SXOV		SOXV		Total	
	#	%	#	%	#	%
X>O	7	70.0	3	30.0	10	100
X=O	3	60.0	2	40.0	5	100
X<O	3	30.0	7	70.0	10	100
Total	13	52.0	12	48.0	25	100

Table (79) only shows that when the RD of X is higher than the RD of O, it is most likely SXOV. This indicates that RD is not relevant to determining the order of clause-internal

Xs and Os. Table (80) shows no significant difference between the two clauses. The difference in occurrences are only by one token in all three categories of relative TP between the S and O. Table (81) shows that when the SW of X is higher than the SW of O, it is most likely SXOV, and when the SW of O is higher than the SW of X, it is most likely SOXV. This result corresponds to the finding of Yamashita that the placement of syntactically heavier accusative objects in front of dative objects. In sum, the results suggest that SW is relevant to clause-internal constituent ordering, while RD and TP are not.

4.4. Summary

In this chapter, the relevance of various factors in the ordering choice of SOV vs. OSV is tested for the tokens in the ≥3NPs constructions. For predicting what comes first among the subject, object and other phrases, the study proved that factors demonstrative, the particle *wa*, the semantic type of locatives and temporal expressions, i.e. scene-setting elements such as at-locatives and time expressions, topical elements play an important role in predicting the clause-initial element. The examination of the X and O in SXOV and SOXV clauses revealed that to some extent TP and SW play a role in determining what precedes what for clause-internal constituents while RD does not.

Table (50) gives a summary of occurrences when each factor favors S-initial, O-initial and X-initial orders.

Table (82) Comparison of factors

Factors	S-initial	O-initial	X-initial
	%	%	%
RD	63.2	52.9	62.5
ERD	62.5	54.5	50.0
TP	50.0	19.0	62.5
SW	42.1	20.0	37.5
The particle wa	69.2	50.0	83.3
Demonstrative	88.9	81.8	61.9
Overall Occurrence	50.9	14.9	34.2

RD is one of the most reliable factors determining what the initial element will be. The factors particle *wa* and demonstrative are the two strongest factors for S-initial. For O-initial, the factors ERD and demonstrative are the two strongest factors. The factor particle *wa* is the strongest factor for X-initial. While TP is not relevant in predicting S-initial or O-initial, it was found to be a strong motivation for X-initial. This is because the Xs in X-initial clauses are mostly scene-setting elements (i.e. 81.3%), which set up the context for the upcoming discourse. The factor SW is not a strong factor, particularly not for S-initial; where it is not a factor al all, however, the data shows that TP and SW work complementarily with RD. Scene-setting at-locatives and temporal expressions occur in the initial position more frequently than non-scene-setting locatives or temporals, although the algorithm based on the single factor of scene-setting elements does not do well. The cohesion of object and verb is useful in predicting the position of objects, i.e. the immediately preverbal position; it, however, cannot be used as a tool to predict what comes first or what precedes what.

Predicting the ordering choice in the ≥3NPs construction revealed the difficulty in formulating algorithms, which are to clearly state when exactly a specific grammatical form can occur. For example, out of 17 objects in O-initial tokens, 16 objects are either *wa*-marked constituents, demonstrative, or entities that are either more recently mentioned, or more cataphorically persistent, or syntactically heavier than the S and other phrases. However, even the best algorithm that I could come up with made the correct prediction only for 12 O-initial tokens. In addition, the fact that the strength of factors varies in subjects, objects and other phrases made it more complex to formulate algorithms. When factors are in conflict, algorithms should be able to tell you which factor overrides another. In this process of building algorithms, we are forced to invert correct predictions by one algorithm into incorrect algorithms by the next algorithm. I believe this is one of the reason for a lower prediction rate for the tokens in the ≥3NPs construction than for those in the 2NPs construction.

Chapter 5

Concluding remarks

In this chapter, I will first briefly recapture the study in previous chapters. I will then discuss what the properties of earlier constituents identified in previous chapters mean, as well as some issues by summarizing the findings in this dissertation.

5.1. Findings

The goal of this dissertation was (1) to identify the factors in determining the constituent ordering, and (2) to find out how those factors interact when the constituent order is determined. As the study moved forwards, there appeared a crucial question: (3) what exactly do those factors determine, what precedes what or what comes first. The study of constituent ordering in the ≥3NPs construction revealed that while those factors that are applied to a set of data are exclusively to predict what comes first, those factors that are applied to the whole data also play a role, to some extent, in predicting the ordering choice of clause-internal constituents.

This study identified three factors, which are applicable to the whole data, for constituent ordering: the recency of previous mentions that is measured by referential distance (RD), the cataphoric persistence that is measured by topic persistence (TP), and syntactic heaviness that is measured by counting the number of defined syntactic units (SW). The relative measurements of RD, TP and SW indicated that more recently mentioned, more cataphorically persistent and syntactically heavier entities are introduced earlier in a clause.

In addition, all the entities introduced earlier in a string are not necessarily simultaneously more recently mentioned, more cataphorically persistent, and syntactically heavier than the entities introduced later, but rather the earlier entities are either more recently mentioned, more cataphorically persistent or syntactically heavier than the later entities. For example, only 12% of SOV and OSV orders in the present data were predicted by all three factors. Related to RD, an extended RD (ERD) that measures the RD of the initial element of constituents is introduced. For predicting the ordering choice of SOV vs. OSV, ERD proved to be a more useful tool than RD with respect to formulating an algorithm. Algorithm (13) based on extended referential distance (ERD), topic persistence (TP) and syntactic weight (SW) predicted the choice of SOV vs. OSV for 80.2% of the data.

This study identified five other factors which are applicable: the use of the particle *wa,* demonstrative, cohesion of the object and the verb, scene-setting expressions, and topical element. The data in this study showed a strong tendency for constituents marked with the particle *wa*, demonstrative constituents, scene-setting at-locatives and time-temporals, and topical elements to appear in the clause-initial position. The cohesion of the object and the verb is observed among tokens in the light verb construction and idiomatic predicates. In the light verb construction and idiomatic predicates, the object tends to appear in the immediately preverbal position. Studying the constituent ordering in the ≥3NPs construction made it clear that the cohesion of object and verb can be used as a tool to predict the position of the object but not the choice of word orders. Therefore, an algorithm using the cohesion of object and verb as a factor predicts SOV order in the 2NPs construction by default, but it does not necessarily predict the occurrence of XSOV. Algorithm (26) using *wa*-marked subject, *wa*-marked object, demonstrative and cohesive objects as factors, when it is incorporated into Algorithm (13) based on the ERD, TP and SW, predicted the ordering choice between SOV and OSV for 87.4% of the data. For predicting the constituent ordering choice in the ≥3NPs construction, additional factors, scene-setting locatives, temporal expressions and topical elements, are identified. Algorithm (33) based on the factor particle *wa*, demonstrative, time expressions, topical elements and RD predicted 66.7% of the data in the ≥3NPs construction.

This study found that the recency of previous mentions is the most reliable factor for the default ordering, that the cataphoric persistence is an effective factor particularly for predicting the ordering where the subject or other phrases than the subject and object appear in the initial position, and that syntactic weight is the most relevant factor for object-initial ordering. The recency and cataphoric persistence are argued as factors that are applicable to non-verb-final languages (e.g. Arnold et al. 2000, Yamashita and Chang 2001). Syntactic weight is a factor that is applicable to both left-branching and right-branching languages (e.g. Hawkins 1992, 1994, Arnold et al. 2000). This study indicated that while the predictions by the recency of previous mentions and cataphoric persistence tend to overlap, the predictions by the recency and syntactic weight are in the complementary distribution.

