FOCUS STRUCTURE AND MORPHOSYNTAX IN JAPANESE: *WA* AND *GA*, AND WORD ORDER FLEXIBILITY

by

MITSUAKI SHIMOJO

August 1, 1995

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

1995

TABLE OF CONTENTS

NOTES ON TRANSCRIPTIONS

Japanese transliteration in given using romanization known as *Kunrei-siki* 'Official System', except for the following changes: *Chi* is used for *ti*, *tsu* for *tu*, *sha*, *shu*, *sho* for *sya*, *syu*, *syo* respectively, and *cha*, *chu*, *cho* for *tya*, *tyu*, *tyo* respectively. *Ti* is used instead of *chi* in foreign loan words since it is not palatalized. Long vowels are indicated by two vowel symbols in a row, instead of using a circumflex.

Proper nouns in the made-up examples are spelled as in English and not given in the English gloss. For proper nouns in the conversation data, I used either pseudonyms or random initials.

In citing example sentences throughout this thesis, I may change the Romanization and the English gloss for grammatical morphemes in order to keep the orthography and the grammatical labels consistent; otherwise, I do not make any change in the original, including the grammaticality judgment by the author.

The post-nominal particles *wa* and *ga* are not glossed in the sentence examples since their functions are discussed throughout this thesis.

A sentence-final question mark [?] in sentence examples indicates a rising interrogative contour, which makes the sentence a question without the sentence-final question particle *ka*. In the transcription of the conversation data, a slash [/] indicates a recognizable pause between sentence elements. However, the clause unit boundary for the purpose of the Referential Distance measurement is indicated by the numbered line break, not by the slash, if there are more than one clause unit in the example. Speakers in conversation are identified by capitalized letters, e.g. A, B, C..., if there are more than one speaker in the conversational context. The order of utterances is shown by the numerals, e.g. 1, 2, 3...; therefore, A1 indicates the first utterance (unit) of speaker A and A2 the second utterance of the same speaker, i.e. speaker A. Variants of a single sentence are indicated by small letters, e.g. a, b, c...

An asterisk [*] indicates that the sentence is clearly ungrammatical, and the symbol [#] shows that the sentence is grammatical but it is clearly inappropriate in the discourse context in question. A question mark [?] is used to indicate the awkwardness of the sentence due to either grammatical constraints or discourse-pragmatic constraints.

Following abbreviations are used throughout this thesis.

ADJ	Adjective	NPM	No Previous Mention
ADV	Adverb	NUC	Nucleus
AFD	Actual Focus Domain	OBJ	Object/Object Marker
AP	Adjectival Phrase	PAS	Passive
ARG	Argument	PFD	Potential Focus Domain
CL	Numeral Classifier	PM	Previously Mentioned
CMPL	Complementizer	PoCS	Postcore Slot
COND	Conditional	POT	Potential
CONJ	Conjunction	PP	Prepositional/Postpositional Phrase
CONS	Concessive	PrCS	Precore Slot
COP	Copula	PRED	Predicate
DAT	Dative/Dative Marker	PrFD	Primary Focus Domain
DEF	Definite	PRO	Pronoun
FP	Final Particle	PST	Past Tense/Perfective Aspect
GEN	Genitive	Q	Question Marker
HON	Honorific	QT	Quotative Marker
HOR	Hortative	RD	Referential Distance
IF	Illocutionary Force	RDP	Right-detached Position
IMP	Imperative	RE	Resultative
LDP	Left-detached Position	REF	Referring Element
LS	Logical Structure	S	Sentence
NEG	Negative	SFD	Secondary Focus Domain
N	Noun	SUB	Subject
NOM	Nominalizer	TNT	Tentative
NP	Noun Phrase	V	Verb

ACKNOWLEDGMENTS

First and foremost, I wish to thank Robert D. Van Valin, Jr., who has served as chair of my committee, for his insightful advice and constant encouragement at every stage of this dissertation. Without his patience and encouragement on my studies, I would never have finished this dissertation.

I also wish to thank Matthew S. Dryer, who willingly agreed to serve on my committee and has been enthusiastic in his support of my studies. I must acknowledge my debt to Matthew particularly because his recent study on *focus* was largely the inspiration for this dissertation. I am also grateful to David A. Zubin for not only his serving on my committee but also his constant support in a wide variety of forms throughout. In particular, I owe David to a great extent for his providing direction on my quantitative studies ever since my master's project.

Throughout my graduate career in Buffalo, I have benefited from the expertise of all the faculty at the Department of Linguistics at SUNY at Buffalo, in particular those who patiently taught me the basics of linguistics: Jeri J. Jaeger, Madeleine Mathiot, Karin E. Michelson, Leonard Talmy, David P. Wilkins, and Wolfgang W÷lck. Without their intellectual and moral encouragement, I would not have continued my study for a Ph.D.

I would also like to thank my fellow students for their support and input, particularly, Isao Honda, David Houghton, David Kemmerer, Tomiko Kodama, Craig Kopris, Wataru Nakamura, Sang-ryong Oh, Ki-seong Park, Kazue Watanabe, Noriko Watanabe, Heather Weber, Yumi Yamasaki, and Byong-seon Yang.

I am greatly indebted to the Department of Linguistics and the World Languages Institute at SUNY at Buffalo for their financial support for a considerable part of my graduate career. In particular, I wish to thank Mark A. Ashwill, Stefanos Papazaharias, Kazue Masuyama, Sachiko Hiramatsu, Akemi Sakamoto, Setsue Shibata, and Akiko Sugiyama for their constant support in my teaching Japanese. In addition, I wish to thank

my students, particularly Fred Eyer and Susanna Mark in the upper-level Japanese courses, for their indulgence as well as their inspiring questions, which greatly motivated me to pursue my studies on the Japanese language.

I would also like to express my thanks to the members of the Chancel Choir at Central Presbyterian Church of Buffalo, particularly Roy W. Clare, Benjamin and Carol Munson, and Jack Anthony, for their constant fellowship and moral support since the very beginning of my stay in Buffalo. Last but not least, I wish to thank my family for their understanding, patience, and constant prayer. Without their care, the completion of this dissertation would have been unthinkable.

ABSTRACT

Truth-conditionally equivalent sentences may be different in *meaning* due to different ways of recognizing more informative and less informative parts of a sentence. This study defines the informative part of a sentence as *focus* and less- or non-informative part as *nonfocus*. The focus-nonfocus contrast is in turn defined in terms of the notion *activation*; the focus of a sentence involves elements corresponding to the least activated referent in the addressee's consciousness in the immediately preceding context of the utterance. Given this background, this study investigates the post-nominal particles *wa* and *ga*, so-called *topic marker* and *subject marker* respectively, and non-canonical word order, particularly the postposing construction, in Japanese. These seemingly disparate phenomena in morphology and syntax are related to each other in that their functions are closely connected with the distribution of focus.

The presentation of the research takes on the following organization. Chapter 1 introduces the focus theory in connection with the notion of activation and discusses a continuum for degree of activation. Chapter 2 investigates the use of *wa* and *ga* in mini-discourses, primarily question-answer pairs, and proposes a contrast that *wa* marks nonfocus and combines with the open proposition which is the focus of the sentence, while *ga* can manifest any of the other possible focus patterns. Chapter 3 investigates conversational Japanese with Giv n's referential distance [RD] measurement, extending the use of the method by measuring the RD for propositions, and further supports the principle proposed in Chapter 2. Chapter 4 outlines possible non-canonical word order in Japanese and applies the focus theory to postposing. I propose a principle that the postposing is acceptable if and only if the element in postverbal position is less important than the preceding elements, which is further supported by the results from the RD measurement in conversational Japanese. Based on the findings in the previous chapters, Chapters 5 and 6 examine *wa*/*ga* and postposing in Role and Reference Grammar,

particularly focusing on the use of *wa*/*ga* in embedded clauses and postposing of and out of embedded clauses.

CHAPTER 1

INTRODUCTION

1.1. OBJECTIVES AND OUTLINE OF THE THESIS

The thesis will investigate the interaction between focus structure and two morphosyntactic phenomena in Japanese: first, the well-known post-nominal particles *wa* and *ga*, which are so-called *topic marker* and *subject marker* respectively, and second, non-canonical word order, particularly focusing on the postposing construction. I will claim that these seemingly disparate phenomena on two different levels, morphology and syntax, are related to each other in that their functions are largely motivated by information dynamics. The investigation of functions of *wa* and *ga* is an indispensable part of a study of the focus structure on the sentence level, and therefore, a functional account of the system of word order variation at least requires the provision of a solid framework to describe the functions of *wa* and *ga*.

The thesis will consist of six chapters, excluding the final chapter, which summarizes the discussion throughout the thesis. The organization of the discussion will be as follows. Chapter 2 investigates the use of *wa* and *ga* in mini-discourses, primarily question-answer pairs, in order to capture the functional contrast between the two on the basis of the notion *activation*, which will be discussed in the next section of the present chapter. Chapter 2 is concerned with the following:

A. Contrasts between *wa* and *ga* in terms of degree of activation.

B. Treatment of functional variations of *wa* and *ga* proposed in previous studies, e.g. topic *wa* vs contrastive *wa*, exhaustive listing *ga* vs neutral description *ga* as proposed by Kuno (1973).

C. Activation pattern of sentences with a *wa*-marked WH phrase.

The theoretical framework employed for defining the focus structure throughout this thesis is based on the view that non-focus is associated with activation, not pragmatic

presupposition, a view which has been most explicitly defended by Dryer (1994). (The same type of claim has been made in Chafe (1976) and Myhill (1992).)

Chapter 3 will further examine and evaluate the principles proposed for the functional contrast between *wa* and *ga* on the basis of Japanese conversation data. In this chapter, I use a quantitative methodology, more specifically the measurement method for *referential distance* [RD] (cf. Giv n (ed.) 1983) to measure the degree of activation, hence the degree of focus, of referents and propositions. The analysis here will be based on the assumption that a referent or proposition of recent mention, i.e. of small RD, is more active in individuals' consciousness than a referent or proposition which has not been mentioned recently or not been mentioned at all in the preceding discourse context, i.e. a referent or proposition of large RD. In addition to the conventional RD measurement method, which measures the distance between a referent and the most recent coreferential expression, this thesis will extend the use of the measurement method by measuring the distance between a proposition and the most recent semantically-identical proposition in the previous discourse. The quantitative methodology will be discussed at length. Particularly, this chapter will be concerned with quantitative description of the following.

A. Activation patterns of sentences with *wa*/*ga*.

B. The use of *wa*/*ga* in subordinate clauses in conversational Japanese.

In Chapter 4, I will extend the analysis to the syntactic level, specifically non-canonical word order, particularly focusing on postposing constructions. This chapter will attempt to describe the following concerns:

A. Description of possible non-canonical word order in Japanese.

B. Restriction on non-canonical word order and postposing.

C. Representation of postposing constructions in terms of focus structure.

First, I will describe the types of possible non-canonical word order and extend the discussion particularly to the postposing construction. Next, I will further examine the

postposing construction to capture the principles involved in terms of focus structure. The discussion will be given in terms of both mini-discourse data and the quantitative analysis of the same conversational Japanese database as used in Chapter 3.

Primarily on the basis of the discussion in the preceding chapters, Chapters 5 and 6 will investigate the same morphosyntactic phenomena in the framework of Role and Reference Grammar [RRG] (cf. Foley and Van Valin 1984, Van Valin 1993a), along with the notion of activation. Chapter 5 focuses on *wa* and *ga* and Chapter 6 the word order variation. Chapter 5 will be concerned with the following:

A. Representation of *wa* and *ga* in the constituent projection.

B. Representation of *wa* and *ga* in the focus structure projection.

C. The use of *wa*/*ga* in complex sentences.

First, the discussion will be given in terms of the layered structure of the clause in RRG and its recent formulation of focus domain. Also, the analysis will be extended to the cases of *wa* marking non-NPs, e.g. *wa*-marked predicate and adverbs, in order to examine how those *wa*-marked non-NPs are represented in RRG. After laying out the basic focus structure of *wa*/*ga* sentences, I will discuss the use of *wa*/*ga* in complex sentences in connection with degree of matrixhood of subordinate clause in order to account for the quantitative data from the previous chapter.

Chapter 6 will extend the RRG analysis to non-canonical word order. This chapter is concerned with the following:

A. Representation of non-canonical word order in simple and complex sentences.

B. Representation of postposing constructions in terms of both the constituent projection and the focus structure projection.

C. Constraints on postposing and their RRG account.

First, I will lay out the basic constraints on non-canonical word order in general in terms of both the RRG constituent structure and the focus structure. Then, I will focus on

postposing constructions, showing how they are represented and how the basic constraints on postposing are accounted for in RRG. Finally, I will investigate constraints on extraction of and out of a subordinate clause and propose principles from two different angles, the structural point of view and the discourse-pragmatics point of view. Before I discuss the morphosyntactic issues in Japanese as outlined above, I will lay out in the following section the theoretical background for the notions of activation and focus which are used throughout this thesis and will be combined with the RRG framework later in this thesis.

1.2. ACTIVATION AND FOCUS: THEORETICAL BACKGROUND

It has long been noted in the literature that truth-conditionally equivalent sentences are not necessarily the same in meaning. For example, the sentences in (1) express the same semantic proposition; however, they are different in *meaning*,[1] due to the placement of focal accent.

(1)
a. HE met Mary.
b. He met MARY.

The difference in meaning expressed by sentence pairs as in (1) has normally been captured by recognition of informative and noninformative parts of a sentence; the constituent with focal accent is more informative than the constituent without such accent. However, different researchers use different notions to define the difference between the more informative part and the less informative part. For example, the former is *focus* and the latter is *presupposition* in Chomsky (1971) and Jackendoff (1972). For Prince (1981a, 1986), the former is focus, while the latter is not only presupposition but also *shared knowledge*. For Halliday (1967), Chafe (1976), and Clark and Haviland (1977), the former is *new* and the latter is *given*.

Against the background of using different terminology, a further complication arises because same terminology is used to refer to distinct notions. For example, Prince (1979) mentions three types of *givenness* which are found in the literature: (1) *predictability/recoverability*, (2) *shared knowledge*, and (3) *saliency*. The notion of predictability/recoverability is found in Kuno (1972, 1973), i.e. given information is the information which is predictable or recoverable by the preceding context. The notion of shared knowledge is typically treated as the same as *presupposition*, and this type of notion is found in works such as Clark and Haviland (1977), Prince (1981a, 1986), Gundel (1985), and Lambrecht (1986, 1992, 1994). For example, Lambrecht defines pragmatic presupposition as "set of propositions lexicogrammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered" (1994: 52). The notion of saliency is represented in the work of Chafe (1974, 1976, 1987, 1994). According to Chafe (1994: 72), for example, information is *new* if it is newly activated in the addressee's consciousness at the point of utterance in the conversation, while *given* information is already active in the addressee's consciousness at the point of utterance.

Chafe's characterization of *new* and *given* above provides the basic assumption behind the notion of *activation*, which will be used throughout this thesis. Saliency is concerned with information which gets activated or deactivated in the individual's consciousness. Chafe (1987: 22) states that "our minds contain very large amounts of knowledge or information, and that only a very small amount of this information can be focused on, or be 'active' at any one time." On the basis of this assumption, Chafe (1987: 25) gives three states of activation: *active concepts*, *semi-active concepts*, and *inactive concepts*. An active concept is one which is currently focused in a person's consciousness, an inactive concept is one which is currently not focused at all (it is only in a person's long-term memory), and a semi-active concept is in the middle of the scale; it is in a person's peripheral consciousness.

Dryer (1994: 6) makes a similar distinction to that of Chafe; however, he further elaborates the notion by recognizing a continuum for degree of activation with four different phases: *focus of attention*, *activated*, *semi-deactivated*, and *nonactivated*. An advantage of a continuum rather than discrete stages as in Chafe's is that we can recognize information which is even more activated, i.e. focus of attention, among the activated ones, and these relative degrees of activation will turn out to be crucial for investigation of Japanese in later chapters. In order to avoid the terminological confusion of the term *focus*, the present study uses the term *center of attention* to refer to the state *focus of attention* in Dryer's (1994) sense.

In addition to *semi-deactivated* on the intermediate level of activation, Dryer (1994) proposes a second type, *accessible*. Accessible entities are different from semi-deactivated entities in that the former becomes semi-activated by some related entity being activated. Dryer (1994: 5) describes this process by stating "the activation of one entity may cause entities which are related to the first entity by inference or other types of association to be highly accessible to activation in the sense that they may be just below the threshold of activation, and thus not activated, but still quite distinct from entities which are fully nonactivated". To illustrate the claim above, Dryer (1994: 5) cites the following.

(2) John came into the room with a woman we had never met. We wondered where his wife was.

In (2), the preceding context makes the referent 'his wife' more accessible to activation for the hearer than 'his wife' without a related preceding context, for example, 'his wife' mentioned out of the blue.

Figure 1 summarizes the continuum of activation, with referents which are center of attention being the highest in activation and nonactivated referents being the lowest in activation, and the present study assumes this continuum as the basis for discussion throughout this study. Finally, while referring to a referent or a proposition causes it to

become activated in individual's consciousness, a referent or a proposition gradually decays in activation if it does not continue to be referred to (Dryer 1994). This is plausible if we assume that only a very small amount of information can be active at a time while new information keeps being activated.

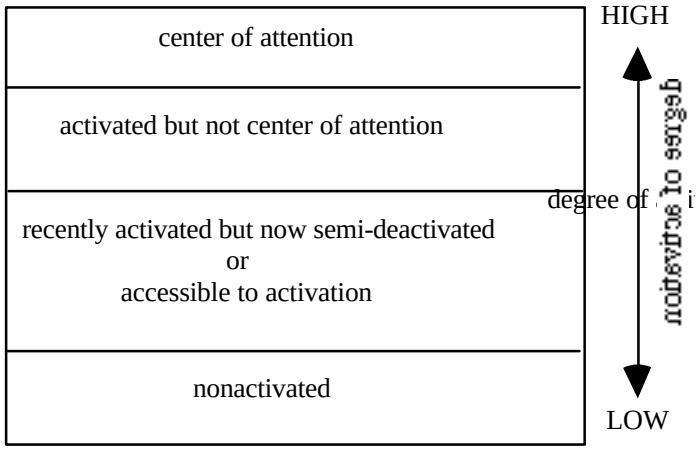


Figure 1: Continuum of activation

Given the association of givenness with activated information, the present study follows Dryer (1994), who explicitly claims that non-focus is associated with activation, not pragmatic presupposition.[2] Namely, the non-focused parts of an utterance correspond to activated information and focused parts to information that is nonactivated or less activated in the immediately preceding context. Obviously, this position rejects the prevailing claim that non-focus is associated with pragmatic presupposition or shared knowledge, as seen above. Dryer's claim above that nonfocus involves activation instead of presupposition comes primarily from analyses of question-answer pairs in English. Below I will briefly summarize his argument for the claim above with examples from Dryer (1994).

The classic examples used in the literature to illustrate the association between focus and pragmatic presupposition are WH-question and answer pairs as in (3).

(3)
A: Who saw John?

B: MARY saw John.

In (3A), 'Mary' is the focus and *saw John* is the nonfocus and it is the case that there is presupposition that someone saw John, which corresponds with the nonfocus of the sentence. However, since WH-questions normally presuppose the existence of an entity which will be the answer to the question, i.e. *someone* in someone saw John in (3A), it is not clear whether this presupposition is due to the focus-nonfocus structure of the sentence in (3B) or due to the presupposition caused by the WH-question in (3A). More useful examples to examine this point are WH-question and answer pairs in which challenges the possible presupposition.

(4)
A: Who saw John?
B: NOBODY saw John.

(5)
A: Who if anyone saw John?
B: MARY saw John.

The distribution of focus and nonfocus is the same in (4B) and (5B) as it is in (3B); the focal accent shows the focus and the rest of the sentences is the nonfocus. However, (4) and (5) are different from (3) in that the sentence in (4B) and (5A) does not presuppose that someone saw John. In (4), the answer asserts exactly the opposite, and in (5) the WH-question itself cancels the presupposition by the expression *if anyone*. The examples as in (4) and (5) indicate that pragmatic presupposition is not necessarily associated with nonfocus. On the other hand, the examples above have in common that the focus involves the activated proposition that someone saw John. Whether the proposition is presupposed or not, it is activated by being mentioned in the immediately preceding context.

To further support the claim above, Dryer (1994: 12) cites negative sentences, questions, and conditionals as well, as follows.

(6)
A: Did anyone see John?
B: I don't know. I know MARY didn't see him.

(7)
A: John thinks that Bill is in the house, but I know that he isn't.
B: Is SAM in the house?
A: NOBODY is in the house.

(8) I don't know whether anyone saw John, but if MARY saw him/John, I will be very angry.

The second sentence in (6B) does not pragmatically presuppose that someone saw John since the question is uttered to find out whether it is the case. The question in (7B) does not presuppose that someone is in the house, and the sentence in (8), due to the conditional, does not presuppose the proposition that someone saw John. In all the examples above, the placement of nonfocus is independent of the speaker's belief of the proposition involved, and they support the claim that pragmatic presupposition is not a necessary property of nonfocus. On the other hand, in all the examples above, it is the case that nonfocus is associated with activated proposition which is mentioned in the immediately preceding context.

A crucial set of data which supports the claim that nonfocus involves activated propositions rather than pragmatic presupposition is the following kind, which is cited by Dryer (1994: 13).

(9)
A: Is Gore the President and Clinton the Vice-President?
B: No, CLINTON is the President; (Gore is the VICE-PRESIDENT.)

(10)
A: Is Gore the President and Clinton the Vice-President?
B: No, Clinton is the PRESIDENT; (GORE is the Vice-President.)

The crucial point here is that the (9B) and (10B) have exactly the same presuppositions and the assertions, yet the focus-nonfocus patterns are different; 'Clinton' is the focus in (9B), while 'President' is the focus in (10B) (ignoring the second part of (9B) and (10B)). The difference in the place of focus is reflected by the fact that (9B) is the response to the first question in (9A) and (10B) is the response to the second question in (10A). As clear

by the discussion above, the characterization of the sentences in (9B) and (10B) in terms of pragmatic presupposition fails to account for the placement of focus and nonfocus here.

The examples such as the ones in (9) and (10) may appear to be a problem for the framework of activation since the elements of focus are activated, being mentioned in the immediately preceding context in (9) and (10). However, the difference between focus and nonfocus here is accounted for by different degrees of activation. Dryer (1994: 14) notes "things that are activated may differ in degree of activation in that some things may be more highly activated in the sense that the individual's attention is focused on them." In the first response in (9B), the proposition *X is the president* is more highly activated than 'Clinton', by placing the question whether Gore is the President, i.e. the first question in (9A), on the center of attention. Conversely, in (10B) B's center of attention is the second question in (10A), i.e. whether Clinton is the Vice-President; the proposition *Clinton is X* is more highly activated than the proposition *X is the President*.

One last type of examples which Dryer (1994: 15-16) gives to support the theory of focus in terms of activation is the kind in which focus falls within a part of an utterance which is presupposed. First, Dryer cites the example in (11) from Halliday (1967: 237).

(11)
A: Have you told John that the window got broken?
B: It was John who BROKE it.

In (11), the proposition *someone broke the window* is presupposed, and the part of the sentence corresponding to this presupposition receives the focal accent in (11B). The same pattern is found in non-cleft sentences, as in (12).

(12)
A: Have you told John that the window got broken?
B: John was the one who BROKE it.

As in (9) and (10), examples such as (11) and (12) are explained by different degrees of activation. In (11A) and (12A), the question makes the referent 'John' more activated than

the individual who broke the window; therefore, the focus falls on the less activated part of the sentence than the more activated part *John*.

I briefly outlined Dryer's (1994) argument for the theory of focus tied to the notion of activation, instead of pragmatic presupposition, and the present study will assume this claim and the analyses of some morphosyntactic phenomena of Japanese in the succeeding chapters will be based on this basic assumption. One point which I should emphasize here is, as Dryer (1994: 16) notes, the mapping between focus/nonfocus and the degree of activation; nonfocus necessarily corresponds to an activated proposition, while activated proposition does not necessarily corresponds to nonfocus. Although Dryer does not explicitly say so, it is also the case, as indicated by the discussion above, that the nonactivated proposition necessarily corresponds to focus, while focus does not necessarily correspond to the nonactivated proposition. In essence, the focus of a sentence involves the part of the sentence corresponding to the least activated element, and this relative degree of activation plays a crucial role to account for cases where the traditional focus theory based on pragmatic presupposition fails.

CHAPTER 2

WA AND *GA*: ACTIVATION AND FOCUS

2.1. INTRODUCTION

The postpositional particles *wa* and *ga* in Japanese, which are traditionally defined as *topic/theme* marker and *subject/nominative* marker respectively (Kuno 1973), have long been controversial in Japanese linguistics. This controversy is partially due to analyses of different levels (e.g. sentence level vs discourse level) and analyses on different notions (e.g. *given* vs *new*, *predictable* vs *unpredictable*, *theme* vs *non-theme*, *important* vs *unimportant*). The primary purpose of the present and the next chapter is investigation of *wa*, including both subject marking and non-subject marking, and *ga* on both on the sentence-pair level and on the discourse level. For the analyses of the two *wa* and *ga*, I will use the notion of *activation* and the theory of focus tied to activation that *non-focus* is associated with activation, as outlined in Section 1.2.

The discussion for this chapter will proceed as follows. Section 2.2 lays out the morphosyntactic characteristics of Japanese which have a bearing on *wa* and *ga*. Section 2.3 examines previous approaches to *wa* and *ga* in terms of both sentence-level analysis and discourse-level analysis. Section 2.4 will examine the function of *wa* and *ga* on the basis of the notion of activation in order to capture their fundamental functional distinction. This section first investigates the general use of the two, and secondly the use of *wa* with WH phrases, and finally evaluates the distinction between *thematic wa* and *contrastive wa*, which has been posited by Kuno (1972, 1973), and examines whether such distinction is captured in terms of their activation structures. Finally, the discussion of this chapter is summarized in Section 2.5.

2.2. MORPHOSYNTACTIC CHARACTERISTICS OF JAPANESE WITH REFERENCE TO *WA* AND *GA*

Typologically, Japanese is classified as an SOV (Subject-Object-Verb) language (e.g. Greenberg 1963), and it well conforms to the "dependent-head" pattern with all types of constituent (Shibatani 1990: 257). Japanese also has the "dependent-marking" characteristic (e.g. Nichols 1986), as manifested by postpositional particles which express nominal relations, as shown in (1).

(1) kinoo wa Hanako ga uchi de Taro to eega o mita
yesterday home at with movie OBJ see:PST
'(As for) yesterday, Hanako saw a movie with Taro at home.'

Those postpositional particles, as in (1), which indicate the semantic relationships of nominals are so-called case particles, which are called *kaku-zyoshi* in the traditional Japanese grammar, and *wa* and *ga* are among those of this type.[3] Unlike most other particles of this type, the two are often interchangeable without changing the truth-conditional value of a sentence; more specifically, *wa* can appear with elements which are marked by *ga*, but not vice versa. (2) shows an example in which *ga* and *wa* mark the subject of the sentence.

(2)
a. Hanako ga kuruma o katta
car OBJ buy:PST
'Hanako bought a car.'

b. Hanako wa kuruma o katta
'Hanako bought a car.'

Despite the noun phrase *Hanako* marked by different particles, (2a) and (2b) are truth-conditionally equivalent; the same conditions which must be met for the former to be true must be met for the latter to be true.

The same sort of interchangeability is found in clausal subjects, as in (3).

(3)
a. Hanako ga kuruma o katta no ga odoroki da

car OBJ buy:PST NOM surprise COP
'It was surprising that Hanako bought a car.'

b. Hanako ga kuruma o katta no wa odoroki da
car OBJ buy:PST NOM surprise COP

Although *ga* is canonically used to mark the nominative case, i.e. it marks the subject of intransitive and transitive clauses, *ga* is also used to mark the object of "stative verbals, that is, a handful of transitive verbs (such as *dekiru* 'be able to', *wakaru* 'understand', *iru* 'need'), all transitive adjectives (such as *hosii* 'wants' *tebetai* 'be anxious to eat') and all transitive nominal adjectives (such as *suki* 'be fond of', *nigate* 'be bad at')" (Kuno 1973: 55).[4] As in subject-marking *ga*, *wa* can replace *ga* which marks the object, without changing the truth-conditional values of a sentence.

(4)
a. sensee no setsumee de sono bun no imi ga wakatta
teacher GEN explanation with that sentence GEN meaning understand:PST
'(I) understood the meaning of the sentence with the teacher's explanation.'

b. sensee no setsumee de sono bun no imi wa wakatta
'(I) understood the meaning of the sentence with the teacher's explanation.'

As implied by the traditional labels *topic* particle and *nominative* particle for *wa* and *ga* respectively, *wa* marks a greater range of elements than *ga* does. Unlike *ga*, *wa* can mark a wide variety of elements, as illustrated below. For example, with the object of "nonstative" verbs:

(5) kono kuruma wa kyonen katta
this car last-year buy:PST
'(I) bought this car last year.'

With indirect objects:

(6) Hanako ni wa sono hon o ageta
DAT that book OBJ give:PST
'(I) gave the book to Hanako.'

With PPs:

(7) nihon de wa sushi o tabeta
Japan in OBJ eat:PST
'(I) ate sushi in Japan.'

(8) gakusee no toki ni wa amari tenisu o sinakatta
student GEN time at not-very tennis OBJ do:NEG:PST
'When (I) was a student, (I) didn't play tennis very often.'

(9) basu de wa daigaku ni konai
bus by university to come:NEG:PST
'(I) don't come to the university by bus.'

With adverbs:

(10) sonnani hayaku wa hasirenai
that fast run:POT:NEG
'(I) cannot run that fast.'

Wa can mark clausal elements as well, though the range of elements which *wa* can mark is limited compared with phrasal elements. For example, with temporal clauses:

(11) denwa o sita toki wa Hanako wa hon o yondeita
telephone OBJ do:PST when book OBJ reading:PST
'Hanako was reading a book when I called (her).'

(12) tegami o kaiteiru aida wa razio o kiiteita
letter OBJ writing while radio OBJ listening:PST
'(I) was listening to the radio while writing the letter.'

With comparative clauses:

(13) Hanako to iku yori wa Taro to itta hoo ga ii
with go than with go:PST way good
'It's better to go with Taro than with Hanako.'

With clausal complements:

(14) Ken wa koko ni kuru to wa itteita
here to come CMPL saying:PST
'Ken was saying that he would come here.'

It has also been observed that *wa* can mark predicates. For example, with verbs:

(15) baa ni itta kedo osake o nomi wa sinakatta
bar to go:PST but alcohol OBJ drink do:NEG:PST

'(I) went to a bar, but (I) didn't drink alcohol.'

With adjectives:

(16) sono mondai wa muzukasiku wa nakatta
that problem difficult NEG:PST
'The problem was not difficult (to solve).'

With the copula:

(17) Hanako wa sensee de wa nai
teacher COP NEG
'Hanako is not a teacher.'

Also, it is possible for *wa* to mark the *te* form of verbs.

(18) Hanako ni tegami o kaite wa mita ga henzi ga nakatta
to letter OBJ write see:PST but reply exist:NEG:PST
'(I) tried writing a letter to Hanako, but there was no reply.'

Despite the wide range of elements which *wa* can mark, there are cases in which *wa* cannot be inserted, as illustrated in the following. First, *wa* cannot appear with noun modifying elements in general, whether it is phrasal or clausal.

(19) * akai wa kuruma o katta
red car OBJ buy:PST
'(I) bought a red car.'

(20) * [Hanako ga katta] wa kuruma o mita
buy:PST car OBJ see:PST
'(I) saw the car which Hanako bought.'

(21) * [Hanako ga kekkonsuru tteiu] wa uwasa o kiita
marry NOM rumor OBJ hear:PST
'(I) heard the rumor that Hanako will get married.'

Also, *wa* cannot mark conditional clauses, reason clauses, nor concessive clauses, as shown in (22), (23), and (24) respectively.

(22) * mosi Hanako ga kitara wa denwasimasu
if come:COND call
'If Hanako comes, (I) will call (you).'

(23) * Taro ga konakatta node wa Hanako ga okotta
come:NEG:PST because become-upset
'Because Taro didn't come, Hanako became upset.'

(24) * Ken ga denwasitemo wa Hanako wa Ken to eega ni ikanakatta
call:CONS with movie to go:NEG:PST
'Although Ken called (her), Hanako didn't go to a movie with Ken.'

In summary, unlike *ga*, *wa* marks a wide range of elements, and this would explain why the previous studies often limited their scope to only those cases in which *wa* marks the subject of sentences. The present study will begin with *wa* and *ga* marking subjects and extend the analysis to both *wa* marking subjects and *wa* marking non-subjects on the basis of conversational Japanese. In the following sections throughout this thesis, I use the term *subject* to refer to elements which are eligible for *ga* marking.[5] I use the term *wa-marked subject* to refer to elements which are eligible for *ga* marking but marked by *wa* in a given sentence and *wa-marked nonsubject* to refer to elements which are not eligible for *ga*-marking but marked by *wa* in a given sentence.

2.3. PREVIOUS APPROACHES TO FUNCTIONS OF *WA* AND *GA*

Studies of *wa* and *ga* have been done in a wide variety of frameworks, from the sentence level to the discourse level. This section will examine major studies among those in order to see how *wa* and *ga* have been defined in the literature; first, the distinction between *wa* and *ga*, and second, different functions of each, which have been suggested in some of the previous works.

2.3.1. Contrast between *Wa* and *Ga*

The classic definitions of *wa* and *ga* is the contrast between old or given information and new information, as found in Mikami (1963), Kuno (1972, 1973), Ono (1973), Chafe (1976), Hinds and Hinds (1979), Inoue (1980), Hinds (1987). Namely, *wa* marks old or given information and *ga* new information when they occur in the subject

position of a matrix clause. Although the terms old/given and new are not exactly the same across the different researchers in this framework, they all appeal to the notion of *newness* of information to account for the use of *wa* and *ga*.[6] Hinds and Hinds (1979), for example, reports a referential progression pattern in their study of oral narratives that *ga* is used for first mention, while *wa* for subsequent mention. Kuno (1972: 277) cites the following to illustrate his claim about *wa*/*ga* distinction in terms of old/new information.

(25)

1. gootoo ga boku no ie ni haitta
   robber I GEN house into enter:PST
   'A robber broke into my house.'

2. sono gootoo *ga/wa boku ni pisutoru o tsukitsukete
   the robber I to gun OBJ point

   kane o dase to itta
   money OBJ give:IMP QT say:PST
   'The robber, pointing a pistol at me, said, "give me money."'

Kuno argues that *sono gootoo* in (25.2) is clearly coreferential with *gootoo* in (25.1) and *sono gootoo* is predictable; therefore, the use of *ga* in (25.2) results in ungrammaticality. This simple old/new dichotomy may explain some portion of the data; however, it does not give a full account and there arise problematic cases as we explore the data, as seen in later sections.

Hinds (1987) elaborates the old/new distinction for the functions of *wa* and *ga* on the basis of *Assumed Familiarity hierarchy* suggested by Prince (1981b). By this hierarchy shown below, Prince attempts to explain the relationship between the form of the text and kinds of assumptions about the hearer/reader from the point of view of a speaker/writer.

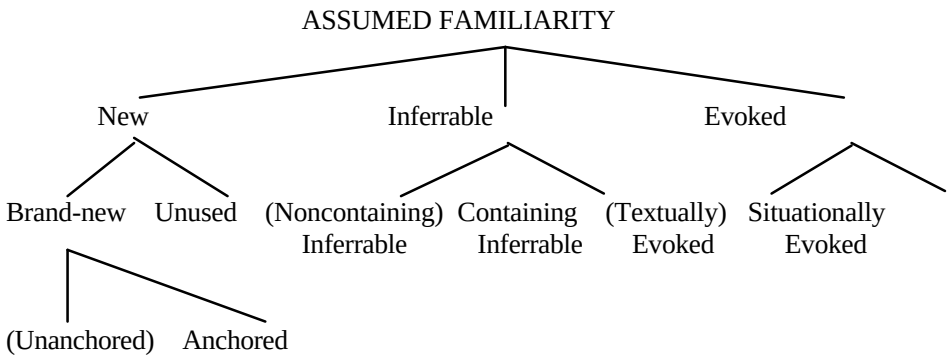


Figure 1: Assumed Hierarchy by Prince (1981b)

Following Prince's categorization, Hinds (1987) observes the use of *wa* and *ga* as in Table 1. According to Hinds, although both *wa* and *ga* are used for Noncontaining Inferrable, Evoked and Inferrable are typically marked by *wa*. Brand-new Anchored information may be marked by either *wa* or *ga* and Brand-new Unanchored is the only category which takes *ga* only. Hinds finds a general correlation between the use of *wa*/*ga* and the degree of *familiarity*; however, the problematic cases are categories which take either *wa* or *ga*. As long as these cases are unexplained by the notion of familiarity, there is a necessity to appeal to other factors.

Table 1: *Wa*/*ga* in Prince's hierarchy

Prince's Hierarchy	Occurrence of *wa*/*ga*
Situationally Evoked	*wa*
Textually Evoked	*wa*
Containing Inferrable	*wa*
Noncontaining Inferrable	*wa*/*ga*
Unused New	*wa*/*ga*
Brand-new Anchored	*wa*/*ga*
Brand-new Unanchored	*ga*

Contrary to the claims based on the old/new dichotomy above, it has been observed (e.g. Clancy 1980, Clancy and Downing 1987, Maynard 1981, 1987, Watanabe 1989, Watanabe 1990) that there are cases which the simple old/new distinction of *wa* and *ga* does not account for. For example, Maynard (1981: 115-116) shows that Kuno's

example above becomes perfectly acceptable with *ga* in the second sentence if we put it in the following context.

(26)

1. gootoo ga boku no ie ni haitta
   robber I GEN house into enter:PST
   'A robber broke into my house.'

2. sono gootoo ga boku ni pisutoru o tsukitsukete
   the robber I to gun OBJ point

   kane o dase to itta
   money OBJ give:IMP QT say:PST
   'The robber, pointing a pistol at me, said, "give me money."'

3. sono toki tomodachi no Yamanaka-san ga heya ni haitte kita
   that time friend GEN room into enter come:PST
   'Then my friend Yamanaka came into the room.'

4. Yamanaka-san wa doa no soba ni atta raihuru o tsukamu to
   door GEN near at exist:PST rifle OBJ grab as soon as

   atarikamawazu uchidasita
   wildly begin-to-shoot:PST
   'As soon as (he) grabbed the rifle that was by the door, Yamanaka began to shoot wildly.'

Maynard explains that the use of *ga* in (26.2) is perfectly acceptable because the discourse introduces a thematic element *Yamanaka-san* in (26.3), whose *wa* marks the entity as the theme of the text. She claims, also in Maynard (1987), that *wa* is used to mark *thematic* information, while *ga* is used to mark *subordinate*, *non-thematic* information, and that it is not the new/old distinction that determines the use of *wa* and *ga* but the writer's thematic choice of one entity over the others. According to Maynard, *wa* places an entity on what she calls the *thematic stage*, which is defined as "the conceptual framework within which the story is told, presented and performed" (1981: 124). In (26) above, the writer chooses *Yamanaka-san* to be the theme of the discourse, not *gootoo*; therefore, the use of *ga* in (26.2) is perfectly acceptable. In essence, Maynard claims that only given information which the writer chooses to be the theme will be marked by *wa*.