One of the important findings in this study is that the default order (in terms of frequency) such as SOV in the 2NPs construction and S-initial in the ≥3NPs construction as well as non-default orders such as OSV in the 2NPs construction and O-initial and X-initial in the ≥3NPs construction are motivated by the same factors. This is an improvement from such research that left the motivation of the default order unexplained while claiming non-default orders are unmarked orders that encode particular features. The statistical data in this study showed that Japanese speakers use the SOV order, not because it is ‘canonical order’, but because there is a reason for speaking in that order. I.e. the subject in SOV and S-initial clauses is more recently mentioned, more cataphorically persistent or syntactically heavier than non-initial constituents (i.e. the direct object in SOV and S-initial clauses and other phrases in S-initial clauses). While the default order and non-default orders are motivated by the same factors, the fact that all algorithms predicted the default order with a higher correct prediction rate than non-default orders indicates that the 'defaultness' per se is a function for constituent ordering.

I performed a binary logistic regression analysis for the 167 tokens in the 2NPs construction on the seven variables, postnominal markings, demonstrative, cohesion of predicates, and relative ERD, RD, TP and SW[1]. The output yielded a similar correct prediction rate for the choice between SOV and OSV orders to the one that Algorithm (26) produced. When a logistic regression analysis was run again with the additional information, i.e. the raw score of ERD, RD, TP and SW for the subject and the object, in addition to the seven variables, the program produced a slightly better correct prediction rate[2]. Under my account, the raw score of ERD, RD, TP and SW for each subject and object was not taken into account. That is, the information on the behavior of individual tokens, such as the behavior of a token with a low RD and a low TP vs. the behavior of a token with a low RD and a high TP, was discussed only on a token-by-token basis. The result of the statistical tests reconfirms the complexity in formulating algorithms which reflects the complexity of human minds.

---

[1] I am very thankful to Douglas Roland for running the program for me.

[2] Since the size of data is so small for the computational analysis that the output of the computer calculation should be taken only for reference. A different set of samples may produce a different output.

5.2. Activation

Givón (1983, 1988) correlates the measurement of referential distance (RD) with ‘predictability/unpredictability’ and that of topic persistence (TP) with ‘importance’. What measurement measures for what function is a crucial issue. Tomlin (1995) emphasizes the importance of form-function relations in linguistic research, and criticizes the lack of clear form-function interactions in Givón’s referential measurements. Givón (1993), while noting a shortcoming of RD, used RD as an index of continued activation and the terminated activation of the currently active referent. I argue that the measurement of RD is a useful index for approximation of relative activation between entities. RD does not measure an absolute cognitive status such as ‘focal attention’. The measurement of RD, however, reflects the degree of activation of the cognitive entity that a linguistic entity or proposition corresponds to. What is relevant is the relative degree of activation of referents in the speaker’s mind or in the speaker’s assumption about the hearer’s mind. More recently mentioned referents (i.e. a lower RD) are more active than less recently mentioned referents (i.e. a higher RD) in the speaker’s mind. The relative degree of activation among the corresponding referents can suggest which referent is more activated in the hearer’s mind that the speaker estimates. The recency of previous mentions is not sole index for the relative degree of activation. The frequency of previous mentions and the presence of potential interference or associating linguistic entities in the preceding discourse may influence the relative degree of activation. The use of demonstratives often reflects the presence of non-linguistic signals such as visual and kinetic signals. The use of particular particles may imply the presence of entities accessible to activation. The relative activation plays an important role to determine the choice of ordering in that more activated entities earn the earlier position in a clause. In many cases, the degree of activation for clause-initial constituents is relevant. In some cases, the degree of activation for the initial element in clause-initial constituents is relevant.

Arnold et al. (2000) argue that conceptually accessible or lexically primed (i.e. accessible at the phonological level) entities tend to be introduced earlier in the sentence by speakers. When the antecedent of a constituent or an initial part of the constituent has the exact same phonological form, it is a case of phonological priming[3]. Phonological priming is more like a physical reflex, reacting to the stimulus. Cases of phonological priming were observed in the present data. The following is one example wherein the very first entity in the subject noun phrase in (1), *"waratte iitomo"* 'It is all right to laugh', (which is a title of a TV variety show), has an antecedent with exact mention in the immediately preceding clause.

(1)

"waratte iitomo"	o	o-tukurininatta	Y-san	to	iu	kata	ga
'It is all right to laugh"	ACC	HON-maku.HON.PST	Y-Mr.	QT	say	person.HON	NOM

anata	o	mitomete-kudasatta-n-desho,	saissyoni
2sg	ACC	recognize-give.PST-NMLZ-CONJ,	first

'Mr. Y who created 'It is all right to laugh' gave you recognition first, didn't he?'

Highly activated entities are conceptually accessible. If there is a physical reflex to the phonological priming, it may be plausible to say there is a cognitive reflex to the conceptual priming. Tomlin showed an experiment result that 'focal attention' has a great effect on what

[3] In my data, phenomena which can be termed as 'conceptual priming', were observed: one concept is realized in two adjacent clauses in different linguistic forms such as a pair of transitive vs. intransitive verbs or affirmative vs. negative mode. For example, the proposition (They are looking forward to X answering incorrectly) in (a) can be said to be conceptually primed by the proposition (They don't want X to answer correctly) in the immediately preceding clause. The two propositions are anchored in one concept, but one is realized in an affirmative clause and another in negation.

(a)

17-3b	K-san	hazureru	no	min'na
	K-Ms.	answer.incorrectly	NMLZ	everyone

tanosimi-ni-site-iru	n-desu	yo
look.forward.to	NMLZ-COP	FP

'Everybody is looking forward to Ms. K's answering incorrectly.'

speakers speak initially. We may interpret 'focal attention' as a kind of a cognitive reflex, if not biological.

5.3. Cataphoric persistence

Topic persistence (TP) is an index of cataphoric persistence of an entity. In Givón (1983, 1988), the measurement of TP corresponds to the notion of 'importance'. In Givón (1993), the function of the cataphoric measurement is more clearly defined. He argues that more important and topical, thus cataphorically persistent referents receive grammatical coding (form) to ground them to the upcoming discourse in the hearer's mind (function). Givón demonstrates that referents introduced by the indefinite article *this* had a higher mean TP value than referents introduced by the indefinite article *a/an*, and while the former plays a central role in cataphoric discourse, the latter does not. That is, the measurement of TP reflects the speaker's intention about an entity with respect to its behavior in the upcoming discourse. If cataphoric grounding is a function of initial constituents, it makes sense that subjects in the initial position have higher cataphoric persistence than other grammatical roles in the initial position. Scene-setting at-locatives and temporal constituents, and topical elements are other devices for cataphorical grounding, and they tend to appear in the initial position.

Because spoken language is dynamic, the speaker does not have control over how the conversation evolves in the following discourse. That is, the speaker's signal as to how he will talk about an entity and the conversation plan that the speaker laid out at the time of utterance may not be understood or received properly by the conversation partner. Arnold (1998) argues that the hearer computes the probability of an entity being mentioned in the upcoming discourse from a linguistic signal that the speaker sends. Arnold calls this probability "probabilistic activation", and 'important' entities are salient in terms of probabilistic activation. Although the speaker's signal that Arnold is concerned with is the choice of reference forms between names and third pronouns, it could be the choice of particles or constituent ordering in a language like Japanese where names are commonly used instead of pronouns.