Here one should note that Maynard's notion *staging* is a strategy which a narrator uses to control the thematicity throughout the narrative utilizing the effect of *wa* and *ga*, as she explains as follows.

> Through this "staging" strategy, a narrator accomplishes an organization of narrative information in accordance with his/her perspective. It is as if the narrator places participants at different spots on the narrative stage for different durations of time in the consciousness of the narrator and his/her intended audience, the reader....What the narrator wishes to accomplish through "staging" is to discriminate the thematized participants from non-thematized ones in such a way that thematized participants remain activated, evoked, and stored in the reader's consciousness. (Maynard 1987: 61)

What Maynard (1981, 1987) seems to focus on is how the narrator can manipulate the thematization utilizing the functions of *wa* and *ga* within the possible range of the use of these two particles and what effect results from this manipulation. However, this approach does not account for the use of *wa* and *ga* on the mini-discourse level, which we will look at in Section 2.4.1, and this suggests the necessity of capturing fundamental functions of the two morphemes on the more local level.

Contrary to the claim that *wa* is the theme indicator, as found in Maynard (1981, 1987), Clancy and Downing (1987) concluded, based on data from oral narratives, that the use of *wa* is for the most part locally motivated rather than motivated by overall discourse theme. They report that a large number of uses of *wa* in their oral-narrative data do not appear to be thematic at the discourse level but *wa* simply exhibits a semantic contrast and linking between local linguistic elements, hence, *wa* as, what they call, a "local cohesive device." Although the elements in this kind of contrastive contexts linked by *wa* do not have to be thematic, the use of *wa* for contrastivity makes it look like a theme marker since "important characters generally have a larger role in the plot, there will be a tendency for *wa*-marking to occur on thematic participants" (Clancy and Downing 1987: 47).

Watanabe (1989: 142) gives a quantitative analysis of written narratives in Japanese on the basis of the quantitative methodology developed by Giv n (ed.) (1983) and makes the following claims as to the fundamental difference of *wa* and *ga* in the subject position.

(i) The post-positional particle *wa* marks definiteness.

(ii) If a referent is indefinite, then it is marked by *ga*.

Following Du Bois (1980) and Giv n (1984), Watanabe (1989: 135) defines definiteness as follows.

> The notion of definiteness and indefiniteness come from the speaker's assumption about the hearer's knowledge in human discourse. Speakers code a referent as indefinite if they think that they are not entitled to assume that the hearer can assign it a unique referential identity. Conversely, speakers code a referent as definite if they think that they are entitled to assume that the hearer can assign it a unique reference.

According to Watanabe, there are three ways for an NP to be definite: (i) it is generically shared (e.g. unique referents such as 'the sun') or culturally shared (e.g. 'father' because of the shared cultural knowledge that every person has a father), (ii) it is deictically-shared (e.g. 'eye' with the deictically related referent 'lion' being mentioned in the immediately preceding context), and (iii) it is contextually shared (i.e. being mentioned in the anaphoric discourse context). Given the criteria above, the results shows that 99.5% of the *wa*-marked NPs are definite, while about 50% of *ga*-marked NPs are indefinite. Watanabe investigated NPs marked by *wa*/*ga* in the data from two other different angles: *old* vs *new* (first-time mention vs non-first-time mention) and *resident* vs *transient* (entity which appeared at least three times in the preceding discourse vs entity which appeared at most twice); however, she concluded that the difference of *wa* and *ga* is most clearly shown by the definite vs indefinite measure.

Following Giv n's (1989) proposal of interpreting referential coherence as mental processing, Watanabe (1989: 159-160) suggests that the old information is "a filing

address or file name, into which new information is to be fed." The decision to be made for *wa*/*ga* (and the zero anaphora) choice depends on *importance* (i.e. degree of persistence in the cataphoric discourse context) and *definiteness* (as defined above) of the referent. The decision making flows as follows.

> If the referent is important, then the file is called for and activated, otherwise it's not activated. When the referent is important and will be activated, if the referent is definite, then the existing file is activated, and if the referent is indefinite then a new file is activated.(Watanabe 1989: 160)

Watanabe indicates the use of *wa* and *ga* (and the zero) in this decision making flow as in Figure 2. It is noteworthy that Watanabe (1989) attempts to account for the correlation between the use of *wa*/*ga* and cognitive function; however, Watanabe (1989) is different from the present study in terms of the use of the notion of activation in the following regards.

In principle, Watanabe's (1989) notion of activation applies to *mental storage files*, into which new information is to be fed. Watanabe uses the notion of importance to supplement her accounts for the use of *ga*. It is only *important* elements that are activated by being mentioned in the discourse; *non-active unimportant* NPs marked by *ga* are not activated at all; instead, such NPs are stored in a currently active file. (See Figure 2.) Furthermore, Watanabe defines the use of *wa* and *ga* (and the zero) in terms of the absolute values, i.e. active vs nonactive, important vs unimportant, etc., while the present study defines *wa* and *ga* in terms of relative values in activation of referents or propositions, for example, *wa*-marked referents are more active than *ga*-marked referents. In addition, Watanabe claims that the NPs marked by *wa* and *ga* are not active and only the information manifested by zero anaphora is active. This study will claim that *wa* marks information which is active at the point of mention, while *ga* marks nonactive or less active information; therefore, *wa*-marked information is not promoted from nonactive to active, as in Watanabe (1989), but it is already more or less active at the point of mention. Finally, I do not find it necessary to suggest the *filing mechanism* in mental storage, as in

Watanabe (1989), in order to account for the use of *wa* and *ga*. Such a mechanism may be useful for explaining mental organization of information; however, it is beyond the scope of the present study.

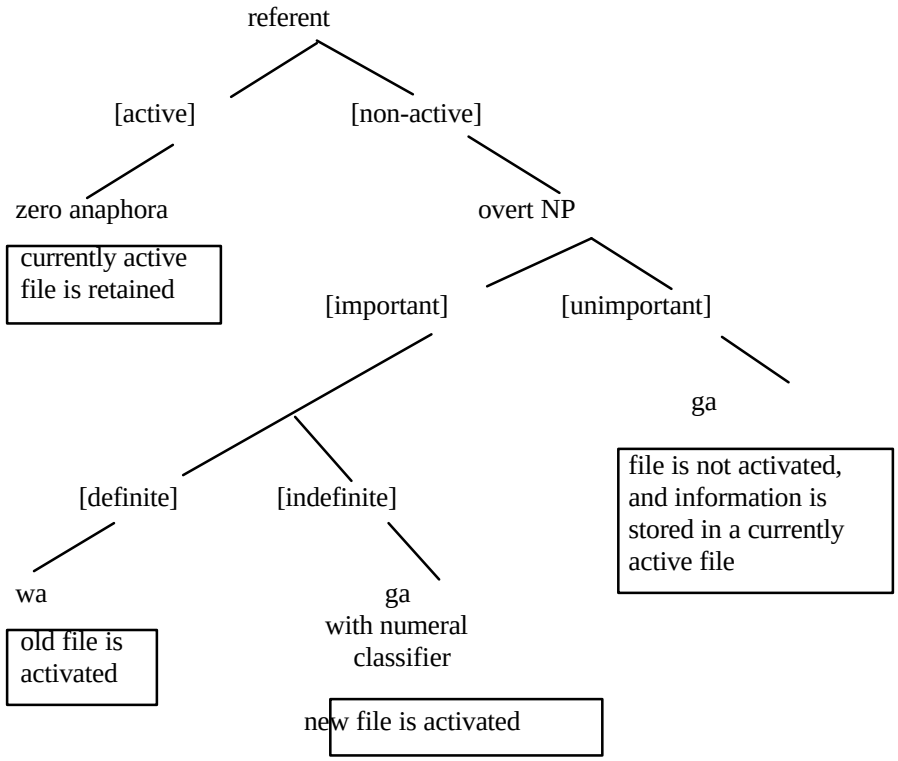


Figure 2: Japanese grammar of referential coherence as mental processing instructions by Watanabe (1989: 162)

2.3.2. Distinct Functions of *Wa*?

Regarding the function of *wa*, Kuno (1972, 1973) distinguishes two discrete functions: *thematic* and *contrastive*. He states that a thematic *wa* phrase is either anaphoric or generic, while there is no such constraint on contrastive *wa*. Consider the following examples from Kuno (1973).

(27) kuzira wa honyuu-doobutsu desu
whale mammal COP
'Speaking of whales, they are mammals. (A whale is a mammal.)'

(28) John wa watakusi no tomodachi desu
I GEN friend COP

'Speaking of John, he is my friend.'

(29) * ame wa hutteimasu
rain falling
'Speaking of rain, it is falling.'

(30) ame wa hutte imasu ga taisita koto wa arimasen[7]
rain fall be but serious matter exist:NEG
'It is raining, but it is not much.'

The NP marked by *wa* in (27) and (28) are generic and anaphoric respectively, and as indicated in the English gloss Kuno assigns the thematic reading in these sentences. On the other hand, *ame* in (29) is neither generic nor anaphoric; therefore, the thematic reading is not possible. However, the same sentence can be grammatical in the context of contrast as in (30). This analysis will find ambiguity between the two readings, thematic and contrastive, when a generic or anaphoric noun phrase is marked by *wa*, as Kuno states.

(31) John wa sono hon o yonda
that book OBJ read:PST
'John read the book.'

(32) John wa sono hon o yonda ga Mary wa yomanakatta
that book OBJ read:PST but read:NEG:PST
'John read the book, but Mary didn't.'

(32) gives the contrastive reading since there is an overt entity *Mary* contrasted with *John*, while (31) can be ambiguous as to whether *John* is contrasted with some entity or not. To disambiguate the sentences such as (31), Kuno claims that NPs of thematic *wa* do not receive prominent intonation, while those of contrastive *wa* do. Namely, *John wa* receives prosodic emphasis if it is contrastive, while it does not if it is thematic. However, a problem with appealing to the prosodic emphasis is that it is not clear whether the speaker must put the prosodic emphasis on *wa*-NPs in order to convey the contrastive reading. For example, the sentence in (32) can be read without prominent intonation on the *wa*-NP in order to convey the contrastive sense. Also, in an appropriate context such as in (33), the sentence (31) can give the contrastive reading without prominent intonation on its *wa*-NP.

(33)

A: asita siken ga aru kedo mada hon o zenzen yondenai
tomorrow exam exist but yet book OBJ at-all read:NEG
'(I) have an exam tomorrow, but (I) haven't read the book at all.'

B: John wa sono hon o yonda tte
the book OBJ read:PST QT
'(I've heard) John has read the book.'

The crucial point here is that *John wa* may or may not receive the prominent intonation to give the contrastive reading since the contrastive sense is clear by the discourse context. Therefore, the prosodic emphasis itself is not a firm criterion to distinguish the two functions of *wa*.[8]

Kuno (1972, 1973) also bases his distinction between thematic *wa* and contrastive *wa* also on the restricted use of *wa* in subordinate clauses. According to Kuno, thematic *wa* cannot occur in a subordinate clause.

(34) Ken wa [Hanako ga/*wa okane o nusunda] no o sitteimasu ka
money OBJ steal:PST NOM OBJ knowing Q
'Does Ken know that Hanako stole the money?'

This restriction applies to any NP in a subordinate clause. With an object NP:

(35) Ken wa [sono okane o/*wa Hanako ga nusunda] no o sitteimasu ka
the money OBJ steal:PST NOM OBJ knowing Q
'Does Ken know that Hanako stole the money (but not something else)?'

In a relative clause:

(36) Ken wa [Hanako ga/*wa mottekita] hon o karimasita
bring:PST book OBJ borrow:PST
'Ken borrowed the book which Hanako brought.'

In an adverbial subordinate clause:

(37) [Ken ga/*wa kaettekita toki] Hanako wa neteita
return:PST when sleeping:PST
'Hanako was sleeping when Ken returned.'

Kuno states that the reading of *wa*-NP as the topic in subordinate clauses such as above is impossible, while the contrastive *wa* is possible. In (34), for example, the sentence with *wa*

in the embedded clause reads "Does Ken know that Hanako, not somebody else, stole the money?" Although Kuno acknowledges the use of contrastive *wa* in embedded clauses, native speakers' judgments vary. A number of native speakers, including myself, rate the above examples as "not totally acceptable" even with the contrastive *wa*. Given the discrepancy as to the native speakers' judgment and the difficulty of judging sentences in isolation, we need to examine actual discourse in terms of the use of *wa* in embedded clauses.

Since Kuno's (1972, 1973) claim of two different functions of *wa*, there have been a number of studies which suggest a single underlying function of *wa*, instead of classifying its functions into thematic and contrastive (e.g. Yoshimoto 1981, Makino 1982, Inoue 1983, Miyagawa 1987, Watanabe 1989, Shibatani 1990). Linguists in this category generally agree with the claim that choosing one entity as the *theme* or *important* entity of discourse automatically entails the sense of contrast, whether overtly with other competing entities in the discourse or covertly without such entities explicitly mentioned, and the degree of contrast is determined by the context. In this point, Shibatani (1990: 265) states "there aren't two distinct *wa*'s, or two distinct meanings associated with *wa*, as suggested by the labels "thematic" *wa* and "contrastive" *wa*; rather, one and the same *wa* has the effect of emphasizing the contrast when the discourse environment provides a background for contrast". Then, what is the basic function of *wa*? Although they word their definitions differently, they seem to address the same point; namely, *wa* separates the entity which it marks from the rest of things present in the context in order to make a certain judgment about the entity chosen. In fact, this function of *wa* which separates the *wa*-marked element and the following which gives the judgment about the *wa*-marked element has long been pointed out by Kuroda (1965, 1972). Kuroda argues that a *wa* sentence represents a *predication*, while a *ga* sentence represents a *non-predicational description*. Predication refers to an act of attributing the information expressed by the predicate to the entity expressed by the (*wa*-marked) subject, and non-predicational

description refers to a representation of a simple recognition of a fact or situation without having a particular constituent of which the situation is predicated. Kuroda refers to the former as the *categorical judgment* and the latter as the *thetic judgment*, following the theory of Franz Brentano and Anton Marty. There are two units in sentences of the categorical judgment, i.e. *wa* sentences: the subject and the rest which gives the predication of the subject, while there are no such separate units in sentences of the thetic judgment, i.e. *ga* sentences.

2.3.3. Functions of *Ga*

Although discussing in detail the different functions of *ga* is beyond the scope of this paper, I will delineate the basic function of *ga* in contrast to the function of *wa* which we have looked at so far.

Kuno (1972, 1973) states that *ga* as subject case marker is either for *neutral description* or for *exhaustive listing*. The following are from Kuno (1973); the sentence in (38) exemplifies neutral description *ga* and (39) exhaustive listing *ga*.

(38) John ga asoko ni tatteiru
there at standing
'John is standing over there.'

(39) John ga gakusee desu
student COP
'John is a student.'

According to Kuno (1973: 51), sentences with neutral description *ga* "present an objectively observable action, existence, or temporary state as a new event", while exhaustive listing *ga* lists individuals which are associated with a certain property. Kuno states that the neutral description *ga* can appear only with action verbs, existential verbs, and adjectives/nominal adjectives which represent changing states, while there is no such restriction on exhaustive listing *ga*.[9] In terms of the shared knowledge of the speakers and the hearer, a sentence with the neutral description *ga* can be uttered out of the blue, while

a sentence with the exhaustive listing *ga* requires certain prior knowledge. A sentence of exhaustive listing *ga* is illustrated by an answer to a WH question, as in the following.

(40)
A: dare ga gakusee desu ka
who student COP Q
'Who is a student?'

B: John ga gakusee desu
student COP
'John is a student.'

In the immediate discourse context of (40B), there is a shared knowledge *X is a student* and the exhaustive listing *ga* marks the individual which is associated with the shared property *student*.

Although the different functions of *ga* may be identified as seen above, Kuno (1972: 273) claims a single underlying function of *ga*; namely, "*ga* as subject marker in the matrix sentence *always* signals that the subject conveys new, unpredictable information."[10] In general, linguists have been in agreement with the idea that *ga* marks *new* information; however, the *newness* is defined in different ways. *New* or *unpredictable* in its literal sense is found in Kuno (1972, 1973), Makino (1982), and Yoshimoto (1982). As discussed above, Maynard (1981, 1987) rejects the distinction of old and new and appeals to thematicity; namely, *wa* marks the theme, while *ga* marks subordinate and non-thematic information, and she claims that the narrator manipulates the thematicity controlling the use of *wa* and *ga*. Watanabe (1990) suggests on the basis of her data of written narrative that the characterization of *ga* as marker for new information is not satisfactory and claims that the use of *ga* must be explained in terms of *deictic center*, a frame of reference, along with its basic function as marker of new information. According to Watanabe (1990), it is crucial from whose viewpoint information is new to determine the use of *ga*. Watanabe's claim above seems to be valid not only in narratives but also in discourse in general; however, this particular issue is not examined in this study.

The discussion of some previous studies above tells us that the simple dichotomy old/new is problematic in the analysis of *wa* and *ga* and there is a necessity for a different framework which accounts for the fundamental distinction between the two. The following sections introduce the notion *activation* and the following analyses will show how some of the claims made in the previous studies above are captured in terms of this framework.

2.4. *WA* AND *GA* IN MINI-DISCOURSE

For the purpose of drawing hypotheses as to the contrast between *wa* and *ga*, we will begin the analysis with data from mini-discourse, primarily question-answer pairs. After laying out the tentative hypotheses, we will discuss to some extent the issues regarding the distinct functions of *wa*, i.e. thematic *wa* and contrastive *wa*, from the activation point of view.

2.4.1. Contrast between *Wa* and *Ga*

In order to investigate the contrast between *wa* and *ga* in terms of activation, this section will begin the analysis with simple question-answer pairs.

(41)
A: kyoo dare ga John o tsurete kita no?[11]
today who OBJ take come:PST FP
'Who brought John (here) today?'

B: Mary ga/*wa tsurete kita yo[12]
take come:PST FP
'Mary brought (John here).'

The context (41A) activates the open proposition *X brought John today*, and there is no other active proposition.[13] Furthermore, the referent of *X* remains nonactive. By uttering

(41B), B supplies the referent 'Mary' for the X-role, and thereby assumes the 'Mary' is not already active in A's consciousness. B marks this information, which is nonactive for the addressee A, with *ga*.

(42) illustrates use of *ga* in a different context.

(42)
A: doo sita no?
how did FP
'What's going on?'

B: ame ga/*wa hutteru yo
rain falling FP
'It's raining.'

In the context of (42A), the only activated proposition is *something is happening*; therefore, B assumes that the entire proposition expressed by (42B) is not active at all in A's consciousness at the time of mentioning (42B). The open proposition *X is falling* and the referent *rain* are both nonactive, and B marks the nonactive information with *ga*. Note that (41B) is a case of exhaustive listing *ga* and (42B) neutral description *ga* in Kuno's terms, and in both cases the *ga* marks nonactive referents.

Now consider the following cases of *wa*.

(43)
A: kinoo Ken to Hanako ga paatii ni itta tte
yesterday and Party to go:PST QT
'(I've heard) Ken and Hanako went to a party yesterday.'

B: sitteru Ken wa/?ga sugu kaetta kedo Hanako wa/?ga hitoban ita tte
knowing soon return:PST but all-night be:PST QT
'I know. (I've heard) Ken left soon, but Hanako was (there) all night.'

At the time of mention of (43B), the referents 'Ken' and 'Hanako' are active in the consciousness of both speakers A and B since they are mentioned in (43A), while the open propositions *X went home soon* and *X was there all night* in (43B) are not active at all in A's consciousness in (43A). The speaker B marks the active referents which fill the open propositions with *wa*. Note that *wa*'s in (43B) are the contrastive ones in Kuno's terms.