5.4. Heaviness

The position of syntactically heavy entities in a clause has been argued in parsing theories as an important factor for facilitating the language processing (e.g. Hawkins 1992, 1994). Parsing theories explain two conflicting tendencies, i.e. the short-before-long tendency in right-branching languages and long-before-short tendency in left-branching languages, in terms of the distance that a top-down parser has to spend until it recognizes the syntactical structure of a clause in question. The shorter the distance is, the easier the language processing is for the hearer. Arnold et al. (2000) see syntactically heavy constituents as more structurally complex and hard-to-produce elements, and argue that English speakers tend to produce their utterances in a way that heavier constituents come later in a string, because it facilitates planning and production on the part of speakers. That is, speakers can gain extra time to formulate their utterances by postponing the elements that are hard to produce while uttering the shorter, less complex elements. Yamashita and Chang (2001) view syntactically heavy constituents as lexically salient entities as opposed to more activated entities that are conceptually salient because syntactically heavy entities are more semantically and pragmatically informational. Yamashita and Chang argue that two conflicting tendencies, short-before-long in VO languages and long-before-short in OV languages, suggest that there exist two levels in production of speech, conceptual level and lexical level. According to their production schema, two levels of salience compete with each other for an earlier position, and whether the lexical salience or the conceptual salience wins which position, earlier or later, is language-specific. In this study on spoken Japanese, syntactic heaviness was proved to be one of the determining factors for constituent ordering, in that syntactically heavier entities come before syntactically lighter entities. Heaviness is particularly significant for predicting non-default orders as opposed to the degree of activation, which is most influential on predicting the default order. In addition, Heaviness is the only factor relevant to the clause-internal constituent ordering. The results of this study did not concur the explanation on the motivation of light-before-heavy ordering by Arnold et al. They, however, confirmed the observation by Hawkins and Arnold et al. that the discourse factor works the best when the syntactic factor is weak, and vice versa.

5.5. Contrastiveness

Wa-marked entities are set-anaphoric, i.e. when the speaker uses the particle *wa*, a competing entity for the *wa*-marked entity exists in the speaker's consciousness at the time of utterance. When a competing entity from the set is linguistically realized, it is Potential Interference. Potential Interference (PI), discussed in Givón (1983, 1988), is a linguistic entity whose referent is syntactically and semantically parallel to the referent of an entity in question. The premise is that parallel elements interfere with the identification of the referent in question. Thus, when there are potentially interfering entities, the referent is introduced earlier to facilitate the identification of the referent by the hearer. In this study, the entity with Potential Interference is often *wa*-marked constituents, and *wa*-marked constituents in this study showed a tendency to appear in the initial position. In addition, the present data shows the relation between a higher RD and *wa*-marked constituents, particularly the direct object marked with *wa*. This result contradicts the view of *wa* as a marker for given information. The positional preference of *wa*-marked constituents, the presence of Potential Interference and a higher RD support the view of *wa* as a marker for 'contrast' (e.g. Shimojo 2005, Clancy & Downing 1987).

5.6. Focus and pragmatic presupposition

Focus involves to the cognitive property of 'attention'. Narrow focus can be defined in terms of nonactive and active distribution in a clause; an entity in narrow focus is a nonactive entity when other entities or the predicational proposition in the same clause are active by being mentioned in the preceding discourse. Since the entities in narrow focus are nonactive, they are expected to be first mentions or have a higher measurement of referential distance (RD). There are, however, cases where narrow focus does not represent previously nonactive information. The following example is from Shimojo (personal communication).

(2) (Comparing two items in a shop…)[4]

Customer: Which one of these would be better for me?
Store clerk: This one (would be better).

The data in this study indicates that the use of the particles *mo* and *de* implies the presence of pragmatic presupposition. Pragmatic presupposition involves the speaker's beliefs with respect to the hearer's beliefs, as opposed to 'activation' that involves the speaker's beliefs about what is activated in the hearer's mind. Dryer (1996) states that activation plays an influential role in the position of focal accent (in English) but pragmatic presupposition does not.

---

[4] Shimojo argues that "picking a subset out of a given set" type of context is normally associated with previously activated narrow focus, and thus what is relevant here is not anaphoric activation but identifiability of information.

One of the interests in this study was whether narrow focus or pragmatic presupposition demonstrates any positional preference in a clause. The current study did not find any connection between the earlier position with the use of *ga* or narrow focus. Nor did the constituents marked by particles *mo* and *de* show any positional preferences to be placed earlier in a clause.

Arnold (1998) introduced an experimental result indicating that when the well-established discourse topic and the focus co-exist, the focus is most likely to be referred to as names than pronouns, and thus the well-established discourse topic is more conceptually salient. The well-established discourse topic is a highly active entity in terms of recency and frequency of previous mentions, while the narrow focus is nonactive entity in terms of active and nonactive distribution. Arnold's experimental result suggests, in the case of conflict between discourse topic and focus, that highly activated entities are more conceptually salient than entities in focus.

5.7. Syntactic and semantic information

The present data contain some tokens in which the positional preference is syntactically or semantically constrained. Syntactic integrity is one of the reasons for the positional preference identified in this study. The semantic cohesion between the direct object and the verb in the light verb construction and idiomatic predicates motivates the syntactic integrity of the structure, and consequently, the direct object is anchored in the immediately preverbal position. Syntactic integrity is a weak constraint, since there is no restriction on preposing or postposing the cohesive direct objects. In this respect, it merely induces a positional tendency.

Some constituent ordering in the present data is determined by semantic constraints, i.e. semantic dependency and truth-conditionality. Semantic dependency is a case where the meaning of constituent introduced later in a clause depends on the meaning of a constituent introduced earlier in a clause. One example involves the semantic dependency of the object on the subject in my data. The subject noun is *amerika* 'U.S.A.' and the object noun phrase is *gobangai no tokoro* 'Fifth Avenue area'. 'Fifth Avenue' is 'the Fifth Avenue of New York city in U.S.A.', and its interpretation depends on the semantics of subject noun. Often the dependent constituent introduced earlier and the dependee constituent introduced depicts a whole-part relationship, such as 'Fifth Avenue' and 'U.S.A.'. Truth-conditionality is a case where scrambling the constituent introduced earlier and the constituent introduced later in a clause would change the propositional meaning of the clause.

(3)

22-4a	daredemo	iitibanni	sigoto	o	si-tai	wake	desyo
	anyone	first	job	ACC	do-want	reason	CONF

'Everybody wants to be the first one to get a job, right?'

Although the indefinite pronoun subject *daredemo* in (3) is by definition first mention, it points vaguely to a group of people in the preceding discourse. In order to mean 'everybody wants to be the first one to get a job' as shown in the English equivalent, the order of the subject and the object has to be the way it is in (3). When the subject and the object are scrambled, the reading is confusing between the above reading and the reading, 'everybody wants to work first'. In cases of semantic dependency and truth-conditionality, the choice of constituent ordering is not available. When the meaning of one entity depends on another, the depended has to come before the dependent.

The semantic type of locatives and temporal expressions play a role in determining the constituent order. At-locatives as opposed to to- or from-locatives, as well as time-temporals as opposed to duration- or frequency-temporals, tend to occur in the initial position. They are scene-setting elements that set up the locational and temporal orientation of discourse. Scene-setting elements often override the factors of the recency of previous mentions, cataphoric persistence and syntactic heaviness.

5.8. Production-based motivation vs. comprehension-based motivation

Placement of syntactically heavy entities is where researchers' views split into two conflicting perspectives, production-based and comprehension-based. Some argue that speakers produce syntactic structures that are economical to process in time and effort on the hearer's part, and thus preposing syntactically heavy entities in left-branching languages, and postposing heavy entities in right-branching languages facilitate the hearer's comprehension. Others argue that producing syntactically heavier entities first (in case of right-branching languages) reduces the load in working memory on the speaker's part, or speaking shorter entities first (in the left-branching languages) buys time for the speaker to organize syntactically heavier phrases which are structurally more complex, and consequently facilitate speakers' production.