As in the case above, the use of *wa* without a contrastive sense exhibits the same activation pattern; the *wa*-marked referent is active in the consciousness of both speaker and hearer in the preceding context, while the open proposition which the *wa*-marked entity fills is nonactive in the addressee's consciousness in the preceding context. In (44), the *wa*-marked referent 'John' is active, being mentioned in (44A), while the open proposition *X walked* in (44B) is nonactive in (44A).

(44)
A: kinoo John ga uti ni kita yo
yesterday house to come:PST FP
'John came to (my) house yesterday.'

B: John wa/?ga aruite kita no?
walk come:PST FP
'Did John walk (to your house)?'

The contrast between *wa* and *ga* as seen above is found not only with *wa*/*ga*-marked referents but also with *wa*/*ga*-marked propositions. (45)-(47) illustrate *wa* and *ga* in cleft sentences.

(45)
A: John wa dare o mita no?
who OBJ see:PST FP
'Who did John see?'

B: John ga mita no wa/*ga Ken da
see:PST NOM COP
'It was Ken that John saw.'

The proposition *John saw*, which is marked by *wa* in (45B), is active in (45A), while the open proposition *X is Ken* in (45B) is nonactive in (45A).

In (46), on the other hand, the preposed constituent in the cleft sentences is marked with *ga*. The proposition *reading a book* in (46B) is nonactive in (46A), while the open proposition *X is John* (46B) is active in (46A).

(46)
A: John wa dono hito
which person

'Which person is John?'

B: hon o yondeiru no ga/*wa John desu
book OBJ reading NOM COP
'The one who is reading a book is John.'

Technically speaking, however, the person who is reading a book is likely to be more or less active in A's consciousness in (46) if the referent is present in the situation; namely, A's question in (46) can be paraphrased as in (47).

(47) asokoni hito ga nanninka iru kedo, John wa dono hito
over-there person some be but which person
'There are some people over there, but which person is John?'

In the context of using (47), the person who is reading a book may be more active than other referents which are not present in the situation. However, the crucial point is that in the immediately preceding context (46B) *the person who is reading a book* is not as activated as the open proposition *X is John*, and the speaker marks this less active proposition with *ga*.[14]

The observation above suggests that the speaker prefers *wa* when *wa*/*ga*-marked information is active and the open proposition which the *wa*/*ga*-marked information fills is nonactive in the addressee's consciousness in the immediately preceding context, and the speaker prefers *ga* when the *wa*/*ga*-marked information is not active in the addressee's consciousness in the immediately preceding context, regardless of the activation status of the open proposition which the *wa*/*ga*-marked information fills. The contrast is schematically shown in (48).

(48) [active] *wa* [nonactive]

[nonactive] *ga* [active or nonactive]

This schema shows that the choice of *wa* and *ga* is motivated by two independent factors. In terms of *wa*/*ga*-marked information, *wa* appears with active information, while *ga* with non-active information. In terms of the open proposition which the *wa*/*ga*-marked referent fills, on the other hand, it is only *ga* that can appear with an active proposition.

Although the contrast between *wa* and *ga* is roughly captured by the schema above, it is necessary to further examine the use of *wa*/*ga* because there are cases in which the speaker marks with *ga* information which is already active in the addressee's consciousness in the immediately preceding context.

(49)

A1: kinoo dare ga Hanako to deetosita ka sitteru deshoo?
yesterday who with date:PST Q knowing COP:HOR
'(You) know who went out with Hanako yesterday, don't you?'

B1: siranai. dare?
know:NEG who
(I) don't know. Who?

A2: sitteru deshoo? hora
knowing COP:HOR you-know
'(You) know (who), don't you? You know.'

B2: aa. John ga/*wa Hanako to eega ni itta n deshoo?
ah with movie to go:PST NOM COP:HOR
'Right. John went to a movie with Hanako, didn't he?'

In (49), the speaker A asks the speaker B who went out with Hanako yesterday. A knows who it is, i.e. 'John', but A asks this question to B just to remind B of the referent. Until the point just before uttering (B2), the referent 'John' is not in B's consciousness; the referent 'John' comes into B's center of attention just before B's uttering (B2). On the other hand, 'John' is in A's center of attention in the preceding context of (B2). In (B2), 'John' is marked with *ga*, though it is active in the addressee's consciousness in the preceding context. In fact, 'John' is active not only in A's consciousness but also in B's consciousness, hence mutually active, at the time just before the utterance. The use of *ga* here contrasts with the use of *wa* in (50B2). The sentences are identical except for the *wa*/*ga* marking; however, the preference of *wa*/*ga* is determined by the activation status of the referent 'John' in the preceding context.

(50)

A1: kinoo dare ga Hanako to deetosita ka sitteru deshoo?

yesterday who with date:PST Q knowing COP:HOR
'(You) know who went out with Hanako yesterday, don't you?'

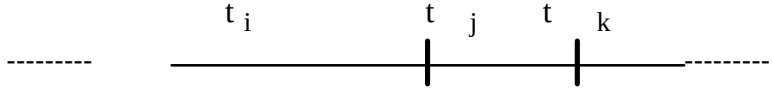
B1: John deshoo?
COP:HOR
'(It's) John, isn't it?'

A2: soo odoroki
yes surprise
'Right. (It's) surprising.'

B2: John wa/?ga Hanako to eega ni itta n deshoo?
with movie to go:PST NOM COP:HOR
'John went to a movie with Hanako, didn't he?'

The referent 'John' is mutually active, i.e. active in both A and B's consciousness, not only at the time just before B's uttering (B2) but also in the preceding context, i.e. at least at the interval of preceding two clauses (B1) and (A2).

The observation above suggests that it is useful to break down the activation span of a referent into at least three stages: (1) the time of utterance, (2) a brief interval just before the time of utterance, and (3) an interval before and adjacent to (2), i.e. an interval before the utterance but excluding the brief interval just before the time of utterance. Figure 3 illustrates the three intervals.



t_k: during and after utterance

t_j: a brief interval just before t_k

t_i: an interval before and adjacent to t_j

Figure 3: Three intervals of utterance

On the basis of the time scale illustrated above, I posit two different types of entities in terms of activation status: (i) mutually active entity, and (ii) non-mutually active entity. An entity is mutually active if it is active or more active in the consciousness of

both speaker and addressee at an interval t_i, i.e. the interval before the utterance excluding a brief interval just before the utterance, while an entity is non-mutually active if it is not active or less active in the consciousness of both speaker and addressee at an interval t_i. The observation of mini-discourse data above suggests that mutually active entities are marked with *wa*, while non-mutually active entities are marked with *ga*. More precisely, the contrast between *wa* and *ga* is captured as in (51).

(51)

a. *Wa* marks an entity which is mutually active for speaker and addressee at an interval t_i, continuing into t_k.

b. *Ga* marks an entity which is not mutually active at t_i. There are two types in this category: (i) The entity is active for speaker at t_i, but not for addressee, and becomes activated for the addressee at t_k, i.e. as he hears the utterance, and (ii) the entity is not active for speaker at t_i, but becomes active at t_j, i.e. as he realizes the referent of the entity before the utterance.

The first type of (51b) is exemplified by the cases as in (41) and (42), and the second is exemplified by the cases as in (49). In sum, *wa* is differentiated from *ga* in terms of the activation status of a referent at the interval t_i, and the *immediately preceding context* of an utterance henceforth refers to this interval, unless noted otherwise.

Having laid out the issue of timing of activation, I examine the use of *wa* and *ga* in a greater variety of mini-discourse contexts in order to further evaluate the characterization of the two, as seen in (48). (52) is from Kuno (1972: 277).

(52)

1. gootoo ga boku no ie ni haitta
   robber I GEN house to enter:PST
   'A robber broke into my house.'

2. sono gootoo wa/*ga boku ni pisutoru o tsukitsukete
   the robber I to gun OBJ point

   kane o dase to itta
   money OBJ give:IMP QT say:PST
   'The robber, pointing a pistol at me, said "Give me money."'

Kuno explains that the use of *ga* is ungrammatical if there is only one robber present in the discourse since *gootoo* in (52.1) is coreferential with *gootoo* in (52.2); therefore, it is predictable. However, the use of *ga* above is perfectly acceptable if the speaker is simply describing a sequence of events, for example, in reporting the crime scene to the police step by step. The possibility of *ga* here in the same example is discussed also by Maynard (1981: 115-116); though she gives an explanation in terms of thematicity. Namely, as a part of the *staging* strategy, as mentioned above, the narrator can mark an NP by *ga*, not *wa*, if the narrator intends not to choose the NP as theme of the text. Whatever the explanation would be, the example here, in spite of Kuno's grammaticality judgment, seems to suggest that *ga* can mark NPs which are active. Note that given the fact that *ga* is also an acceptable choice in (52.2), neither predictability nor shared knowledge gives a satisfactory account for the data here.

Ga marking active entities appears to contradict the principle stated in (48), since the principle predicts that *ga* marks nonactive entities. However, I will claim that the same principle can be used to account for these seeming problematic cases in terms of relative degrees of activation of *wa*/*ga*-marked entities and open propositions. To demonstrate this point, I examine three sets of mini discourse below.

(53)

A1: kinoo no yoru kyuuni madogarasu ga wareru oto ga kikoeta n da yo
yesterday GEN night suddenly window-glass break noise hear:PST NOM COP FP
'Last night I suddenly heard the glass window broken.'

B: e hontoo sorede doo natta no?
oh really then how become:PST FP
'Really? Then what happened?'

A2: gootoo ga/#wa boku ni pisutoru o tsukitsukete
robber I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'The robber, pointing a pistol at me, said "Give me money."'

In (53A2), *ga* is the only possible choice because the 'robber' and the open proposition *X pointed a pistol at me and said "Give me money"* are equally nonactive in B's consciousness in the immediately preceding context.

In the identical sentence (54A2), on the other hand, *wa* is not totally unacceptable, though *ga* is better than *wa*.

(54)
A1: kinoo gootoo ga boku no ie ni haitta n da yo
yesterday robber I GEN house to enter:PST NOM COP FP
'A robber broke into my house yesterday.'

B: e hontoo sorede doo natta no?
oh really then how become:PST FP
'Really? Then what happened?'

A2: soitsu ga/?wa boku ni pisutoru o tsukitsukete
the-guy I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'The guy, pointing a pistol at me, said "Give me money."'

In the preceding context of (A2), the 'robber' is already active, having been introduced in (54A1). However, the hearer's utterance in (54B) suggests that the 'robber' is not his center of attention since the question *What happened next?* does not assume a particular agent for the subsequent events. In (54A2), *soitsu* is focused with *ga*, because the 'robber' is not yet the center of attention for B. This point becomes clear with the following example.

(55)
A1: kinoo gootoo ga boku no ie ni haitta n da yo
yesterday robber I GEN house to enter:PST NOM COP FP
'A robber broke into my house yesterday.'

B: e hontoo sorede nanika sareta no?
oh really then something do:PAS:PST FP

'Really? Then did (the guy) do something to you?'

A2: soitsu wa/#ga boku ni pisutoru o tsukitsukete
the-guy I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'The guy, pointing a pistol at me, said "Give me money."'

In (55A2), *wa* is the more appropriate choice than *ga*. (54) and (55) are identical in that the 'robber' is active, being mentioned in (A1) in the immediately preceding context of (B1); however, the question in (55B1) suggests that the 'robber' is B's center of attention. Furthermore, the open proposition *X pointed a pistol at me and said "Give me money"* is not active in B's consciousness, i.e. not mutually active. Here the *wa*/*ga*-marked referent is more active than the open proposition; therefore, *wa* is preferred over *ga*.

Interestingly, the contrast among (53) through (55) is further illustrated by the acceptability of zero anaphora for the entity *gootoo/soitsu* in (A2). Ellipsis is totally unacceptable in (53), it is marginally acceptable in (54), and it is perfectly acceptable in (55). This is illustrated by (56), (57), and (58) respectively.

(56)
A1: kinoo no yoru kyuuni madogarasu ga wareru oto ga kikoeta n da yo
yesterday GEN night suddenly window-glass break noise hear:PST NOM COP FP
'Last night I suddenly heard the glass window broken.'

B: e hontoo sorede doo natta no?
oh really then how become:PST FP
'Really? Then what happened?'

A2: {gootoo ga / # } boku ni pisutoru o tsukitsukete
robber I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'A robber, pointing a pistol at me, said "Give me money."'

(57)
A1: kinoo gootoo ga boku no ie ni haitta n da yo
yesterday robber I GEN house to enter:PST NOM COP FP
'A robber broke into my house yesterday.'

B: e hontoo? sorede doo natta no?
oh really then how become:PST FP
'Really? Then what happened?'

A2: {soitsu ga / ? } boku ni pisutoru o tsukitsukete
the-guy I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'The guy, pointing a pistol at me, said "Give me money."'

(58)
A1: kinoo gootoo ga boku no ie ni haitta n da yo
yesterday robber I GEN house to enter:PST NOM COP FP
'A robber broke into my house yesterday.'

B: e hontoo? sorede nanika sareta no?
oh really then something do:PAS:PST FP
'Really? Then did (the guy) do something to you?'

A2: { / ?soitsu wa} boku ni pisutoru o tsukitsukete
the-guy I to gun OBJ point

kane o dase tte itta n da
money OBJ give:IMP QT say:PST NOM COP
'The guy, pointing a pistol at me, said "Give me money."'

The contrast in acceptability of zero anaphora above suggests that the 'robber' is not even in the attention of the addressee in (56), it is in the attentional field, i.e. active, but not the center of attention in (57), while it is not only in the attentional field but also in the center of attention in (58).

In terms of the relative degree of activation in the addressee's consciousness, the *wa*/*ga*-marked entity, i.e. the 'robber', and the open proposition are equally nonactive in (53), the *wa*/*ga*-marked entity is a little more active than the open proposition in (54), while the *wa*/*ga*-marked entity is far more active than the open proposition in (55). In

terms of the *wa*/*ga* choice, *ga* is preferred in (53), *ga* is better than *wa*, though *wa* is marginally acceptable, in (54), and *wa* is preferred in (55).

The relative degrees of activation explain the use of *wa*/*ga* in Kuno's (1972) example discussed above, which is repeated below as (59). In (59B), *wa* and *ga* are both appropriate; however, the choice between the two depends on the speaker's assumption as to the addressee's consciousness. Namely, *ga* is appropriate when the *ga*-marked referent is not in the addressee's center of attention, even though it is active, having been mentioned in the preceding context, while *wa* is more appropriate when the entity is in the addressee's center of attention.

(59)
A: gootoo ga boku no ie ni haitta
robber I GEN house to enter:PST
'A robber broke into my house.'

B: sono gootoo wa/ga boku ni pisutoru o tsukitsukete
the robber I to gun OBJ point

kane o dase to itta
money OBJ give:IMP QT say:PST
'The robber, pointing a pistol at me, said "Give me money."'

The same line of argument can be applied to the well-known example from Kuno (1972: 276) in (60).

(60)
1. John to Mary to Bill ga tazunete kimasita
and and visit come:PST
'John, Mary, and Bill called on me.'

2. John ga/wa kudamono o kuremasita.
fruits OBJ give:PST
'John gave (me) fruits.'

3. Mary ga/wa okasi o kuremasita.
cookies OBJ give:PST
'Mary gave me cookies.'

4. Bill *ga/wa hon o motte kite kuremasita

book OBJ bring come give:PST
'Bill brought me a book.'

According to Kuno, *ga* is grammatical in (60.2) and (60.3) because *John* in (60.2) and *Mary* in (60.3) are unpredictable since the person who gave fruits can be any of the three people and the person who gave cookies can be either of the two people left. In (60.4), on the other hand, *Bill* is predictable since it is the last choice; therefore, *wa* is the appropriate choice. The fact that *wa* is perfectly acceptable in (60.2) and (60.3) is not explained by the predictability account, as Kuno (1972: 276) regards them as exceptions by saying "the NP-*wa* in the subject position sometimes represents new information."

However, (60) can be straightforwardly accounted for by the notion of activation. The use of *wa* in (60.2), (60.3), and (60.4) is explained by the fact that 'John', 'Mary', and 'Bill' are all active in the immediately preceding context, having being mentioned in (60.1). The speaker may assume that the addressee's center of attention includes the people who came to visit and the open propositions *X gave me fruits*, *X gave me cookies*, and *X brought me a book* are all nonactive in the addressee's consciousness in the preceding context. In other words, the use of *wa* is appropriate for a reply to the question *What did they do next?* On the other hand, at the point of mentioning (60.2), (60.3), and (60.4), the speaker can also assume that *John*, *Mary*, and *Bill* are not the center of attention for the addressee in the immediately preceding context; therefore, the speaker can treat *John*, *Mary*, and *Bill* as a part of the proposition to be focused. In other words, the use of *ga* assumes a reply to the question *What happened next?* Although Kuno rates the use of *ga* in (60.4) as unacceptable, *ga* in (60.4) is perfectly acceptable in the context of (60) if it is mentioned in an 'event-description' context. Again, the choice between *wa* and *ga* depends on the speaker's assumption as to the addressee's consciousness in the immediately preceding context.

Given the discussion as to the interaction between the use of *wa*/*ga* and the relative degrees of activation in *wa*/*ga*-marked elements and open propositions which they fill, I

examine one more set of data before summarizing the finding. The data of the following kind, which the *predictability* approach fails to account for, is well explained by the subtle interaction between a *wa*/*ga*-marked entity and an open proposition in terms of relative degrees of activation.

(61)

A: Sally ga amerikazin de Ken ga igirisuzin na no?
American and British NOM FP
'Is Sally American and Ken British?'

B: ie Sally ga igirisuzin de Ken ga amerikazin desu
no British and American COP
'No, Sally is British and Ken is American.'

B': ie Sally wa igirisuzin de Ken wa amerikazin desu
no British and American COP
'No, Sally is British and Ken is American.'

B": ie Sally ga igirisuzin de Ken wa amerikazin desu
no British and American COP
'No, Sally is British and Ken is American.'

B"': ie Sally wa igirisuzin de Ken ga amerikazin desu
no British and American COP
'No, Sally is British and Ken is American.'

In giving a negative answer to (61A), there are four possible patterns as in (61B)-(61B"') depending on which of *wa* and *ga* is used to mark each NP.[15] While these four are all grammatical, the notion of predictability accounts only for (B"), which is the *...ga...wa* pattern. Namely, the first NP is unpredictable, while the second is predictable.

On the other hand, all four sentences in (61B)-(61B"') are explained by the notion of activation as follows. In the immediately following context of (61A), there are four active units which are appropriate for the discussion here: the two referents 'Sally' and 'Ken', and the two open propositions *X is American* and *X is British*. In terms of the

degrees of activation in entities marked by *wa*/*ga*, the *...wa...wa* pattern is appropriate because *Sally* and *Ken* are both activated by the first sentence. On the other hand, the *...ga...ga* pattern is explained by the fact that the two propositions *X is American* and *X is British* are both active. Therefore, (61) is considered as a case of conflict between two independent principles, as stated in (48): (i) *wa* marks an active entity, and (ii) an active open proposition is filled with a *ga*-marked entity. We can assume that the result of this conflict is simply choosing the one which the speaker assumes is even more active than the other in the addressee's consciousness. Therefore, (B) results if the speaker assumes the propositions are more active than the two NPs in the addressee's mind, while (B') results if the speaker assumes the two NPs are more active than the two propositions in the addressee's mind. Given this logic, it is also possible to split the patterns in the answer to (61A). Namely, one possibility is that the 'Sally' is more active than the proposition *X is British* for the first clause, while the proposition *X is American* is more active than the 'Ken'. The other possibility is that the proposition *X is British* is more active than 'Sally' for the first clause, while the 'Ken' is more active than the proposition *X is American*. These two patterns are manifested in (B'") and (B") respectively.

Furthermore, the interaction of between *wa*/*ga*-marked entities and open propositions in terms of activation is illustrated by (62).

(62)

A1: kinoo steeki to sarada o tabeta kedo, ano resutoran wa yokunakatta
yesterday steak and salad OBJ eat:PST but that restaurant good:NEG:PST
'I ate steak and salad yesterday, but that restaurant wasn't good.

B: doosite
why
'Why?'

A2: steeki wa/*ga oisikatta kedo, sarada ga/?wa mazukatta
steak tasty but salad taste-bad:PST
'The steak was good, but the salad was bad.'

In the immediately preceding context of (62A2), there are two active entities and one active proposition which are relevant for the discussion here: 'steak', 'salad', and *X was bad*. (The open proposition *X was bad* was not literally expressed in (62A1); however, *ano resutoran wa yokunakatta* gives the same sort of proposition.) In the first clause of (62A2), *wa* is appropriate and *ga* is not because 'steak' is active and the open proposition *X was good* is not active in the immediately preceding context. The second clause in (A2), on the other hand, is a case of conflict between *wa* and *ga* since 'salad' and *X was bad* are both active; therefore, *ga* and *wa* should be both acceptable. However, this particular example is an interesting case because of B's question *Why?*, which indirectly seeks the information X to fill the open proposition *X was bad*. Therefore, the speaker A can assume that the open proposition is somehow more active than the 'steak' and 'salad' in B's consciousness, which accounts for the preference for *ga* over *wa*, though *wa* is not totally unacceptable here. This point becomes obvious with the contrast with the following example.

(63)

A1: kinoo steeki to sarada o tabeta kedo, ano resutoran wa yokunakatta
yesterday steak and salad OBJ eat:PST but that restaurant good:NEG:PST
'I ate steak and salad yesterday, but that restaurant wasn't good.

B: soo
so
'Is that so?'

A2: steeki wa/*ga oisikatta kedo, sarada ga/wa mazukatta
steak tasty:PST but salad taste-bad:PST
'The steak was good, but the salad was bad.'

In (63), B's response further activates neither 'steak'/'sarada' nor *X was bad*; therefore, *wa* and *ga* are equally acceptable in the second clause of (A2).

The discussion in this section suggests that what is crucial to the choice of one over the other in the *wa*/*ga* set is the relative degrees of activation rather than the degree

of activation of each independent of each other. Given this assumption, I modify the principles in (48) as in (64).

(64)

WA: A *wa*-marked referent is more mutually active than the open proposition with which the referent combines.

GA: A *ga*-marked referent is less mutually active than the open proposition with which the referent combines or they are equally mutually nonactive.

The principles in (64) are schematically shown in (65).

(65)

a. A $\overset{wa}{>}$ B A is more active than B.

b. A $\overset{ga}{<}$ B B is more active than A.

c. A $\overset{ga}{=}$ B A and B are equally nonactive.

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

(65) shows three patterns depending on relative degrees of activation between information marked by *wa*/*ga* and the open propositions. In Kuno's (1972, 1973) terms, (65a) indicates the cases of both *thematic* and *contrastive wa* and (65b) and (65c) indicate *exhaustive listing ga* and *neutral description ga* respectively. What is missing in this paradigm is the case in which A and B are equally activated. In this type, the speaker makes an utterance which consists of only active elements; two active elements are combined in a single sentence. An example of this type is the occurrence of *wa* with a WH phrase, which will be examined in the following section. Also, this particular focus type will be further discussed in the following chapter on the basis of conversational Japanese.

2.4.2. *Wa* and WH Phrase

It has long been pointed out (e.g. Kuno 1972, 1973) that the *thematic wa* cannot appear with a WH phrase.

(66) dare *wa/ga kita no?
who come:PST FP
'Who came?'

However, it is generally said that if *wa* has a clear contrastive sense, the sentence becomes acceptable.

(67) dare wa kite, dare wa konakatta no?
who come who come:NEG:PST FP
'Who came and who didn't come?'

The contrast between (66) and (67) leads us to a question "What are the conditions which allow *wa* to occur with a WH phrase?" According to Miyagawa (1987), there are two conditions for *wa* to appropriately appear with a WH phrase.

(68) Conditions for the appropriate use of *wa* with a WH phrase by Miyagawa (1987)

a. The speaker and the hearer share the knowledge of the existence of an identifiable set of individuals in the immediate conversational context.

b. Every member of this set must be exhaustively represented in the *WH wa* question.

The sentence in (66) is ungrammatical because the thematic *wa* in general refers to a definite individual and this contradicts the nature of the WH question, in which the WH phrase lacks such definite reference. However, the use of *wa* in (67) is acceptable only if a set of definite individuals is given in the preceding context; in other words, one cannot ask a question such as (67) out of blue; it requires the following kind of context in order for it to be acceptable.

(69)
A: kinoo Ken to Hanako to Taro ga kuru yotee datta kedo,
yesterday and and come plan COP:PST but

hutari sika konakatta
two only come:NEG:PST

'Ken, Hanako, and Taro were going to come yesterday, but only two came.'

B: dare wa kite, dare wa konakatta no?
who come who come:NEG:PST FP
'Who came, and who didn't?'

In the mini-discourse above, the speaker and the hearer share the set of individuals [Ken, Hanako, Taro] in the immediately preceding context of (69B), which makes the use of *wa* in (69B) appropriate. Given the nature of *wa* in the discussion above, Miyagawa regards *wa* as *set-anaphoric*; namely, *wa* requires an existing set of entities which it refers to. To further illustrate this point, Miyagawa gives the following data.

(70)
A: Taro to Hanako to dareka ga saakasu ni itta
and and someone circus to go:PST
'Taro and Hanako and someone went to the circus.'

hitori wa nizi ni itte, moo hutari wa sanzi ni itta
one 2-o'clock at go more two 3-o'clock at go:PST
'One went at 2, and the other two went at 3.'

B: *dare wa nizi ni itte, dare wa sanzi ni itta no?
who 2-o'clock at go who 3-o'clock at go:PST FP
'Who went at 2, and who went at 3?'

Although speakers A and B share the knowledge of the set [Taro, Hanako, someone] before B's utterance, the use of *wa* in (70B) is unacceptable. This is explained by the condition that each of the individuals in the shared set must be identifiable and *someone* in the set prevents the set from being totally identifiable.

The second condition of Miyagawa's follows from the set-anaphoricity of *wa*; every member of a set must be represented in the question.

(71)
A: kinoo Ken to Hanako to Taro ga kuru yotee datta kedo,
yesterday and and come plan COP:PST but

hutari sika konakatta
two only come:NEG:PST
'Ken, Hanako, and Taro were going to come yesterday, but only two came.'

B: *dare wa kita no?
who come:PST FP
'Who came?'

In (71B) above, the question represents only two people, namely, the people who came, out of the set of three individuals. However, this condition applies only to *wa* with a WH phrase.

(72)
A: kinoo Ken to Hanako to Taro ga kuru yotee datta kedo,
yesterday and and come plan COP:PST but

hutari sika konakatta
two only come:NEG:PST
'Ken, Hanako, and Taro were going to come yesterday, but only two came.'

B: Ken wa kita no?
come:PST FP
'Did Ken come?'

As is clear from the data above, the second condition in (68) implies that *wa* with a WH phrase must show an explicit contrast between members of a set by exhaustively picking them out; however, it is not the case with *wa* with a referential NP.

We have so far looked at the conditions under which *wa* can appear with a WH phrase. Although Miyagawa (1987) appeals to the notions of *shared knowledge* and *identifiability* for the use of *wa*-marked WH phrases, as found in his principles in (68), Miyagawa's observation can be accounted for by the notion of activation, more specifically the principle that *wa* should mark an active elements. First, as seen in the contrast between (71B) and (72B), there is a difference between a WH phrase and a referential NP as to the acceptability of *wa*: a referential NP can be marked by *wa*, while a WH phrase cannot, even though they both occur in the identical discourse context, i.e. (71A) and (72A). In (72B), *wa* marks the active element *Ken* since it is mentioned in (72A). In (71B), on the other hand, the element marked by *wa*, i.e. the WH phrase *dare*, is only a subset of the active set of individuals [Ken, Hanako, Taro]. In (71B), the set [Ken, Hanako, Taro] is

active as a whole, being mentioned in the immediately preceding context; however, the subset [Ken] is less active than the whole set since the preceding context in (71A) does not place only [Ken] in the center of attention. In other words, [Ken] in (71) may be activated but not center of attention, while [Ken, Hanako, Taro] as a whole should be the center of attention. Another crucial factor is the fact that the *wa*-marked WH phrase in (71B) is an *unspecified* subset in that it can be any individual of the whole set, which makes a sharp contrast with the referential NP *Ken* in (72B).

The discussion above requires an account regarding the second principle of Miyagawa's in (68): "Every member of this set must be exhaustively represented in the WH *wa* question." If *wa* cannot mark an unspecified subset of an active set of individuals, why does the use of *wa* become acceptable if the *wa*-marked WH phrases exhaustively represent the members of an active set of individuals? This point is illustrated by the contrast between (69), in which *wa* is acceptable with WH phrases that exhaustively represent the whole set of individuals, and (71), in which *wa* is not acceptable where the WH phrase represents only a subset of the whole. Again, the basic principle that *wa*-marked element is active is relevant here. In (71), the WH phrase represents only an unspecified subset of the whole set of individuals which is active. In (69), on the other hand, *wa* is acceptable because the WH phrases collectively represents the whole set of individuals; therefore, *wa*'s in a single utterance, i.e. (69B), collectively marks the active set of individuals [Ken, Hanako, Taro].

At this point, it is useful to compare the use of *wa* and *ga* with WH phrases to see how the use of each can be explained by the principles proposed in (64). As seen in (69), the proposition expressed by the predicate is active in a WH-*wa* construction. The speaker uses the WH-*wa* to identify an entity for an already-existing proposition; therefore, the proposition must be active at the point where the WH-*wa* question is uttered. However, this picture contradicts the principle in (64) since the predicate of a *wa* sentence should

not be active. In order to explain the use of *wa* here, we need to look at the use of *ga* with a WH phrase. Observe the following.

(73)

A: kinoo Ken to Hanako to Taro ga kuru yotee datta kedo,
yesterday and and come plan COP:PST but

hutari sika konakatta
two only come:NEG:PST
'Ken, Hanako, and Taro were going to come yesterday, but only two came.'

B: dare ga/wa kite, dare ga/wa konakatta no?
who come who come:NEG:PST FP
'Who came, and who didn't?'

Regarding the choice of *wa*/*ga* here, native speakers' preferences vary; some prefer *ga*, rejecting the use of *wa*, while the others accept *wa*; they do not have any preference as to the choice. In other words, *ga* is invariably acceptable in sentences such as (73B), while the native speakers' judgment varies as to the use of *wa*. The difference in acceptability of *wa* and *ga* here is useful to further evaluate the principles stated in (64). As discussed above, the use of *wa* is explained by the fact that the *wa*-marked WH phrases represent the active set of individuals, even though each WH phrase represents an unspecified subset of the whole set. In terms of the degree of activation for the predicate, however, the use of *wa* in (73B) is not appropriate by the principles in (64); the predicates in (73B) represent active open propositions, which is mentioned in the immediately preceding context, i.e. (73A): *X and Y came and Z didn't come*. Although *Z didn't come* is not explicitly mentioned in (73A), it is apparent by the explicit contrast made by the utterance. The discussion above indicates that sentences such as (73B) exhibit another logical possibility which is not covered by the principle in (64); namely, the element marked by *wa*/*ga* and the proposition represented by the predicate are both equally active, and *ga* is invariably preferred in this activation pattern, as in (73B). The principles in (64) and (65) are modified as in (74) and (75) respectively by adding the fourth possibility in which *ga* appears.

(74)

WA: A *wa*-marked referent is more mutually active than the open proposition with which the referent combines.

GA: A *ga*-marked referent is less mutually active than the open proposition with which the referent combines or they are equally mutually active or nonactive.

(75)

a.	A > B *wa*	A is more active than B.
b.	A < B *ga*	B is more active than A.
c.	A = B = active *ga*	A and B are equally active.
d.	A = B = nonactive *ga*	A and B are equally nonactive.

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

In Chapter 3, I will further evaluate the principles above in terms of data from conversational Japanese and show that the activation pattern added above is in fact fairly common in the conversational data base of the present study.

2.4.3. Reexamining *Thematic Wa* and *Contrastive Wa*

As discussed earlier in this thesis, a number of people have argued against Kuno's distinction between thematic *wa* and contrastive *wa*, claiming instead that there is a single fundamental function of *wa*. One of these arguments has been made by Miyagawa (1987). As seen in the previous section, Miyagawa (1987) states that *wa* is set-anaphoric; *wa* must have an existing set of entities to refer to in the preceding context. This set anaphoricity of *wa* is also the case in non-WH constructions. Miyagawa cites the following to illustrate this point.

(76) Taro wa uchi e kita kedo, Hanako wa konakatta
house to come:PST but, come:NEG:PST
'Taro came to (my) house, but Hanako didn't come.'

(77) Taro wa uchi e kita
house to come:PST
'Taro came to (my) house.'

Under the same discourse context in which Taro and Hanako contextually comprise an identifiable set, (76) exhaustively refers to the set [Taro, Hanako], while in (77) only one of the members is explicitly picked out. Even though only 'Taro' is mentioned in (77), *Taro wa* evokes the whole set and implies that Hanako did not come to the speaker's house. If (77) is uttered in a context in which the set consists of only 'Taro', the *wa* becomes *thematic* since the entire set is referentially represented. Miyagawa (1987:200) explains the difference between the thematic reading and the contrastive reading of *wa* in the following way.[16]

> If only a portion of the set is referentially picked out, it is put in contrast to the other members, resulting in a contrastive reading. If, on the other hand, the members of the set are exhaustively referentially picked out, and are associated with the same property, the result is thematic *wa*.

As seen in Section 2.3.2, the thematic *wa* is the same as the contrastive *wa* in terms of the activation pattern; namely, the entity marked by *wa* is active, while the predicate is not active. In this respect, there is a striking contrast between the function of *wa* and that of *ga*. As Miyagawa explains above, *wa* picks out some active entity, either out of many in a set or the only one active entity in the context, in order to contribute to the communication by adding a nonactive or less active entity in the rest of the sentence. In this regard, the contrastive *wa* is not different from the thematic *wa;* they both play the same role of picking out an active entity to add some predication to it. The contrast between thematic and contrastive seems to arise by the contextual effect, as Miyagawa

suggests above. This point is illustrated by the fact that the same sentence can be either thematic or contrastive depending the preceding context.

(78)
1. kyonen Boston ni itta
   last-year to go:PST
   '(I) went to Boston last year.'

2. Boston wa totemo samukatta
   very cold:PST
   'It was very cold in Boston.'

(79)
1. kyonen Boston to Chicago ni itta
   last-year and to go:PST
   '(I) went to Boston and Chicago last year.'

2. Boston wa totemo samukatta
   very cold:PST
   'It was very cold in Boston.'

In the preceding context of (78.2), there is only one active place referent 'Boston'; therefore, *wa* in (78.2) pick out the only choice, which gives this *wa* the thematic sense. In (79), on the other hand, there are two equally-active place referents in the preceding context of (79.2) and *wa* picks out only one out of the two, which gives the contrastive reading, i.e. gives the implication that it was not very cold in Chicago, since the predication *it was very cold in X* is made only for the entities which was picked out, not for the other. This may suggest that the sense of thematic/contrastive of *wa* depends not only on the number of competitive entities in a set, in Miyagawa's (1987) sense, but also on the degrees of activation on each competitive entity. Observe the following.

(80)
1. kyonen Boston ni itta

last-year to go:PST
'(I) went to Boston last year.'

2. hurui ie o takusan mita
old house OBJ many see:PST
'(I) saw many old houses.'

3. sorekara Chicago ni itta
then to go:PST
'Then (I) went to Chicago.'

4. Chicago wa totemo samukatta
very cold:PST
'It was very cold in Chicago.'

Wa in (80.4) is not as contrastive as (79.2); (79.4) does not imply that Boston was not very cold as strongly as (79.2).[17] There are two competing active place referents in (80); however, two entities have different degrees of activation in the immediately preceding context of (80.4). Given the basic assumption that an entity normally gradually decays in activation unless it keeps being referred to (Dryer 1994), we can assume that 'Chicago' is more active than 'Boston' in the immediately preceding context of (80.4) since 'Chicago' is the immediate mention, while 'Boston' was last mentioned three clauses before.

In summary, thematic *wa* and contrastive *wa* in Kuno (1972, 1973) manifests the same function of marking active information. These two different effects are caused by external factors such as the number of competing entities and the degrees of activation of each competing entity. Moreover, the thematic *wa* and the contrastive *wa* are the two extremes on the same scale. We should not regard them as two discrete functions of *wa*. Rather, there are different degrees of contrastiveness in *wa*, just like different degrees of activation, which are affected by various contextual factors.

2.5. CONCLUSION

The present chapter examined the use of *wa* and *ga* on the mini-discourse level, mostly in question-answer pairs, and have demonstrated that the notion of activation,

along with the concept of relative degrees of activation, accounts for the use of *wa* and *ga* in question-answer pairs, where the traditional approaches in terms of old/new, identifiable/non-identifiable, and predictable/unpredictable information face a problem. Furthermore, the account of the use of *wa* and *ga* on the minimum-discourse level is crucial due to that fact that there is only *local* contextual information available in question-answer pairs where the notions in terms of global discourse structures, such as *thematicity*, have a little role to play.

Given the concept of relative degrees of activation, I defined the functional contrast between *wa* and *ga* in terms of a *wa*/*ga*-marked entity and the open proposition which the entity fills, which is repeated as (81).

(81)

WA: A *wa*-marked referent is more mutually active than the open proposition with which the referent combines.

GA: A *ga*-marked referent is less mutually active than the open proposition with which the referent combines or they are equally mutually active or nonactive.

The contrast between *wa* and *ga* stated in (81) is schematically illustrated in Figure 4 in terms of relative degrees of activation. Line (a) shows the activation pattern for *wa*, which indicates that a *wa*-marked referent is more active than the open proposition. Lines (b) and (c) shows the two extremes of activation pattern for *ga*; (b) shows one extreme where the *ga*-marked referent and the open proposition are almost equally active or nonactive; however, the *ga*-marked referent may be a little more active than the open proposition, and (c) shows the other extreme where an open proposition is more active than the *ga*-marked referent. (b) is exemplified by the cases in which while an open proposition is nonactive, the *ga*-marked entity is active but not the center of attention, as seen in (54). There are in-between cases between the two extremes (b) and (c), including those in which a *ga*-marked entity and the open proposition are both equally active or nonactive.