Anaphoric and cataphoric salience which are commonly associated with earlier position are often interpreted as functions to facilitate the hearer's identification of the referent. We can interpret the speaker's producing more activated entities first in two ways, either from a comprehension-based or production-based perspective. The first entity that the speaker produces is either an entity that is more activated in the speaker's mind or an entity that the speaker believes more activated in the hearer's mind. The speaker produces a more activated entity first because he is concerned about the mental representation of the discourse in the hearer's mind, or simply the speaker refers to an entity that is more activated in his mind as if simply "he puts what he has in mind into words" (Chafe 1976:51). Ferreira and Yoshita propose a production-based perspective for given-before-new preference in Japanese. They argues that speakers do not place an entity in the first position because the entity is given information (i.e. the entity's givenness per se), but they do so because given information is more available information, and speaking more available information first buys time for speakers to retrieve less available information or to form new information.

Cataphoric saliency is mostly viewed as comprehension-based motivation. Placing cataphorically more salient entities first is also pertinent to the speaker's estimate of the mental representation of the discourse in the hearer's mind or consciousness. By using particular linguistic forms, the speaker sends a message to the hearer about the status of a referent in upcoming discourse to facilitate the hearer's organization about referents.

Examples of repairs in the present data suggest a comprehension-based perspective. For example, the speaker tends to give more general ideas or descriptions of an entity first. When the speaker realizes that the referent of the idea or entity he had presented was not easily identifiable by the hearer, he tries to repair his utterance to facilitate the hearer's comprehension by giving a more concrete idea or a detailed description of the entity. The following (4) is an example of repairs. The speaker of (4) first uses a demonstrative pronoun *sore* 'it', and then repairs with a noun phrase accompanied by a demonstrative adjective *sono kuruma* 'that car'. 'It' and 'that car' denote the same referent, however, 'that car' is more concrete description of the referent than 'it'.

(4)	sore	o	desu	ne,	sono	kuruma	o	desu	ne,	K-san	mo	mottete
	it	ACC	COP	FP	that	car	ACC	COP	FP	K-Mr.	also	possess.PPG.TE

'Mr. K also owns it, that car.'

The examples of repairs demonstrated the speaker's concern about the hearer's comprehension, and thus proved that the speech planning is not strictly production-based.

5.9. Final words

Clause-initial entities were found to be more activated, more cataphorically persistent and syntactically heavier than non-initial entities. Initial entities are anaphorically, cataphorically and physically salient. More salient entities in the speaker's consciousness are spoken first to facilitate either production or comprehension. When the three types of salience are in conflict, the anaphoric salience wins. However, most of the time, the factors are complementarily distributed. This result indicates that the factors are working together as a system but not as an individual agent.

Appendix A: Cohesive objects in the light verb construction

Object	frequency				
	SOV	OSV	3NPS	4NPS	Total
~kata 'the way of ~'	2	1	1		4
arubaito 'a part-time job'	2				2
booeki 'import-export'	1				1
en 'a party'		1			1
giron 'discussion'				1	1
hanabi 'firework'	1				1
hanasi 'talk'				1	1
hantai 'opposition'	1				1
kaigo 'nursing'			1		1
kakeoti 'eloping'				1	1
kakkoo 'appearance'	1		1		2
kantoku 'a movie director'			1		1
kao 'face'	1				1
kekkon 'a marriage'	1				1
konsaato 'a concert'			1		1
kossetu 'a fructure'	1				1
koto 'matter'	3	1	2		6
kuroo 'hardship'	1				1
kyarakutaa 'characters'			1		1
kyasutaa 'a newscaster'	1				1
meikyappu 'makeup'			1		1
miseban 'shop-sitting'	1				1
mono 'things'		1			1
nakoodo 'match making'	1	1	1		3
no 'nominalizer'	4		1		5
okkake 'being groupie''		1			1
petto 'a pet'					
risaitaru 'a recital'			1	1	2
ryokoo 'a travel'					
sakusi 'writing the lyrics'	2				2
sensei 'a teacher'	1				1
sibai 'a play'	1		1		2
sigoto 'a job'	4		1	1	6
sikai 'being a host or being MC'	2				2
siki 'conducting'	1				1
simei 'nomination'			1		
syatyoo 'CEO'	1				1
tabi 'a travel'			1		1
tai'in 'a member of a brigade'			1		1
taiki 'waiting'				1	1
tenisu 'tennis'			1		1
tumemono 'stuffing'			1		1
utagassen 'a singing show'	1				1
yuusyoo 'winning a championship'			1		1
Total	35	6	20	6	67

Appendix B: Tokens

SOV Tokens

1. 母親が今日はえり子は具合が悪いので休みますっていう手紙を書いて、
2. すごくうちの両親ってすごくわたしをかわいがってくれまして
3. お母さん達が子供達を連れてきた、
4. で、お母様がヒアノをおひきになるって。
5. 事務所の社長さんかなんかがもう引退しますっていうようなことをおっしたのかしら。
6. プロデューサーの方が気を使われて、
7. ここにつけてた（腰のあたりをさす）竹竿がギシギシギシギシ音を、をこうきしんでですね、で、
8. 子供がドンドンドンドントイレのドアを叩いて、
9. あの、笑っていいとものをお作りになった横沢さんという方があなたを認めてくださったんでしょ、
10. その中の一人がわたしのことを気に入ってくださったらしくて、
11. アメリカがなんか五番街のところをはなふぶき、
12. はい　芸術学科というのがあの映像とかそれからあの舞台要するにあの演劇系とか舞台美術とかそういうのをやってるんです。
13. ま、いろんな方が、もちろん、母親が死んで悲しいってことは書いてらっしゃるけど、
14. あなたがあのう、十朱さんの絵なんかかいてあげたりなんかして。
15. あの、前田とか東金之助っていうのが５０万円づつだしてくれました。
16. 古館が司会、ねえ、してらして
17. 奥さんが、それこそ、カレーライス作ったって
18. 小沢せいじさんが指揮なさって
19. やっぱりご自分がこれだけ長く評論書いていらっしゃって、
20. 酔っぱらいの人がなんか面白い歩き方してたんですよ。
21. その人がいろんなことやったんですけど、
22. うちの父が店番しております。
23. 監督がたぶんあなたで決めるからみたいなこと言ってくれて、
24. もう一人サラリーマンの方がうーんってわたしのこと見て、
25. 世の中の人全部がそういう育て方してるわけでもないでしょ
26. あなたが、あれ、作詞なさったの。
27. 饗美が、僕と女房と別々に、その、呼び出して、
28. 娘があれして
29. でもお客様が傘さしてらっしゃるんでね、
30. だけどあなたがたが問題だすじゃない？
31. みんながいろんなことして、
32. あたしが最後のメロン食べてて
33. 武田たいじゅんなんて人が先生やってました。
34. あの方がねあの冷蔵庫なんか開けてね
35. ご主人がなんにも言わないで、
36. 要するに、かれは要するにハーバードを出て、
37. この方は、あの、劇団を持っていらっしゃいまして、
38. 渡辺えり子さんは結婚をしてらっしゃいませんで、
39. そんな、子供はすごい気を使ってる訳ですよ、
40. そいで、ことに女の人はねいろんな苦労をしてて、
41. 僕はありのままを書いたつもりなんです。
42. 僕はそれを断わったんです。