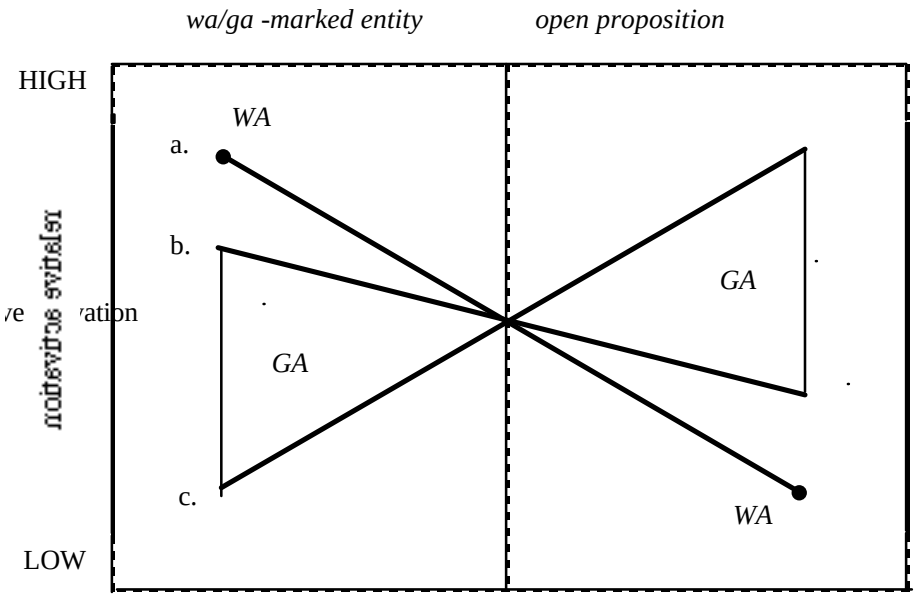


Figure 4: The contrast between *wa* and *ga* in terms of relative degrees of activation

Figure 5 shows the three basic activation patterns of *ga* in terms of relative degrees of activation. (a) indicates the pattern that a *ga*-marked entity and the open proposition are equally nonactive, while (b) indicates the pattern that an open proposition is more active than the *ga*-marked entity. (a) and (b) are labeled as *neutral description ga* and *exhaustive listing ga* respectively in Kuno's (1972, 1973) term and *sentence focus* and *argument focus* respectively in Lambrecht's (1986, 1992, 1994) term.[18] (c) shows the third pattern that a *ga*-marked entity and the open proposition are equally active, which is exemplified by the cases in which WH-phrases may be marked by *wa* and *ga* interchangeably, as seen in Section 2.4.2. This particular pattern of *ga* is further examined on the basis of the conversational Japanese data in the following chapter.

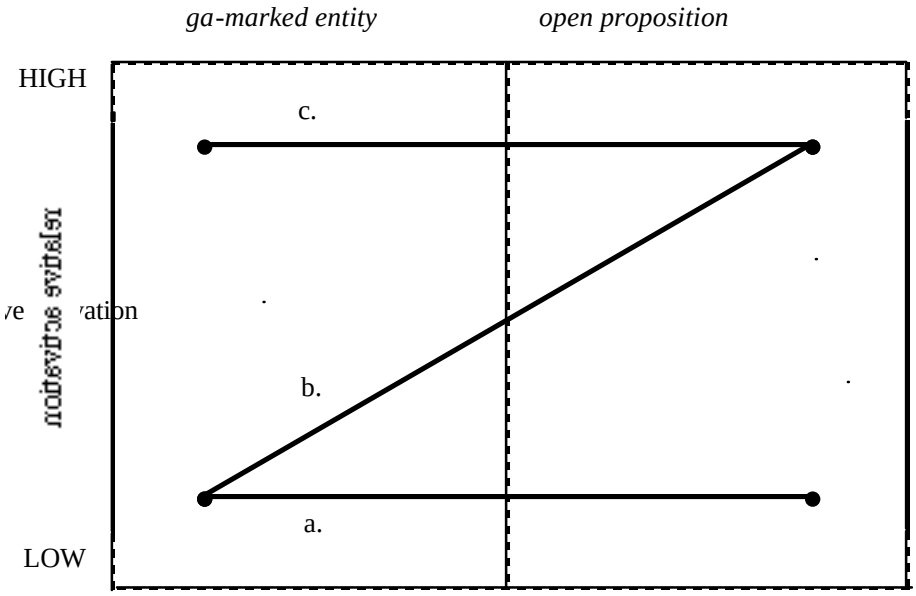


Figure 5: The three basic activation patterns of *ga*

Having laid out the contrast of *wa* and *ga* in terms of activation above, *wa*/*ga* sentences are differentiated in terms of the notion of focus as in (82), on the basis of the assumption that non-focus is associated with activation.

(82)

WA: A *wa*-marked element is non-focus and it combines with the open proposition which is the focus of the utterance.

GA: a. A *ga*-marked element and the open proposition which it combines with are both the focus of the utterance.

b. A *ga*-marked element is focus and it combines with the open proposition which is non-focus of the utterance.

c. A *ga*-marked element and the open proposition which it combines with are both non-focus of the utterance.

The characterization of *wa* and *ga* above is shown schematically in Figure 6 below. In essence, the contrast between *wa* and *ga* is shown by the fairly clear markedness distribution; *wa* is marked, while *ga* is unmarked in that *wa* occurs only in a particular condition as described in (59), while *ga* occurs in a variety of focus patterns, as in (59a)-(59c). It appears odd to say that the whole sentence is non-focus with the third type of *ga*

in (59c); however, this only indicates that in this focus pattern the *ga*-marked element and the open proposition are both active in the immediately preceding context and this is not to say that there is no focal element in sentences of this focus pattern. As discussed in Section 1.2, among active elements there is a difference in relative degrees of activation; one element may be even more highly active than the other active element, i.e. the former is the center of attention, while the latter is not, though they are both active. In (59c), the focus of the sentence is determined by this subtle difference in relative degrees of activation. The *ga*-marked element is the focus of the sentence if the open proposition is the center of attention, while the open proposition is the focus of the sentence if the *ga*-marked referent is the center of attention. In this way, the focus pattern in (59c) is neutral since it is flexible in determination of the focal element of the sentence, depending on the difference in relative degrees of activation among active elements and propositions. I will henceforth call the focus type in (59c) *neutral focus* and sentences of this focus type *neutral-focus sentences*. This focus pattern along with the other types will be further discussed in the following chapters.

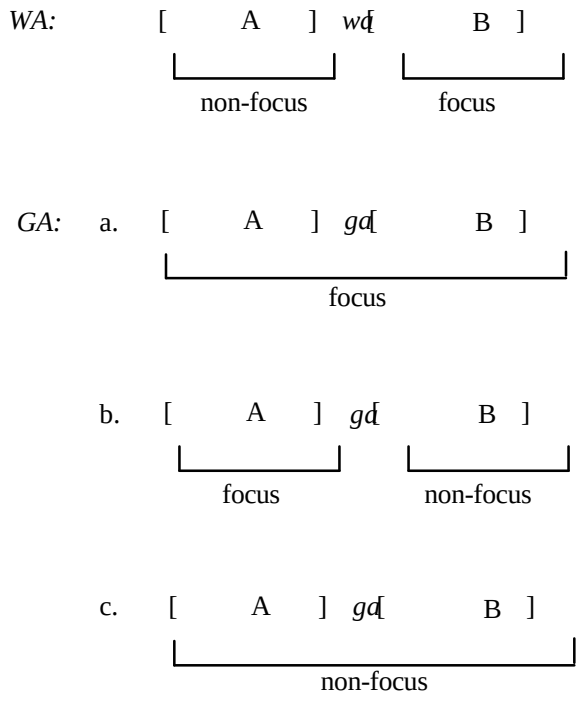


Figure 6: Focus patterns of *wa*/*ga* sentences

CHAPTER 3

WA AND *GA*: QUANTITATIVE ANALYSIS OF CONVERSATIONAL JAPANESE

3.1. INTRODUCTION

The present chapter will investigate the use of *wa* and *ga* quantitatively on the basis of data from conversational Japanese and examine whether the functions of *wa* and *ga* proposed in Chapter 2 are supported empirically. In order to measure quantitatively the degrees of activation, this chapter utilizes the *referential distance* [RD] measurement, one of the major quantitative analysis methods found in Giv n's work (cf. Giv n (ed.) 1983). In the RD measurement, the number of clauses is counted backward to the most recent appearance of the coreferential expression. Thus, an RD of 1 clause indicates that a referent in question in the present clause was mentioned in the immediately preceding clause. Given the notion of referential distance, this study assumes that the RD measurement is a useful way to measure the degrees of activation quantitatively, though the previous mention of an identical NPs is not the only factor which makes an element in question activated.[19] On the basis of the assumption that an entity becomes activated and deactivated, as discussed in the previous chapter, the present study assumes that an entity of recent mention is more active than an entity of non-recent mention; therefore, the smaller the RD is, the more activated an element is.

In addition to the measurement of RD of *wa*/*ga*-marked elements, this study will measure the RD for the predicate in *wa*/*ga* clauses, i.e. an open proposition represented by the predicate.[20] The measurement of RD for propositions has not been done in the literature before; however, it will provide a useful clue to the investigation of *wa* and *ga* on the basis of the activation status of elements predicating the *wa*/*ga*-marked elements.

Following the quantitative analysis methodology discussed in Section 3.2, Section 3.3 presents the results from the RD measurement for *wa* and *ga* in matrix clauses. The

results will be discussed to examine *wa*-marked subjects and *wa*-marked non-subjects, and also *wa*-marked subjects and *ga*-marked elements, i.e. subjects. Given the RD measurement of *wa*/*ga*-marked elements, the results will be presented to examine the activation status of predicates of *wa*/*ga* utterances. Section 3.4 discusses the RD measurement results for *wa* and *ga* in embedded clauses. The results will be presented for *wa*-marked subjects, *wa*-marked non-subjects, *ga*-marked subjects, and predicates of *wa*/*ga* embedded clauses. Section 3.5 discusses the cases in which *wa* marks elements of large RD to investigate the cases which the RD measurement does not capture. Finally, the discussion of this chapter is summarized in Section 3.6.

3.2. QUANTITATIVE ANALYSIS METHODOLOGY

The data investigated for the present section is the conversation text which I transcribed from a TV talk show, which is broadcast daily nation-wide in Japan. In this show, a female host interviews a guest of the day for about 35 minutes. The database for this section consists of the transcribed texts from four different days, which is about two hours and twenty minutes in total. The talk show guests in the database are four native speakers of Japanese, two males and two females, and spoke standard Japanese in the show. The conversation data here may not be labeled as *natural* conversation in conversation analysts' sense in that the host leads the conversation on the basis of a roughly predetermined set of topics.[21] In spite of this topical guideline, however, the conversation in this talk show exhibits the nature of *natural* conversation to a large extent in that there is no predetermined conversational turns and the participants are free to give spur-of-the-moment utterances at any time; therefore, I assume that the investigation of the use of *wa* and *ga* in the database here captures the picture of *wa* and *ga* in *natural* conversational Japanese without the presence of observers.

For the RD measurement of this study, I used the following guidelines applied to judgments as to what to include and what not to in the database. The following instances of *wa*/*ga* were not included in the database.

(i) *Wa*/*ga* marking pronouns

(ii) *Wa*/*ga* in direct quotations

(iii) *Wa*/*ga* marking a WH phrase

(iv) *Wa* marking a temporal expression (e.g. 'today', 'now', 'that time', 'until now', 'first') which is not the subject of the clause

(v) *Wa*/*ga* in utterances which repeat utterances in the immediately preceding context

The use of pronouns is triggered by the active antecedent referents; pronouns are typically used to refer to an active entity or proposition; therefore, the RD for *wa*/*ga*-marked pronouns will not be considered to exclude a possible independent factor, i.e. the correlation between the use of pronouns and *wa*/*ga*.

Wa/*ga* in direct quotations were excluded since the reported speech is from a different discourse. Quotations were labeled as "direct" if (i) there is a change in referring expressions in terms of the speaker's perspective (e.g. a switch of pronouns from 'you' to 'I') in the reported speech, and (ii) there is a sentence-final particle (i.e. *ne*, *yo*, *wa*, etc.) at the end of the reported speech.

Wa/*ga* with a WH phrase was not included in the RD measurement since *wa* can appear with a WH phrase only in a restricted context, as discussed in Chapter 2.

Temporal expressions were not included unless they are subject of the sentence, i.e. if they can be marked by *ga*. Non-subject temporal expressions were not included because they are typically marked by *wa* regardless of their RD status. According to my preliminary RD count, there are 83 non-subject temporal expressions marked by *wa* in the database, whose token distribution is illustrated by Table 1.[22] As shown in the table, a total of 77% of the non-subject temporal expressions marked by *wa* are either of RD 20+

or of the first mention, and this results suggest that the presence or absence of preceding co-referential expressions is not much responsible for the use of *wa* for temporal expressions. The use of *wa* for temporal expressions will be further discussed in Section 3.5 below.

Table 1: Token distribution of *wa*-marked non-subject temporal expressions in terms of RD

RD	# of *wa* tokens	
1	8	(10%)
2 - 10	8	(10%)
11 - 19	3	(4%)
20+	20	(24%)
NPM	44	(53%)
Total	83	(100%)

Finally, the *wa*/*ga* tokens in utterances which repeat utterances in the immediately preceding context were not included. I regarded an utterance as repetition if the repeated utterance contains (i) the same NP and predicate as in the previous utterance and (ii) the same particle, i.e. *wa* or *ga*, is repeated. Given these criteria, the repeated utterances were typically the repetition of utterances within no more than two clauses back, and interestingly there was no repetition token for *wa* in the database.[23]

Although those categories of *wa*/*ga* above were not included in our database, noun phrases marked by *wa*/*ga* in the above categories were treated as preceding co-referential expressions in measuring the RD. Also, zero anaphora was counted as a preceding co-referential expression if such was the case.

For identification of utterance boundaries, the transcribed text was divided into units followed by a pause. Pause can be identified relatively clearly in the conversation data, as discussed in Maynard (1989, 1993); therefore, it is a reliable criterion for identifying the utterance boundary.[24] However, pause alone leads to morphosyntactically unnatural divisions in some cases; therefore, I used the following as secondary criteria. An

utterance string was treated as one unit if both of the following criteria were met, even though there was a clear pause within the string.

(i) The utterance makes one clausal unit.

(ii) There is no clause-final falling pitch within the clausal unit.[25]

An unnatural division of units typically arises when a pause divides a morphosyntactically contiguous unit, as shown in (1). (The pause boundary is shown by a slash [/].)

(1) S-san no gekidan/ wa ima made K-san ga irasita kedo
GEN th.-company now until belong:HON:PST but
"Ms. K was in Mr. S's theatrical company until now, but..."

In (1), the pause divides the *wa* and the NP which the *wa* marks, which makes the unit boundary unnatural. Also, there is no clause-final falling pitch on the phrase preceding the pause.[26]

The same set of criteria above was applied to the cases in which units are accompanied by post-verbal elements. Namely, the post-verbal elements were considered to constitute a single unit with the preceding elements if there is no intervening pause between the verb and the post-verbal elements. If there is a pause preceding the post-verbal elements, I regarded the post-verbal unit as separate from the preceding elements.

Given the utterance-boundary identification procedure described above, a large number of units were already smaller than a clause. Since a clause is the analyzing unit in the RD measurement, an independent phrase was counted as a clause. Units containing more than one clause, typically one pause unit made up of more than one simple sentence or a complex sentence, were further divided into clausal units at this point. I used the following guideline for divisions of complex sentences.

A clause coordinated with a following clause by conjunctive morphemes, e.g. *ga* 'but', *si* 'and', *tari* 'and', was counted as an independent clause token. For example,

(2) [zenbu anoo hoteru wa kimatteru si]

all uh hotel being-arranged and

[oshokuzi mo kimatteru wake deshoo]
meals also being-arranged reason COP
'Therefore, (they) chose all the hotels and arranged the meals too.'

An adverbial subordinate clause followed by conjunctive morphemes, such as *kara* 'because', *node* 'because', *toki(ni)* 'when', *nagara* 'while', *aida(ni)* 'while', *ato(ni/de)* 'after', *mae(ni)* 'before', *to* 'when/if', *ba* 'if', and *made* 'until', *madeni* 'by', was also counted as an independent clause token. For example,

(3) [demo yappari ano hito wa daikantoku desu kara]
but as-expected that person v.-director COP because

dokkade monosugoi chikara ga detekichau
somewhere marvelous power come-out
'But, because that person is a virtuoso director as you know, (his) marvelous skills are shown somewhere (in his movie).'

Clauses which modify a head noun and nominalized clauses were regarded as modifiers of nouns; therefore, they were not counted as independent clause tokens. (4) and (5) show an example of each respectively.

(4) gaikoku e ikimasu to
f.-country to go if

warito [tochuu de deru] hito iru n desu nee
quite-a-few middle in leave people exist NOM COP FP
'If you go to foreign countries, (you find) quite a few people who leave (the theater) in the middle of a movie.'

(5) tonikaku [atasino kao to mikurabeteru] no ga
anyway my face with comparing NOM

zutto damatte bukimina shunkan desita
all-through quiet weird moment COP:PST
'Anyway, (his) comparing may face with (the picture) went on in silence and was a weird moment.'

I also considered complement subordinate clauses dependent on the matrix verb, such as 'say' and 'think'; therefore, they are not counted as an independent clausal unit. By

the same token, direct quotation was regarded as a part of the clausal unit of the matrix clause, even though the reported speech included a complex sentence.

In addition to the conjunctive morphemes as listed above, Japanese has *te*-connective, the extensively-used connective form of verbs and adjectives.[27] *Te* can connect elements of any syntactic level; it can link two verbal elements to form a complex predicate, as in (6), as well as to link phrases/clauses, as in (7).

(6) atakusi wa zutto anaunsaaruumu de taiki o si**te** imasita yo
I continuously announcer-room in stand-by OBJ do be:PST FP
'I was continuously standing by in the announcer room.'

(7) ima supittsu atarasiku kat**te**
now spitz newly keep

moo hooboo kizudarakena n desu tte
already everywhere full-of-scratch NOM COP QT
'(I heard you) got a new spitz, and there are scratches everywhere.'

For the purpose of identifying the RD measurement unit, this study utilizes Hasegawa's (1992) analysis of the syntactic level of the *te*-linkage.[28] According to Hasegawa, if each linked clause has its own syntactic subject, the linkage is at the clause level. (8) is from Hasegawa (1992: 190).

(8) Joan ga gitaa o hii**te** Hiro ga utau
guitar OBJ play sing
'Joan will play the guitar, and Hiro will sing.'

However, given the fact that Japanese allows extensive ellipsis, subject NPs are not necessarily present; both of the linked clauses may lack an overt subject, even though the two subjects have disjoint reference; therefore, the linkage is at the clause level. In the case of *te*-linkage, however, Hasegawa (1992: 191) claims that if the linked clauses lack at least one overt subject, the subjects are coreferential. On the basis of this claim, Hasegawa gives the following generalization.[29]

(i) If both subjects are overt, they have disjoint reference; therefore, the linkage is at the clause level.

(ii) If either or both of the subjects is missing, they are coreferential; the linkage is at the core level.

For the purpose of identifying the RD measurement unit for this study, I used the same criteria as above; namely, the linkage is at the clause level, i.e. the sentence as a whole contains two clause tokens for the RD measurement if the subjects have disjoint reference, while the linked units as a whole count as one clause token if the subjects are coreferential. The judgment of the syntactic level was fairly easy to make, given the contextual information provided by the conversation data.

After going through the procedure for identification of the clausal unit described above, the database still contained every utterance from the original conversation text. As a final step of preparing the text for the RD measurement, I excluded *back-channel* utterances in that those utterances were ignored in counting the number of clauses back to the preceding co-referential expression. Back-channel utterances are in general propositionally empty utterances made by the hearer during the speaker's conversation turn; therefore, their effect on the activation degree of active propositions in the preceding conversational context is considered to be minimum. For identifying the back-channel utterances, I followed the characterization given in Yngve (1970) and Maynard (1993); namely, Yngve (1970: 568) states that there is back channel when "the person who has the turn receives short messages such as 'yes' and 'uh huh' without relinquishing the turn" and Maynard (1993: 58) elaborates this by giving the definition that "back channel is short messages (including nonverbal behavior) which the listener sends during the interlocutor's speaking turn, and the short messages for which the interlocutor shows a reaction of relinquishing the speaking turn are not regarded as back channel." Furthermore, Maynard adds that the back channel behavior includes a short message which is sent immediately following the speaker's turn and is followed by a pause before the listener takes the next speaking turn since it is assumed to be a response to the immediately preceding turn.