43. わたくしは今度の芝居をして
44. おとうさまはもちろん貿易会社の社長さんをやってらっしゃるんでしたっけ。
45. あなたはま、ちょっとほんとにお店をま、お開き、念願のお店開いたんだけど、
46. わたしはこうデザインを考えて、
47. あなたはずいぶんアルバイトをして
48. ま、あなたを支持して下さった先生はあなたのご活躍をお喜びで、
49. 私は、自身は、反対を向いている、
50. いや、僕は大変なものを見せていただきましたよ、
51. ハルマンさんはその二人の子供をつれて、
52. とにかくあなたはくるぶしのすごいもう考えられない骨折をして、
53. おやじさんはこっちを見てる、
54. あの、おやじさんは仲人を全然やってらっしゃらなかったんです。
55. で、しょうがなく、父は、なんか、斧をもって、
56. わたしたちは、もう、テーブルを出して、
57. わたしは風俗嬢を演じました。
58. 僕はすぐ手品のネタをおしえ、教えちゃうんですけどね。
59. あなたのお父様はお小さいあなたとかみなさんをつれて
60. であたしたちはそれをね見て
61. イングリットバーグマンはその映画をみて
62. 僕は、はあ、僕のことを怖いって言うのもいるんだなってことは知りました。
63. わたくしはもうあの大学は卒業して
64. この方は結局サッカーはあきらめて、
65. わたしはちょっとあまり本よまないんでね、
66. ですから僕は当然その二頭買ったわけですよ。
67. 僕は、とっさに僕は自分の将来ききたいんですけど、
68. あなたはそのパン屋さんのほかにずいぶんアルバイトやってらっしゃるんですってね。
69. で、まあ、僕は本出したんですけども。
70. 好きな人は五万円払えばいいし、
71. ホロビッツは、それ、聞いたんですね、
72. 僕はそれ知りませんでしたけども。
73. あのね、私は割に職業かわったんですね。
74. それであなたは上智大学の比較文化学科おでになって、
75. 三本木君とこのうた君は動じない顔して、
76. （写真が画面に出る）わたくしは実物拝見、（笑い）拝見したことはない、
77. でも、あなたはやっぱりこういう仕事やってみたかったの。
78. あなたは全然違う、地域活性化コンサートみたいなのやってらっしゃるんですって
79. 僕はそれ見てね、
80. ま、岸田今日子さんは、まあ、ご本もお書きになったり、
81. わたくしは、あの、他の政党がたくさんま、できて、いろいろそれなりの理由があってですね、新党が結成されているのも見ておりますけども、
82. あたしはほらだって戦争中も戦後も知ってるでしょう
83. 根津さんなんて、あのー、こうですね、いろいろこう、ちょうどモトクロスバイクのパットみたいなのをされてね、
84. 今日のお客様、ほんとに、CNN のキャスターをなさって、
85. お父様真剣にそれを考えてくださってました。
86. 誰一人反対をしませんでした。
87. あなたなんとか卵っていうのを食べてるんだって？
88. もう、みなさんダライラマのこう手を握ったり、
89. で、あなたなにか、ろっ、何かの傷跡、ってご本を今度、ま、お出しになった

90. やっぱり誰でもその一番に仕事をしたいわけでしょ
91. あたくしあのう紅白歌合戦の司会をやらしてただいたでしょ
92. 皆様お名前はご覧になったことあるかもしれないけど、
93. あたしパンダがね、ビスケットもらった時にね、あのう、アヒルにやったって話しは聞いたことありますけどね。
94. それで、あの、あの方、なんかサッチャーさんの周辺のことかいたりなんかなすったりね。
95. 割とあなた普通のこう大学教授とか、普通の人というよりは、普通の人なんだけど、なんか結末大変になっちゃうってのずいぶんおやりになってらっしゃるわねえ、
96. でもお父様少しお金送ってくださったんですってね。
97. お父様貿易、やってらっしゃたんですものね、
98. みんなものすごくおいしそうなの、たべるんですよね、
99. まあ、みんなお相撲さんの格好してやることになった、
100. わたしあんまり黒柳さんが ほ、ほ、ほ、ってのけぞって笑うの見たことないんですけど
101. あなた、じゃ、好きな物おっしゃってみていただけません？
102. 私全然夢みないのね、
103. フランキーさんあなたのこと気にかけていたそうです。
104. 奥様おこずかいそんなにくださらなかったんですって。
105. わたし達車運転して、
106. お父様斧もって出かけて行って、
107. だからお父様こうあのいろんな事なさって
108. そいであたしねえ　えー　２つしかないねえアナウンサーの部屋とかあのスタジオ見せてね
109. あたし賞金もらって
110. 寅さんまだ面白いギャグもなんにも言ってないんですよ、
111. 紘子さんお書きになるもの大宅壮一賞とか賞もとってらっしゃるし、
112. あなたバンドもつれてらっしゃるんでしょ。
113. 劇団員の人もみなお金をあつめて、
114. 女の人も仕事をしてるし、
115. 男も仕事をしていて、
116. 僕も目を見て、
117. えっと、なんか、わたくしも、ろくろく、その注意を払ってないんですけど、
118. わたしも写真見ててね、
119. 主人もお皿洗ってくれますし、
120. わたしも 酒と涙と男と女っていうの見ましたもの、
121. あたくしも紅白歌合戦ご一緒にやらせていただいたりいたしましたけれども
122. でもう村の人もみんな映画見にいく
123. 姉と私と二人で、あの、作詞をして、
124. みんなで、こう、いろんなこと好き勝手を言い合って、
125. 自分でおにぎりを作ってね
126. みんなでずーっと音楽はなしてらしたでしょ。
127. みんなで旗ふったりしますよね。
128. 家族中みんなでメールもってるの。
129. みんなで花火してたんです。
130. あの、僕と女房で寿司とか食べながらね、

OSV Tokens

1. で、それをその女子学生がみてて、
2. それをわたくしの家内が見て、
3. 柱時計をおばあさまがずれたの。

4. そういう写真週間誌のゲラ刷り？その発売される前の段階のものを各局テレビ局のワイドショーの方々、スタッフの方々が見られるんですよね。
5. それを本木君あの徳川慶喜でご一緒だったあの方が見て
6. その仲人をうちの両親が頼まれでいたしまして、
7. 僕の目の前を松崎真さんが通ったんですよ。
8. ただお別れの宴をねえあの科学番組の連中がずっとやってくれましてねえ
9. 寛三郎の声をねなんとなくぱっとこうあたしがねまねをしてみたけど
10. ですから、自分自身が理想的な自民党であって、いまの自民党はかなりゆがんでいるっという風な考え方をわたしはしております。
11. それをわたくしは拝見しました、
12. あのう、こういう風な酒の歌をあなたはうたって
13. その時の状況を、あのう、ま、いろんな方書いてらっしゃいますけど、
14. 前は女の人の数の方がおいくって、女の人の方が投票に行ったっていうことをね、市川房枝さんおっしゃってらしたんですけど。
15. でも、もうあと、あとどのぐらいの所を、このぐらい時間かかってるんですか。
16. 主役の女の子をあなたいびる、
17. その絵を僕かいたんですよ、
18. その車をですね、昔、力道山も、さんももってて、
19. こういうものはね、生まれつきやりたいと思ってた人がやるべきだってね、
20. 食事は女の人が作って、
21. あのー、メロディーはわたしが作りまして、
22. しかもあの琥珀っていう、いうものの中にそれが入っているってことはみんな知ってて、
23. ま、とにかく真紀子さんが女優さんになろうと思ってらっしたのはみなさんご存じで、
24. 芝居は皆さん喜んで下さいましたか
25. 今日のお客様ほんとに個性的な声とそのたたずまいと、ま、演技力で、ほんとにユニークな女優さんでいっらっしゃることは皆様よくご存じで、
26. それは渥美清さんもよくおっしゃってましたよ。
27. でもねえあのほんとに父がおじいさんになる絵っていうのは誰も家族もう誰もみいこうみうかばらないってんですかうかべられないでしょ
28. いかにみんながわたしに当てさせさくないかってスタッフがよくご存じなんですって。
29. でも、あのう、これ、ちょっとあのう、スタッフが見つけてきたんですけど、
30. まあ、ずいぶん、北川真知子さんの映画わたしもずいぶんはいけんしてたんですけども、
31. あれ私も読んでも鮮明に覚えてますし
32. だって、違うこと皆してるわけ、訳で、
33. これなんてみなさんご覧ください、
34. もー、黒柳さんはずれるのみんな楽しみにしてるんですよ。
35. これ、人、まあ、間違えて、
36. これいっつも私持って歩いて
37. おっかけも３人でやったりなんかするんですけど、

≥3NPs Tokens in SO order

1. Shall we dance をおつくりになった周防まさよし監督があなたを、あの、全部おつくりになった映画にきようしてらっしゃいますよね。
2. わたくしはだいたい、あの、ものを、点字図書館にいつも朗読してる訳じゃないですど、
3. 先生自体が、あの、どんどん、こう、指示の出し方を、あの、また新しい新しい方法にかえてくだすっていて、
4. その、本木さんがとてもそれに、ま、興味をお持ちになって、
5. わたし自身が、ずごく、きれいな声って言われることに抵抗をもってたんですね。
6. チンパンジーが子猫を自分のペットにして

7. 男の子があなたのことふとってるって言ったんですって。
8. 僕は、人形町やなんか、とうとうとう通って、そんで、銀座にいったりなんかするのもなんとも思わなかったし、
9. すれちがったサラリーマンのかたが、一応外見的なことは、あ、いい女っていわれたんですよ。
10. 父はもう私のことはできるだけ政治にたずさわってほしくないっておもったらしんですね、
11. あなたが、ご自分のお母さんのことを、お母さんたすけてって、思わず言ったの。
12. ま、私もだから、ママって、な、娘のことは言って、
13. 仙台からね東京行くって僕ね内示みたいなの受けてたんですね。
14. 「中村屋１７代目」とみんな声かける
15. 僕はねえ、そのあとね、あなたがだんだんだんだんののいろんなとこで活躍するようになったのを脇からはいけんしてて、
16. ご自分のほうも、二人いる方のうちの男の子のほうはむこうにお渡しになって、
17. この方がお書きになりました NHK の音静の海に眠れというのは、プラハ国際テレビ祭り、テレビ祭りでグランプリをおとりになった。
18. うちの母は、あのう、こう、なんか、頭に、こう、あんこっていってなんか詰め物をして、
19. で、それが脳に刺激を送って、
20. それからあなたショパンコンクールで優勝もなさってらっしゃいますよね。
21. みんなわっーとすごい大きな所でコンサートやるんだけど、
22. わたし達は映像でも歌の声を聞くことができる。
23. だからたまたまそのプールサイドで僕自身も本よんでたんですが、
24. え、お店で店員さんが帽子をかぶって
25. そんな中でこの方はほんとにもういろんなキャラクターなさって、
26. 屋上でね三人があのそれぞれ別のことをして、
27. そこでうちの両親が仲人をやって
28. で、ほうぼうの店から僕招待状もらってるんですけども、
29. そうすると舞台の上手の方にもう中村やが盲人やいちの格好をして
30. 僕もうぜんぶ寝技で読むんですよ。
31. その若い人があのなんかえらい人をこう目くばせで呼んで
32. お父様はお父様で気を使って、
33. 子供は子供で気を使って、
34. やっぱりもともとわたくしは兄の影響でねあのうまあ映画みるようになったんでえ
35. 大映と松竹でね共作っていう形で同じ話をねつくったんです
36. この方お父様と一緒に隊員、ウルトラマンの隊員、変身しない隊員やってらしたんですって。
37. 僕は年とともに年とりました。
38. そういう人達と一緒にほんとにあのうみんな力をあわせて、
39. あら、橋田須賀子さんその後意地悪な、あの、いろんな人お書きになったけども、
40. そのお父様が今日店番しててくださってるの。
41. あなたのほんとに昔からそういうのやりたかったし
42. ちょうど僕がその前の日夢を見て、
43. それで、あなたは前に、ほら、子供向けの歌を、あの、まりちゃんのためにといってもいいくらい、cd を、パンツのはきかたとか、おだしになったことあるじゃないい、
44. あの、やはり主人が父が病気になりました時にですね、介護をやっていて、
45. わたしね、最近ね、あの、小説新潮の 10 月号のね、小さい時から考えてきたこと、っていう、あなたの、あの、隔月連載、あれ、読んで
46. この間のセリーヌディオンがなんかのインタビューの時に、ヴォイスドクターっていうようね、あの、言い方をしてらしたんですね。
47. わたしもこの間芝居の旅してたんですが、

48. わたしさっき名前わすれちゃ、間違ったんですけど、
49. 中村屋がその間にメーキャップをして
50. ドイツはまあひところあのニュージャーマンシネマっていうのやってまして
51. 皆さんあのうその日は早く仕事を終えて
52. 戦争中はねえ、殊に戦争に反対の人はなんにもいえないんですよ。
53. 年に一回竹中さんこうお芝居をやられてるんですけど、
54. 今あなた、監督してるんですって
55. ずいぶん前からわたしは興味をもって、
56. その時にあなたお手紙、あ、お、おで、お電話番号をおあげになったんですって。
57. でも、寝る前にあなたご本読んだりしてあげてらっしゃうんでしょ。
58. ある日ね、わたくしテニスをやってたんですよ。
59. だからこの間わたし蟹のお茶碗を送ったの。
60. 今まあ、お父様スペアタイア持って帰って来る。
61. でも、その時フランキー堺さんが　当時は珍しいその、 瞬間カメラ、一瞬にしてとれてピッってひっぱるやつ、それをなんか手にいれたんですって。
62. 1977 年、僕はあの寝袋かついで、
63. あのう今年二人でね、モンゴル旅行したんですよ。
64. 戦後僕らあのイタリアンリアリズムなんていう映画を見ましてね
65. でも、そういうものが逆に言えばあなたを強くしましたかね。
66. お母様がお若くしてね、お父様おなくしになって、
67. そして、いい監督さん、いい演出家の方がこの方じゃなければということで、あの、この方をご指名になるという、
68. だから僕も、それはそうだねってことで、それを置いて
69. こんなにみんな気持ちそろえて
70. そのお金をめあてに寅フグのたじゅうとお市という夫婦がおい夫妻がねおじさんをころして
71. ああいう所はステーキ屋さんもステーキたべさせてくれるんですか。
72. ジュラシックパークってのは私本をよみました。
73. 今ああやって時間出したりするのも全部/彼が時間出してるんですよね、
74. これはりゅうたつという坊主が小金をためて、
75. そこで三百人がテラスでよちゃんとテーブルかこんで
76. その学校で先生がそこの所にそういうパンフレットを置いて、
77. 今まであなたずいぶん長いことお仕事してらっしゃったんだけど
78. どうしてあなたそんなにせいつけたいの。
79. あたくしがね東京で三か月アナウンサーの養成を受けて
80. であたくしはずっと一日中ねあのアナウンサールームでもって待機をしていましたよ
81. それでまあ、わたくし、ちょうどデビュー４０年っていうこともあって、ショパンのプログラムでリサイタルをしたんですけども、
82. 僕も必ず仕事場にきていると今日ママだっきって言ったかってメール打つんです。
83. その三人子供が並んでお父様に手を振った。
84. 私はね、三十八ぐらいの時ですかね、一年間仕事休んで
85. あの方９１、２で、あの３０いくつの女性と駆け落ちやりまして、
86. お父様は真紀子さんとうちで議論なさって、
87. わたしなんかもう起きてからすぐそうだって感じではなししてますし、
88. あなた一年中のほとんどをコンサートで日本中歩く。
89. 女子学生が昼間から不倫してる人のドラマをビデオで録画して夜みている。
90. 大詰めでね琵琶湖湖畔でもって木下藤吉郎が非人達にねお金を恵むんですよ
91. ぼくはそこでジャズコンサートでショウアップするのに椅子持って行く。
92. 僕のおやじ僕が子供の頃ですね毎朝食パンに山盛りの砂糖のっけてね、