Following the characterization of the back channel above, the following criteria were used for identifying the back channel. Utterances were excluded from the database for the RD measurement (i) if they are uttered by the listener during the interlocutor's speaking turn, (ii) if they are uttered by the listener immediately following the speaker's turn and followed by a pause, i.e. if they seem to be simply a response corresponding to the immediately preceding turn. Those utterances excluded by the criteria above were typically short responses to the speaker's utterances such as, *hai* 'yes', *ee* 'yes', *un* 'yeah', *soodesu* '(That's) right', *naruhodo* 'I see', which are typically listed as back channels in studies of Japanese conversation (e.g. Clancy 1982, Mizutani 1983).

Finally, I also excluded short utterances, typically the same kinds of expressions as back channels described above, which *fill* the gap between speaking turns, since those utterances do not carry relevant propositional meanings.[30] To illustrate this point, (9) shows an example of filler utterances which were not included in the RD measurement. There is a pause after (B2) and (A2) and the speakers fill the gap by the short utterances (A2) and (B3).

(9)

A1: de otoosama wa hantainasaranakatta no sono zyoyuu ni naru toki
and father oppose:HON:NEG:PST FP that actress RE become time
'Wasn't your father opposed to your becoming an actress?'

B1: ee monosugoku hantaisimasita nee
yeah terribly oppose:PST FP
'Yes, (he) opposed terribly.'

B2: demo yoku wakarimasu yo sore wa
but well understand FP it
'But (I) can understand it well.'

A2: naruhodo nee
I-see FP
'I see.'

B3: ee
yeah

'Yeah'

A3: de zyoyuu ni onarininatte kara
then actress RE become:HON since
'Then, since (you) became an actress,'

3.3. *WA* AND *GA* IN MATRIX CLAUSES: MEASUREMENT RESULTS

There are two types of sentence elements whose RD is examined in this section: (i) the referent or proposition expressed by *wa*/*ga* marked elements, and (ii) the proposition expressed by the predicate of sentence including the *wa*/*ga* elements. This section further divides the former into two categories: RD for *wa*-marked subjects and RD for *wa*-marked non-subjects. First, the results will be presented in order to compare *wa*-marked subjects and *wa*-marked non-subjects in terms of RD. Then, the RD measurement results for the *ga*-marked elements will be presented in comparison with the *wa*-marked subjects to examine the contrast of the two in terms of the token distribution. Finally, I will discuss the results from the RD measurement for predicate of sentences with *wa*/*ga*.

3.3.1. RD for *Wa*-marked Subjects and *Wa*-marked Non-subjects

Table 2 shows the token distribution of *wa*-marked subjects and non-subjects in terms of RD. Among the total of 241 *wa*-marked elements in matrix clauses, there are 142 *wa*-marked subjects and 99 *wa*-marked non-subjects in the database. Given the activation pattern of *wa* sentences discussed in Section 2.4.1, we expect a pattern that *wa*-marked elements are active, hence small in RD. The expected pattern is reflected in both *wa*-marked subjects and *wa*-marked non-subjects. Table 3 summarizes the pattern by showing the token distribution in terms of the five categories of RD: 1 through 4, 5 through 10, 11 through 19, 20 and over, and no previous mention. For *wa*-marked subjects, we find the expected pattern in the following respects. A total of 61% occurs in the range of 1 through 4 of RD, and 10% in the range of 5 through 10 of RD. There are only 5% and 6% of *wa*-marked subjects in the range of both 11 through 19 and 20+ respectively. The same

pattern is found in *wa*-marked non-subjects. Namely, there are 48% of *wa*-marked non-subjects for the RD of 1 through 4, and 19% for 5 through 10. For the RD of 11 through 19 and 20+, on the other hand, there are only 7% and 8% of the cases respectively.

Although the overall pattern in the results is what we expect from the principle that *wa*-marked elements are active, hence small in RD, there is one puzzling point; namely, the occurrences of *wa* with high RD, i.e. 20+ and NPM. There are 19% of the *wa*-marked subjects and 18% of the *wa*-marked non-subjects for the RD of NPM, which are both considerably higher than the percentage for the RD of 20+, i.e. 6% and 8% respectively, which in turn are higher than the percentage for the mid-range of RD, e.g. RD 10 through RD 19. Discussion will be given regarding those of high RD in Section 3.5 below.

Table 2: Token distribution of *wa*-marked subjects and non-subjects in terms of RD

RD	*wa*:	subject	*wa*:	non-subject
1	52	(37%)	27	(27%)
2	14	(10%)	15	(15%)
3	11	(8%)	5	(5%)
4	9	(6%)	5	(5%)
5	3	(2%)	4	(4%)
6	7	(5%)	0	
7	1	(1%)	3	(3%)
8	1	(1%)	5	(5%)
9	0		2	(2%)
10	2	(1%)	3	(3%)
11	1	(1%)	0	
12	0		1	(1%)
13	1	(1%)	1	(1%)
14	2	(1%)	1	(1%)
15	1	(1%)	0	
16	2	(1%)	0	
17	0		1	(1%)

18	0		0	
19	0		0	
20+	8	(6%)	8	(8%)
NPM	27	(19%)	18	(18%)
Total	142	(100%)	99	(100%)

Table 3: Token distribution of *wa*-marked subjects and non-subjects in terms of RD: 1-4, 5-10, 11-19, 20+, and NPM, and % by total for subjects and non-subjects

RD	*wa*:	subject	*wa*:	non-subject
1 - 4	86	(61%)	47	(48%)
5 - 10	14	(10%)	19	(19%)
11 - 19	7	(5%)	7	(7%)
20+	8	(6%)	8	(8%)
NPM	27	(19%)	18	(18%)
Total	142	(100%)	99	(100%)

3.3.2. RD for *Wa*/*ga*-marked Subjects

This section gives the results for *wa*-marked subjects and *ga*-marked elements, i.e. subjects, in order to compare the token distribution of the two in terms of RD. Tables 4 and 5 illustrate the token distribution and the percentage of *wa* and *ga* in terms of each RD; therefore, the percentages show the choice that speakers make for whether to mark a subject with *wa* or *ga*. Given the functional contrast between *wa* and *ga* discussed in the previous chapter, we expect that nonactive elements, i.e. elements of high RD, are marked with *ga*, and this expected tendency is reflected in the results. Table 5 summarizes the token distribution in terms of the four categories of RD: 1 through 4, 5 through 19, 20 and over, and no previous mention. In terms of the percentage, the table shows a clear tendency that the larger the RD, the more *ga* tokens occur; namely, 37% for 1 through 4 of RD, 50% for 5 through 19, 69% for 20 and over, and 89% for no previous mention. Conversely, there is a tendency that the smaller the RD, the more *wa*-marked subjects occur; 63% for 1 through 4 of RD, 50% for 5 through 19, 31% for 20 and over, and 11% for no previous mention. The results here indicate an overall tendency that *wa*-marked elements are more active than *ga*-marked ones.

Table 4: Distribution of *wa*/*ga*-marked subjects in terms of RD and % by total for each RD

RD	*wa*		*ga*		Total	
1	52	(65%)	28	(35%)	80	(100%)
2	14	(56%)	11	(44%)	25	(100%)
3	11	(58%)	8	(42%)	19	(100%)
4	9	(75%)	3	(25%)	12	(100%)
5	3	(33%)	6	(67%)	9	(100%)
6	7	(54%)	6	(46%)	13	(100%)
7	1	(33%)	2	(67%)	3	(100%)
8	1	(33%)	2	(67%)	3	(100%)
9	0		1	(100%)	1	(100%)
10	2	(100%)	0		2	(100%)
11	1	(50%)	1	(50%)	2	(100%)
12	0		2	(100%)	2	(100%)
13	1	(100%)	0		1	(100%)
14	2	(100%)	0		2	(100%)
15	1	(100%)	0		1	(100%)
16	2	(100%)	0		2	(100%)
17	0		0		0	
18	0		0		0	
19	0		1	(100%)	1	(100%)
20+	8	(31%)	18	(69%)	26	(100%)
NPM	27	(11%)	212	(89%)	239	(100%)
Total	142		301		443	(100%)

The results also reflect an expected pattern in that *ga* marks elements of small RD as well as elements of high RD. In spite of the fairly clear preference for *ga* over *wa* for no previous mention, i.e. 11% for *wa* vs 89% for *ga*, the difference is less clear for 1 through 4 of RD, i.e. 63% for *wa* vs 37% for *ga*.[31]

Table 5: Distribution of *wa*/*ga*-marked subjects in terms of RD: 1-4, 5-19, 20+, and NPM

RD	*wa*		*ga*		Total	
1 - 4	86	(63%)	50	(37%)	136	(100%)
5 - 19	21	(50%)	21	(50%)	42	(100%)
20+	8	(31%)	18	(69%)	26	(100%)
NPM	27	(11%)	212	(89%)	239	(100%)
Total	142		301		443	(100%)

In order to fully evaluate the activation patterns of *wa*/*ga* sentences, we need to examine the RD of predicates as well as the *wa*/*ga*-marked elements. The principles discussed in the previous chapter predict that *wa*-marked elements occur with predicates

of large RD, while there are three different patterns with *ga*: *ga*-marked elements of small RD occur with predicates of small RD, while it is not necessarily the case with *ga*-marked elements of large RD; their predicates may be either high or low in RD. In order to see if the expected trend is reflected in the data, the following section examines the RD of predicates.

3.3.3. RD for Predicates

As noted earlier, in the previous studies the RD measurement has been applied only to referents of noun phrases, not to propositions represented by predicates. For the purpose of this study, however, the RD measurement of propositions, more specifically open propositions represented by a predicate, becomes very useful in order to measure the degree of activation of predicates in order to capture the contrast between *wa* and *ga*.

In the RD measurement for this section, as in the cases for the *wa*/*ga*-marked elements, the RD was measured for the propositions which predicate the *wa*/*ga*-marked elements. In the case of *wa*/*ga*-marked elements, the elements marked are mostly NPs, i.e. phrases headed by a single noun or often a single noun alone; therefore, the RD count was made by the preceding co-referential NPs. In the RD count in this section, on the other hand, the elements which predicate *wa*/*ga*-marked elements consist of propositions; therefore, the RD for propositions was measured by tracing back to the preceding antecedent proposition. In (10.4), for example, *ga* marks the noun 'water' and combines with the open proposition *there is no X*. For this open proposition, there is an preceding antecedent open proposition in (10.1); therefore, the RD for the open proposition in (10.4) is 3.

(10)

1. nanimo nai to kiite wa ita nde
   anything exist:NEG QT hear be because
   'Because I heard that there is nothing (available).'

2. raamen toka raamenpotto toka mottette

noodle and noodle-pan etc. bring
'(I) brought noodles, pans, etc.'

3. heya de wa tsukuru n desu kedo
room in cook NOM COP but
'(I) cooked (noodles) in the room, but'

4. omizu ga nai n desu ne
water exist:NEG NOM COP FP
'There is no water.'

There are cases in which an open proposition does not share the same lexical forms with the antecedent proposition but they share the same propositional content. For example, the utterance in (11.3) has the open proposition *X ga sensee yattemasita* '(lit.) X was doing a professor' and this proposition shares the same propositional content with the proposition expressed by the whole utterance in (11.2), i.e. *there were good professors*. In other words, they both assert the same proposition that there were professors; therefore, the RD for the open proposition in (11.3) is 1.

(11)
1. sensee ippai imasita
prof. many exist:PST
'There were many professors.'

2. ii sensee imasita
good prof. exist:PST
'There were good professors.'

3. T nante hito ga sensee yattemasita
such-as person prof. doing:PST
'There were professors such as Mr. T.'

An open proposition of an utterance can consist of a number of *subset* open propositions. For example, the open proposition predicating the subject *Taro wa* in (12), i.e. the underlined part, can be divided into three subsets: (i) *X is a student*, (ii) *X is from Japan*, and (iii) *X is a student from Japan*.

(12) Taro wa <u>nihonzin no gakusee desu</u>
Japanese GEN student COP
'Taro is a student from Japan.'

In the RD measurement for the present study, a proposition is regarded as the antecedent only if the proposition is explicitly mentioned in the preceding discourse and it contains the whole open proposition predicating the *wa*/*ga*-marked entity in question. In the case of (12) above, only the open proposition *X is a student from Japan* or a proposition which includes this open proposition is counted as the antecedent proposition, and neither *X is a student* nor *X is from Japan* is counted as the antecedent since it is only a subset of the open proposition in question. For example, the utterance in (13B2) has the open proposition *X is the photograph which I carry with me all the time*. (13A) has the open proposition *X is a very small photograph*; however, this proposition contains only the subset of the open proposition in (13B2), i.e. *X is a photograph*, hence, the open proposition in (13A) does not count as an antecedent of the open proposition in (13B2).

(13)

A: sugoi chicchai shasin na no ne kore
very small photograph COP NOM FP this
'This is a very small photograph.'

B1: iya okkii no mo aru n desu
no big NOM too exist NOM COP
'No, there is a big one too.'

B2: kore wa watasi ga tsuneni keetaisiteiru shasin
this I always carrying photograph
'This is the photograph which I carry (with me) all the time.'

Given the criteria for the RD measurement for open propositions, I discuss the measurement results below. Given the principles in (75) in Section 2.4.2, which is repeated as (14) below, we expect that the RD for predicates in *wa* clauses is large, while the RD for those in *ga* clauses may be either small or large.

(14)

a. A > B (*wa*) — A is more active than B.

b. A < B (*ga*) — B is more active than A.

c. A = B = active *ga* — A and B are equally active.

d. A = B = nonactive *ga* — A and B are equally nonactive.

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

The results support the principles above. Table 6 summarizes the token distribution in terms of the RD for predicate, i.e. open propositions.[32] A total of 87% of *wa* tokens occur with predicate of NPM, while the preference for predicate of large RD is not the case for *ga*. A total of 10% of *ga* appear with predicate of RD 1, 15% with RD 4 through 19, 26% with RD 20 and over, and only 43% of *ga* tokens occur with predicate of NPM. The results here suggest that the open proposition of *ga* utterances may be either active or semi-active or nonactive. On the other hand, the open proposition of *wa* utterances is in general not active.

Table 6: Token distribution of *wa*/*ga*-marked subjects in terms of RD of open proposition (predicate)

RD of open proposition	*wa*		*ga*	
1	5	(2%)	28	(10%)
2 - 3	6	(3%)	20	(7%)
4 - 19	6	(3%)	43	(15%)
20+	13	(6%)	74	(26%)
NPM	198	(87%)	123	(43%)
Total	228	(100%)	288	(100%)

At this point, we need to further examine the token distribution of *wa* and *ga* since Table 6 does not show whether there is correlation between *wa*/*ga*-marked elements and predicates in terms of RD. For this purpose, Tables 7 and 8 illustrate the token distribution of *wa* and *ga* in terms of both *wa*/*ga*-marked elements and the RD of predicate respectively.

As seen in Table 6, Table 7 illustrates that *wa*-marked elements occur with predicates of large RD. In addition to this overall tendency, Table 7 reflects the principle that *wa*-marked elements are more active than their open propositions. A total of 91% of the *wa*-marked elements of RD 1 occur with a predicate of NPM, 88% of the *wa*-marked elements of RD 2 and 3 appear with a predicate of NPM, 85% of the *wa*-marked elements of RD 4 through 19 appear with a predicate of NPM, and 81% of the *wa*-marked elements of RD 20+ occur with a predicate of RD NPM. Although it is also the case that a total 83% of the *wa*-marked elements of RD NPM, i.e. 34 tokens, occur with a predicate of RD NPM, i.e. *wa*-marked elements and predicates are equally nonactive, the 34 cases are only 17% of all the predicates of RD NPM (34 out of 198); namely, a total of 83% of the NPM predicates (164 out of 198) appear with *wa*-marked elements of smaller RD. Overall, 174 *wa* tokens, which are 76% of the total of 228 *wa* tokens, conform to the expected pattern, which is indicated by the boxed cells in boldface in Table 7.

Table 7: Token distribution of *wa* in terms of both RD of *wa*-marked elements and RD of predicate. % in terms of the total for RD of *wa*-marked elements.

RD of *wa*-marked elements	Pred. RD 1		Pred. RD 2-3		Pred. RD 4-19		Pred. RD 20+		Pred. RD NPM		Total	
1	2	(3%)	2	(3%)	1	(1%)	2	(3%)	69	(91%)	76	(100%)
2 - 3	0		2	(5%)	0		3	(7%)	38	(88%)	43	(100%)
4 -19	1	(2%)	1	(2%)	4	(8%)	2	(4%)	44	(85%)	52	(100%)
20+	0		1	(6%)	1	(6%)	1	(6%)	13	(81%)	16	(100%)
NPM	2	(5%)	0		0		5	(12%)	34	(83%)	41	(100%)
Total	5	(2%)	6	(3%)	6	(3%)	13	(6%)	198	(87%)	228	(100%)

Table 8 shows three general patterns for the token distribution of *ga*. First, a total of 81% of the *ga*-marked elements of RD NPM occur with the equally nonactive predicate, i.e. the predicate of RD NPM. Second, there is also a pattern that the predicates of *ga* are more active than the *ga*-marked elements. A total of 82% of the predicates of RD 20+ appear with *ga*-marked elements of RD NPM. About a half of the predicates of RD 4 through 19 appear with *ga*-marked elements of RD NPM, and 55% of the predicates of RD 2 and 3 appear with *ga*-marked elements of RD NPM. Finally, *ga*-marked elements

occur with equally active predicates. Of the total of 28 *ga*-marked elements of RD 1, a total of 54% of the predicates of RD 1 occur with *ga*-marked elements of RD 1, i.e. equally active predicates. In total, there are 251 *ga* tokens, 87% of the total of 288 tokens, which conform to the three patterns above, which is illustrated by the boxed cells in bold face in Table 8.

Table 8: Token distribution of *ga* in terms of both RD of *ga*-marked elements and RD of predicate. % in terms of the total for predicate RD.

RD of *ga*-marked elements	Pred. RD 1		Pred. RD 2-3		Pred. RD 4-19		Pred. RD 20+		Pred. RD NPM		Total	
1	15	(54%)	2	(10%)	2	(5%)	1	(1%)	8	(7%)	28	(10%)
2 - 3	2	(7%)	6	(30%)	4	(9%)	3	(4%)	3	(2%)	18	(6%)
4 -19	2	(7%)	1	(5%)	10	(23%)	2	(3%)	8	(7%)	23	(8%)
20+	0		0		7	(16%)	7	(10%)	4	(3%)	18	(6%)
NPM	9	(32%)	11	(55%)	20	(47%)	61	(82%)	100	(81%)	201	(70%)
Total	28	(100%)	20	(100%)	43	(100%)	74	(100%)	123	(100%)	288	(100%)

The activation pattern that the *ga*-marked elements of small RD appear with a predicate of small RD suggests a clue to a fundamental functional difference between *wa* and *ga* Namely, both *wa* and *ga* mark active elements; however, they are differentiated by the pattern that *wa* appears with nonactive or less active predicates and *ga* appears with equally active predicates if it marks active elements. Tables 9 and 10 illustrate the point here more clearly by showing the number of tokens for each of the activation patterns: (1) *wa*/*ga*-marked elements are more active than the predicates, (2) *wa*/*ga*-marked elements are less active than the predicates, (3) *wa*/*ga*-marked elements and the predicates are both RD of NPM, i.e. equally nonactive, and (4) *wa*/*ga*-marked elements and the predicates are both previously mentioned [PM] and equal in RD, i.e. equally highly active or active.