≥3NPs Tokens in OS order

1. で、その馬券をみんなで一万づつ入れて、
2. で、あの人のお芝居を僕は裏方を手伝ったことがありますけども
3. 奥様のことをお子さん達はママと呼び、
4. それを本人もまずいと思ったらしくて、
5. それは生徒はみんなそう言う。
6. それを僕はそこで読んでて
7. それをですね中村寛三郎が楽屋で聞いていた
8. これは作家の小林のぶひこさんがですへん、あの、なんかで書いたんですよね、
9. そ、お父様のことは随分あなたいろんなものに書いてらっしゃるんですけど、
10. それをわたしが小学生の低学年の時に読んで、
11. テレビを見ながら食事っていうことはわたしも子供が小さいときからほとんどやっておりません。
12. それを全部彼はエネルギーの火達磨にして
13. そういう意味でほんとにスポーツマンかっての自分で疑問に思っています。
14. マイナーな登場人物もそれなりにその作家が意味があって必要性があって書いてますでしょ、
15. なんやっての大阪弁で、あなたが言ってらっさるんですけど、
16. あの、酒と涙と男と女という曲をもうずいぶん前にご自分でおつくりになりまして
17. その中にあなたのことを淀川先生がとってもほめてらっしゃるんのね。
18. そこは、そこを、あ、象が歩いてるとか、
19. 初めはお父様のだから鶴田ってお名前あなた名乗っていらっしゃらないんですよね
20. これどういうのでこれみんなで写したの
21. 先ほど、本の題名をわたくし、羊の歌のこと羊の群れって言ってたようなんですけど
22. 今の状態は食事はその相棒っていうのかな、一緒にすんでる人が、あの、女の人が作ってくれるんですね。

References

Aoyagi, Hiroshi. 1998. *On the Nature of Particles in Japanese and Its Theoretical Implications*. University of Southern California: Doctoral dissertation.

Arnold, Jennifer E. 1998. Making salience: the similarity of topic and focus. Ms. University of Pennsylvania.

Arnold, Jennifer E., Wasow Thomas, Anthony Losongco, and Ryan Ginstrom. 2000. Heaviness vs. newness. *Language* 76:28-55.

Blutner, Reinhard, Helen de Hoop, and Petra Hendriks. 2004 (pre-final version). *Optimal Communication*. Stanford: CSLI Publications.

Bock, J. K. 1995. Sentence production: from mind to mouth. In *Speech, Language, and Communication: Handbook of Perception and Cognition*, eds. Joanne L. Eimas and Peter D. Miller, 181-216. San Diego: Academic Press.

Bock, J. K. and W. J. M. Levelt. 1994. Language production: grammatical encoding. In *Handbook of Psycholinguistics*, ed. M. A. Gernsbacher, 945-984. San Diego: Academic Press.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In *Subject and Topic*, ed. Charles Li. New York: Academic Press.

Chafe, Wallace. 1987. Cognitive constraints on information flow. In *Coherence and Grounding in Discourse and Grammar*, ed. Russell Tomlin, 21-51. Amsterdam: John Benjamins.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford: CSLI Publications.

Clancy, Patricia M. 1982. Written and spoken style in Japanese narratives. In *Spoken and Written Language*, ed. Deborah Tannen, 55-76. Norwood, NJ: Ablex.

Clancy, Patricia M. and Pamela Downing. 1987. The use of *wa* as a cohesion marker in Japanese oral narratives. In *Perspectives on Topicalization. The Case of Japanese 'Wa'*, eds. John Hinds, Senko K. Maynard and Shoichi Iwasaki, 3-56. Amsterdam: John Benjamins.

Dryer, Matthew S. 1980. The positional tendencies of sentential noun phrases in universal grammar. *The Canadian Journal of Linguistics* 25:123-195.

Dryer, Matthew S. 1995. Frequency and pragmatically unmarked word order. In *Word Order in Discourse*, eds. Pamela Downing and Michael Noonan, 105-135. Philadelphia: John Benjamins.

Dryer, Matthew S. 1996. Focus, pragmatic presupposition, and activated proposition. *Journal of Pragmatics* 26:475-523.

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

Ferreira, Victor S. and Hiromi Yoshita. 2003. Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research* 32:669-692.

Firbas, Jan. 1964. On defining the theme in functional sentence analysis. *Travaux linguistique de Prague* 1.

Firbas, Jan. 1966. Non-thematic subjects in contemporary English. *Travaux Linguistiques de Prague* 2:239-254.

Foley, Williams and Robert Van Valin Jr. 1985. Information packaging in the clause. In *Language Typology and Syntactic Description*, ed. Timothy Shopen. Cambridge: Cambridge University Press.

Fujii, Noriko and Tsuyoshi Ono. 2000. The occurrence and non-occurrence of the Japanese direct object marker *o* in conversation. *Studies in Language* 24:1-39.

Fujii, Yoko. 1991. Reversed word order in Japanese - a discourse-pragmatic analysis. *Gengo Kenkyu* 99.

Givón, Talmy. 1975. Focus and the scope of assertation: some Bantu evidence. *Studies in African Linguistics* 6:185-205.

Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In *Topic Continuity in Discourse: A Quantitative Cross-Language Studies*, ed. Talmy Givón, 4-41. Amsterdam: John Benjamins.

Givón, Talmy. 1988. The pragmatics of word-order: predictability, importance and attention. In *Studies in Syntactic Typology*, eds. Michael Hammond, E. A. Moravcsik and J. R. Wirth, 243-284. Amsterdam: John Benjamins.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, ed. Joseph H. Greenberg, 79-113. Cambridge, Mass: MIT Press.

Greenberg, Joseph H. 1966. *Language Universals (with special reference to feature hierarchies)*. The Hague: Mouton.

Grimshaw, Jane and Armin Mester. 1988. Light verbs and ø-marking. *Linguistic Inquiry* 19:205-232.

Haig, John H. 1996. Subjacency and Japanese grammar: a functional account. *Studies in Language* 20:53-92.

Hakuta, Kenji. 1982. Interaction between particles and word order in the comprehension and production of simple sentences in Japanese children. *Developmental Psychology* 18:62-76.

Halliday, M. A. K. 1967. Note on transitivity and theme in English, Part 2. *Journal of Linguistics* 3:199-244.

Hasegawa, Yoko. 1992. *A Study of Japanese Clause Linkage: The Connective Te in Japanese.* Stanford: CSLI Publications.

Hawkins, John A. 1992. Syntactic weight versus information structure in word order variation. In *Informationsstruktur Und Grammatik, Linguistische Berichte Special Issue*, ed. J. Jacobs, 196-219.

Hawkins, John A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammars*: Oxford University Press.

Herring, Susan C. and J. C. Paolillo. 1995. Focus position in SOV languages. In *Word Order in Discourse*, eds. Pamela Downing and Michael Noonan. Amsterdam: John Benjamins.

Hinds, John. 1976. *Aspect of Japanese Discourse Structure*. Tokyo: Kaitakusha Publishers.

Hinds, John. 1983. Topic continuity in Japanese. In *Topic Continuity in Discourse: A Quantitative Cross-Language Study*, ed. Talmy Givón. Amsterdam: John Benjamins.

Iijima, Koichi. 1997. Ame nimo makezu, kaze nimo makezu (Against the rain, against the wind). *Gengo* 26:90-96.

Kameyama, Megumi. 1995. The syntax and semantics of the Japanese language engine. In *Japanese Sentence Processing*, eds. Reiko Mazuka and Noriko Nagai, 153-176. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Kim, Alan Hyun-Oak. 1988. Preverbal focusing and type XXIII languages. In *Studies in Syntactic Typology*, eds. Michael Hammond, E. A. Moravcsik and J. R. Wirth. Amsterdam: John Benjamins.

Kitahara, Yasuo. 1970. Jyodooshi no soogo shoosetu ni tuite no koobunteki koosatsu "on Referentiality of Auxiliary Verbs". *Kokugogaku* 83.

Kuno, Susumu. 1972. Functional sentence perspective: a case study from Japanese and English. *Linguistic Inquiry* 3:269-320.

Kuno, Susumu. 1973. Constraints on internal clauses and sentential subjects. *Linguistic Inquiry*:363-385.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy - a reexamination of relativization phenomena. In *Subject and Topic*, ed. Charles Li, 417-444. New York: Academic Press.

Kuno, Susumu. 1978. *Danwa No Bunpoo 'Grammar of Discourse'*. Tokyo: Taishuukan.

Kuroda, S.-Y. 1965. *Generative Grammatical Studies in the Japanese Language*. Massachusetts Institute of Technology: Doctoral dissertation.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: A Theory of Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Lehmann, Winfred P. 1988. Towards a typology of clause linkage. In *Clause Combining in Grammar and Discourse*, eds. John Haiman and Sandra A. Thompson, 191-225. Amsterdam: John Benjamins.

Levelt, Willem J. M. 1989. *Speaking: From Intention To Articulation*. Cambridge, Mass: MIT Press.

Li, Charles. 1976. Subject and topic: A new typology of language. In *Subject and Topic*, ed. Charles Li, 459-488. New York: Academic Press.

Martin, S. E. 1975. *A Reference Grammar of Japanese*. New Haven: Yale University Press.

Mikami, Akira. 1960. *Zoo Wa Hana Ga Nagai 'Elephant-wa Nose-ga Long'*. Tokyo: Kurosio Publisher.

Mikami, Akira. 1969. Sonzaibun no mondai 'On existential sentences'. *Otani-Zyosidaigaku Kiyoo* 3:38-48.

Mitsuya, Shigematsu. 1908. *Kootoo Nohongo Bunpoo 'Advanced Japanese Grammar'*. Tokyo: Meiji Shoin.

Miyagawa, Shigeru. 1987. Wa and the WH phrase. In *Perspectives on Topicalization: The Case of Japanese Wa*, eds. John Hinds, Senko K. Maynard and Shoichi Iwasaki, 185-217. Amsterdam: John Benjamins.

Miyajima, Tatsuo. 1963. Kakari no ichi 'The Position of *kakari'*. *Keiryoo Kokugogaku* 23.

Miyajima, Tatsuo. 1964. Gendai zasshi 90-shu no yoogo yooji 'The use of words and letters in 90 modern magazines'. In *Kokuritu Kokugo Kenkyuusho Hookoku 25*: National Language Research Institute.

Myhill, John. 1992. *Typological Discourse Analysis: Quantitative Approaches to the Study of Linguistic Function*. Cambridge: Blackwell Publishers.

Nichols, Johanna. 1986. Head marking and dependent marking grammar. *Language* 62.

Ohori, Toshio. 1992. *Diachrony in Clause Linkage and Related Issues*. University of California at Berkeley: Doctoral dissertation.

Payne, Doris L. 1985. *Aspects of the Grammar of Yagua*. University of California at Los Angeles: Doctoral dissertation.

Saeki, Tetsuo. 1960. Gendaibun ni okeru gojyun no keikoo 'The tendency of word order in modern Japanese sentences'. *Gengo Seikatu* 12.

Saeki, Tetsuo. 1975. *Gendai Nihongo no Gojun 'Word Order in Modern Japanese'*. Tokyo: Kasama Shoin.

Saeki, Tetsuo. 1976. *Gojun to Bunpoo 'Word Order and Grammar'*. Osaka: Kansai University Press.

Saeki, Tetsuo. 1998. *Yoosetsu Nihonbun no Gojun 'Summary on Word Order in Japanese*'. Tokyo: Kurosio Publisher.

Sakakura, Atsuyoshi. 1966. *Gokoosei no Kenkyuu 'Study on Word Structure'*. Tokyo: Kadokawa Shoten.

Shibamoto, Janet. 1985. *Japanese Women's Language*. New York: Academic Press.

Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.

Shimojo, Mitsuaki. 1995. *Focus Structure and Morphosyntax in Japanese: Wa and Ga, and Word Order Flexibility*. State University of New York at Buffalo: Doctoral dissertation.

Shimojo, Mitsuaki. 2005. *Argument Encoding in Japanese Conversation*. Hampshire and New York: Palgrave Macmillan.

Simon, Mutsuko E. 1989. *An Analysis of the Postposing Construction in Japanese*. University of Michigan: Doctoral dissertation.

Suzuki, Satoko. 1995. The functions of topic-encoding zero-marked phrases: a study of the interaction among topic-encoding expressions in Japanese. *Journal of Pragmatics* 23:607-626.

Testelec, Yakov G. 1998. Word order in Daghestanian languages. In *Constituent Order in the Languages of Europe*, ed. Anna Siewierska, 257-280. Berlin: Mouton de Gruyer.

Tomlin, Russell S. 1995. Focus attention, voice and word order; an experimental, cross-linguistic study. In *Word Order in Discourse*, eds. Pamela Downing and Michael Noonan. Amsterdam: John Benjamins.

Tomlin, Russell S. and Rhodes Richard A. 1979. An introduction to information distribution in Ojibwa. *Papers from the Fifteenth Regional Meeting of the Chicago Linguistics Society*, 307-321. Reprinted in, *The Pragmatics of Word Order Flexibility* ed. by Doris Payne. Amsterdam: John Benjamins.

Uehara, Satoshi. 1998. *Syntactic Categories in Japanese: A Cognitive and Typological Introduction.* Tokyo: Kurosio Publisher.

Ueno, Mieko and Robert Kluender. 2003. Event-related brain indices of Japanese scrambling. *Brain and Language* 86:243-271.

Van Valin, Robert D., Jr. 1984. A typology of syntactic relations in clause linkage. *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society,* 542-558. Berkeley: Berkeley Linguistics Society.

Van Valin, Robert D., Jr. 1987. Recent development in Role and Reference Grammar: the layered structure of the clause and juncture. *Davis Working Papers in Linguistics*. University of California at Davis.

Van Valin, Robert D., Jr. 1993. A synopsis of Role and Reference Grammar. In *Advances in Role and Reference Grammar*, ed. Robert D. Van Valin, Jr, 1-164. Amsterdam: John Benjamins.

Van Valin, Robert D., Jr and Randy LaPolla. 1997. *Syntax: Structure, Meaning and Function.* Cambridge: Cambridge University Press.

Vilkuna, Maria. 1998. Word order in European Uralic. In *Constituent Order in the Languages of Europe*, ed. Anna Siewierska, 173-233. Berlin: Mouton de Gruyter.

Yamashita, Hiroko. 2002. Scrambled sentences in Japanese: linguistic properties and motivations for production. *TEXT* 22:597-633.

Yamashita, Hiroko and Franklin Chang. 2001. "Long before short" preference in the production of a head-final language. *Cognition* 81:B45-B55.