Overall, Table 9 shows the predicted patterns. The most frequent type for *wa* is the one in which *wa*-marked elements are more active then the predicates, which comprises a total 77% of the *wa* tokens. The two most frequent types of *ga* are those in which *ga*-marked elements are more active than the predicates or *ga*-marked elements and the predicates are equally nonactive, which is predicated by the activation patterns of *wa*

discussed earlier. In terms of the third activation pattern expected for *ga*, the total number appears to contradict the expected patterns since there are more tokens for the pattern that *ga*-marked elements are more active then the predicates than for the pattern that *ga*-marked elements and the predicates are equally active, i.e. previously mentioned. However, the predicated pattern for *ga* is in fact reflected in the data and this will become clear if the *ga* tokens are compared with *wa*-marked subject tokens, as in Table 10.

Table 9: Total number of tokens for *wa*/*ga*-marked elements and predicates in terms of the four activation types

Degrees of activation of *wa*/*ga*-elements and predicates	*wa*		*ga*	
RD of *wa*/*ga*-element < RD of predicate	176	(77%)	41	(14%)
RD of *wa*/*ga*-element > RD of predicate	14	(6%)	118	(41%)
RD of *wa*/*ga*-element = RD of predicate = NPM	34	(15%)	100	(35%)
RD of *wa*/*ga*-element = RD of predicate = PM	4	(2%)	29	(10%)
Total	228	(100%)	288	(100%)

Table 10 includes the tokens of only the *wa*-marked subjects and shows the percentage in terms of the total number of tokens for each activation pattern; therefore, the percentages show the choice that speakers make for whether to mark a subject with *wa* or *ga*. For the cases in which *wa*/*ga*-marked subjects are more active than the predicates, 70% of the 137 cases appear with *wa*, which is the pattern predicted by the principle stated in (11). For the other three patterns, on the other hand, *ga* is predominantly preferred to *wa.* This trend is shown especially clear with the cases in which *wa*/*ga*-marked subjects and the predicates are previously mentioned and equal in RD, i.e. equally more or less active; all the tokens for this pattern appear with *ga*.

Table 10: Total number of tokens for *wa*/*ga*-marked subjects and predicates in terms of the four activation types

Degrees of activation of *wa*/*ga*-subject and predicates	*wa*		*ga*		Total	
RD of *wa*/*ga*-element < RD of predicate	96	(70%)	41	(30%)	137	(100%)
RD of *wa*/*ga*-element > RD of predicate	8	(6%)	118	(94%)	126	(100%)
RD of *wa*/*ga*-element = RD of predicate = NPM	21	(17%)	100	(83%)	121	(100%)
RD of *wa*/*ga*-element = RD of predicate = PM	0	(0%)	29	(100%)	29	(100%)

3.3.4. Discussion

The results from the RD measurement in the preceding sections can be summarized by the following.

(i) *Wa* marks elements of low RD, therefore, relatively active entities.

(ii) *Ga* marks elements of high RD, therefore, relatively nonactive entities.

(iii) *Wa*-marked elements fill open predicates which have high RD, therefore, of relatively low activation.

(iv) *Ga*-marked elements fill open predicates which have either low RD or high RD, therefore, of either low activation or high activation.

The overall tendencies above support the principles stated in (14), which is repeated as (15) below. (15a) is exemplified by the cases in which *wa* marks an entity of low RD and is predicated by an open proposition of high RD. (15b) is exemplified by the cases that the RD of predicates is smaller than that of *ga*-marked elements. (15c) is exemplified by the cases in which *ga* marks an entity of low RD and is predicated by an open proposition of low RD. (15d) is exemplified by the cases in which *ga* marks an entity of high RD and is predicated by an open proposition of high RD.

(15)

a. A $>_{wa}$ B — A is more active than B.

b. A $<_{ga}$ B — B is more active than A.

c. A $=_{ga}$ B = active — A and B are equally active.

d. A $=_{ga}$ B = nonactive — A and B are equally nonactive.

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

In Section 2.4.2, I discussed the fact that the activation pattern in (15c) is found in sentences with a WH phrase where *wa* and *ga* are interchangeably used; however, (15c) is typically exemplified by sentences without a WH phrase in the conversational database. The claim that A and B are equally active in (15c) appears to go against the Gricean maxim of quantity "Make your contribution as informative as is required" (Grice 1975: 45), since utterances of this type consist of only active elements at the point of discourse. However, the utterances of this type in the database commonly make a new proposition by combining active propositions. The following example from the database illustrates this point.

(16)

1. S-san no osibai taitee ichizikan zengo nan desu yone
   GEN show usually 1-hour about NOM COP FP
   'Mr. S's show is usually about one-hour long.'

2. doosite ka to omou to
   why Q QT think if
   'If (I) think why'

3. yappari yakusha no nikutai ga sore sika zizokudekinai to omou n desu
   as-expected actor GEN body that only last:POT:NEG QT think NOM COP
   '(I) think that the actors can continue (the show) for only that long.'

4. uun kitsui kara ne
   um tough because FP
   'Because (it's) hard.'

5. sorede yappari S-san ga ichizikan kurai de matometerassharu n dakedomo
   so as-expected 1-hour about in make:HON NOM but
   'So, Mr. S makes (his show) about one-hour long (as you know), but...'

Ga in (16.5) marks the active entity *Suzuki-san*, which was mentioned four clauses before, and it is also predicated by an active proposition *X makes the show about one-hour long*. However, the proposition conveyed by the whole utterance (16.5) is not active in its preceding context, i.e. *Mr. Suzuki makes the show about one hour long*. Putting it in a

different way, (15b) consists of two elements: the active entity or proposition A and the active *open* proposition B, i.e. 'X *ga* predicate' [X: variable], and the *closed* proposition which combines A and B with *ga*, i.e. 'A *ga* predicate', is asserted.

3.4. *WA* AND *GA* IN EMBEDDED CLAUSES: MEASUREMENT RESULTS

As noted earlier, it has long been pointed out that there is a constraint as to the use of *wa* in embedded clauses. However, in the previous studies the analysis of *wa* in embedded clauses has been mostly limited to the mini-discourse level, i.e. sentences in isolation or question-answer pairs, and the use of *wa* in embedded clauses in actual discourse has not been investigated. The purpose of this section is to probe the database to examine the use of *wa* in embedded clauses; more specifically, this section will pose the following questions for the analysis.

(i) To what extent is the use of *wa* in embedded clauses limited in terms of the token distribution?

(ii) Does the same principle found in the previous sections apply to *wa* in embedded clauses? i.e. Does *wa* in embedded clauses mark active elements which are predicated by less active or nonactive open proposition?

(iii) Is there any difference between subject-marking *wa* and non-subject marking *wa* in embedded clauses in terms of degrees of activation?

(iv) Is there any correlation between the token distribution patterns and types of embedded clauses.

The clauses treated as embedded in the database for the purpose of this section consist of the following types: clauses followed by the conjunctive morphemes *kara* 'because', *node* 'because', *ba* 'if', *tara* 'if', *nara* 'if', *to* 'when/if', *toki* 'when', *temo* 'even though', *kara* 'after', *nagara* 'while', *aida* 'while', *made* 'until', complement clauses accompanied by the nominalizer *no* or *koto*, and clauses which modify a head noun, including both relative clause constructions and noun complement constructions.

The discussion for this section will proceed as follows. Section 3.4.1 will discuss the token distribution of *wa* to compare *wa*-marked subjects and *wa*-marked non-subjects, and Section 3.4.2 will compare *wa*-marked subjects and *ga*-marked subjects. We will extend the analysis to predicates of *wa*/*ga*-embedded clauses in Section 3.4.3. Section 3.4.4 will summarize the results, and Section 3.4.5 will discusses the use of *wa*/*ga* in connection with different types of embedded clauses.

3.4.1. RD for *Wa*-marked Subjects and *Wa*-marked Non-subjects

In Section 3.3.1, we have found that *wa*-marked subjects and *wa*-marked non-subjects in matrix clauses exhibit the same pattern in terms of the degrees of activation; *wa*-marked elements are active. This section illustrates the measurement results for *wa*-marked elements in embedded clauses to see whether the same expected pattern is found here as well. The database included 78 *wa*-marked elements in embedded clauses, and *wa*-marked pronouns and temporal expressions were excluded for the RD measurement for the same reason discussed in Section 3.2. For the RD measurement here, there are 44 *wa*-marked subjects and 34 *wa*-marked non-subjects. The token distribution is shown in Table 11 for each RD, and the distribution is summarized in Table 12 in terms of five RD groups: 1 though 4, 5 through 10, 11 through 19, 20 and over, and NPM.

Table 11: Token distribution of *wa*-marked subjects and non-subjects in embedded clauses in terms of RD

RD	*wa*:	subject	*wa*:	non-subject
1	10	(30%)	2	(11%)
2	4	(12%)	3	(17%)
3	1	(3%)	1	(6%)
4	1	(3%)	0	
5	2	(6%)	2	(11%)
6	1	(3%)	1	(6%)
7	0		1	(6%)
8	0		0	
9	1	(3%)	0	
10	0		0	
11	1	(3%)	1	(6%)

12	1	(3%)	0	
13	0		1	(6%)
14	0		0	
15	1	(3%)	1	(6%)
16	0		0	
17	0		0	
18	1	(3%)	0	
19	0		0	
20+	5	(15%)	2	(11%)
NPM	4	(12%)	3	(17%)
Total	33	(100%)	18	(100%)

Table 12: Token distribution of *wa*-marked subjects and non-subjects in embedded clauses and % by total for subjects and non-subjects

RD	*wa*:	subject	*wa*:	non-subject
1 - 4	16	(49%)	6	(33%)
5 - 10	4	(12%)	4	(22%)
11 - 19	4	(12%)	3	(17%)
20+	5	(15%)	2	(11%)
NPM	4	(12%)	3	(17%)
Total	33	(100%)	18	(100%)

The results in Table 12 shows the expected overall trend though it is exhibited by *wa*-marked subjects more clearly than by *wa*-marked non-subjects. For both *wa*-marked subjects and non-subjects, RD 1 through 4 shows the highest percentage, i.e. 49% (16 out of 33 tokens) and 33% (6 out of 18 tokens) respectively. For larger RD groups, the tokens are divided almost equally; there is no particular pattern found for RD 5 through NPM. The token distribution here is analogous to that found for *wa* in matrix clauses also in that *wa* occurs with elements of large RD, i.e. 5 tokens for 20+ and 4 tokens for NPM for subjects and 2 tokens for 20+ and 3 tokens for NPM for non-subjects. In Section 3.5 below, those cases of *wa* with elements of high RD in the database, along with those in matrix clauses, will be discussed in terms of both the notion of activation and other factors.

3.4.2. RD for *Wa*/*ga*-marked Subjects

This section gives the results for subjects marked by *wa* and *ga* in embedded clauses in order to compare the token distribution of the two in terms of RD. Table 13 illustrates the token distribution and the percentage of *wa* and *ga* in terms of the total for each RD, and Table 14 summarizes the distribution in terms of the four RD groups: 1 through 4, 5 through 19, 20+, and NPM, and the percentage in terms of the total for each RD group. As indicated in Table 12, there is a clear pattern that *ga* is the preferred choice for RD NPM; a total of 94% of the elements of RD NPM occur with *ga*, while only 6% occur with *wa*. This tendency that *ga* is preferred for large RD conforms to the principle that *ga* marks nonactive elements. For small RD, on the other hand, the results do not exhibit the expected trend; *ga* appears with elements of both RD groups 1 through 4 and 5 through 19 more frequently than *wa*, which indicates that there is preference for *ga* over *wa* overall in embedded clauses. This point will be further investigated in Section 3.4.5 below in order to see whether this overall preference for *ga* is the case in any type of embedded clauses and whether the measurement result that *ga* is preferred than *wa* in terms of frequency for the small RD range, as shown in Table 14, is due to the preference for *ga* over *wa* in any particular embedded clause type.

In spite of the result above that *ga* is the preferred choice over *wa* overall in embedded clauses, it should be emphasized that the token distribution of *wa* still conforms to the principle that *wa* marks active elements, which is indicated by the pattern that the smaller the RD is, the more *wa* tokens there are. Table 13 illustrates this point. A total of 49% of the *wa*-marked subjects in embedded clauses appear with elements of RD 1 through 4, 24% with elements of RD 5 through 19, 15% with RD 20+, and 12% with RD NPM.

Table 13: Token distribution of *wa*/*ga*-marked subjects in embedded clauses in terms of RD and % by total for each RD

RD	*wa*		*ga*		Total	
1	10	(36%)	18	(64%)	28	(100%)

2	4	(50%)	4	(50%)	8	(100%)
3	1	(11%)	8	(89%)	9	(100%)
4	1	(20%)	4	(80%)	5	(100%)
5	2	(40%)	3	(60%)	5	(100%)
6	1	(25%)	3	(75%)	4	(100%)
7	0		1	(100%)	1	(100%)
8	0		3	(100%)	3	(100%)
9	1	(100%)	0		1	(100%)
10	0		1	(100%)	1	(100%)
11	1	(25%)	3	(75%)	4	(100%)
12	1	(50%)	1	(50%)	2	(100%)
13	0		1	(100%)	1	(100%)
14	0		1	(100%)	1	(100%)
15	1	(33%)	2	(67%)	3	(100%)
16	0		0		0	
17	0		1	(100%)	1	(100%)
18	1	(100%)	0		1	(100%)
19	0		0		0	
20+	5	(50%)	5	(50%)	10	(100%)
NPM	4	(6%)	60	(94%)	64	(100%)
Total	33	(22%)	119	(78%)	152	(100%)

Table 14: Token distribution of *wa*/*ga*-marked subjects in embedded clauses in terms of RD: 1-4, 5-19, 20+, and NPM, and % by total for each RD group.

RD	*wa*		*ga*		Total	
1 - 4	16	(32%)	34	(68%)	50	(100%)
5 - 19	8	(29%)	20	(71%)	28	(100%)
20+	5	(50%)	5	(50%)	10	(100%)
NPM	4	(6%)	60	(94%)	64	(100%)
Total	33	(22%)	119	(78%)	152	(100%)

Table 15: Token distribution of *wa*/*ga*-marked subjects in embedded clauses in terms of RD: 1-4, 5-19, 20+, and NPM, and % by total for *wa*/*ga*.

RD	*wa*		*ga*	
1 - 4	16	(49%)	34	(29%)

5 - 19	8	(24%)	20	(17%)
20+	5	(15%)	5	(4%)
NPM	4	(12%)	60	(50%)
Total	33	(100%)	119	(100%)

3.4.3. RD for Predicates

The RD measurement for open propositions in this section was conducted in the same manner as in the measurement for open propositions in matrix clauses in Section 3.3.3 above. Table 16 summarizes the token distribution in terms of the RD for predicate. The token distribution patterns here is analogous to those in matrix clauses as discussed in Section 3.3.3. Namely, there are patterns in Table 16 that elements marked by *wa* are nonactive, i.e. large in RD, while elements marked by *ga* are either active or nonactive, i.e. either small or large in RD. A total of 85% of *wa* tokens occur with predicate of NPM, 13% with predicate of RD 20+, and only 2% with predicate of RD 1 through 3. A total of 61% of *ga* tokens occur with predicates of NPM; however, there are also 15% of the cases which appear with predicate of RD 1 through 3, which is about the same as the percentage for RD 20+.

Table 16: Token distribution for embedded clauses in terms of RD of open proposition (predicate)

RD of open proposition	*wa*		*ga*	
1 - 3	1	(2%)	18	(15%)
4 - 19	0		8	(7%)
20+	6	(13%)	21	(18%)
NPM	41	(85%)	72	(61%)
Total	51	(100%)	119	(100%)

Given the overall patterns described above, we will further examine the data in terms of both RD of *wa*/*ga*-marked elements and predicate RD. Table 17 illustrates the expected pattern that the RD of predicate is larger than the RD of *wa*-marked elements. For the *wa*-marked elements of RD 1 through 3, 95% of the cases occur with predicates of RD NPM, and for the *wa*-marked elements of RD 4 through 19, all the tokens occur

with predicates of RD NPM. In total, there are 36 cases in which *wa*-marked elements are more active than the predicate (the boxed cells in bold face in Table 17), which comes to 75% of the total of 48 cases. As noted earlier, Section 3.5 below will discuss the cases in which *wa* marks elements of 20+ and NPM, which were not captured by the RD measurement.

Table 17: Token distribution of *wa* in embedded clauses in terms of both RD of *wa*-marked elements and RD of predicate, and % in terms of total for *wa*-marked elements

RD of *wa*-marked el.	Pred. RD 1-3		Pred. RD 4-19	Pred. RD 20+/NPM		Total	
1 - 3	1	(5%)	0	20	(95%)	21	(100%)
4 -19	0		0	16	(100%)	16	(100%)
20+/NPM	0		0	11	(100%)	11	(100%)
Total	1	(2%)	0	47	(98%)	48	(100%)

Table 18 shows that the RD measurement results indicate the expected patterns for *ga*. First, *ga*-marked elements and predicates are equally active or nonactive. Half of the tokens of predicate RD 1 through 3 appear with the *ga*-marked elements of the same RD, and 60% of the tokens of predicate RD NPM appear with the *ga*-marked elements of NPM. Second, *ga*-marked elements are less active than the predicates. 44% of the tokens of predicate RD 1 through 3 appear with *ga*-marked elements of RD 20+ or NPM. In total, a total of 66% (78 out of the 119 tokens) of the total fall on the expected pattern, which is indicated by the boxed cells in bold face in Table 18.

However, we should note that there is also an opposite trend to the expected pattern described above, which is suggested by the relatively low percentage of the total, i.e. 66%, for the expected patterns. To put it differently, the token distribution also shows a *wa*-like pattern among the *ga* tokens. This point will become clear in Table 19 with the percentage in terms of the total for RD of *ga*-marked elements. Of the total of 30 tokens for the *ga*-marked elements of RD 1 through 3, 57% appear with the predicates of RD NPM, and of the total of 24 tokens for the *ga*-marked elements of RD 4 through 19, 83% occur with the predicates of RD NPM.

The RD measurement for predicates in embedded clauses indicates that *wa* in embedded clauses is in general the same as *wa* in matrix clauses in terms of the token distribution, while there is a difference in *ga* between matrix clauses and embedded clauses in that *ga* in embedded clauses exhibits a *wa*-like pattern to some extent in addition to its canonical patterns. We will further examine *ga* in embedded clauses in terms of types of embedded clauses in order to see whether there is any correlation between the distribution patterns of *ga* and types of embedded clause.

Table 18: Token distribution of *ga* in embedded clauses in terms of both RD of *ga*-marked elements and RD of predicate, and % in terms of total for predicate RD

RD of *ga*-marked el.	Pred. RD 1-3		Pred. RD 4-19		Pred. RD 20+/NPM		Total	
1 - 3	9	(50%)	4	(50%)	17	(18%)	30	(25%)
4 -19	1	(6%)	3	(38%)	20	(22%)	24	(20%)
20+/NPM	8	(44%)	1	(13%)	56	(60%)	65	(55%)
Total	18	(100%)	8	(100%)	93	(100%)	119	(100%)

Table 19: Token distribution of *ga* in embedded clauses in terms of both RD of *ga*-marked elements and RD of predicate, and % in terms of total for *ga*-marked elements

RD of *ga*-marked el.	Pred. RD 1-3		Pred. RD 4-19		Pred. RD 20+/NPM		Total	
1 - 3	9	(30%)	4	(13%)	17	(57%)	30	(100%)
4 -19	1	(4%)	3	(13%)	20	(83%)	24	(100%)
20+/NPM	8	(12%)	1	(2%)	56	(86%)	65	(100%)
Total	18	(15%)	8	(7%)	93	(78%)	119	(100%)

3.4.4. Discussion

Despite the claim which has long been made in the literature that the use of *wa* is restricted in embedded clause, the conversation data for this study includes a fair number of *wa* tokens, i.e. 78 *wa* tokens out of the *wa*/*ga* total of 219, including *wa*/*ga*-marked pronouns and temporal expressions, which indicates that *wa* does appear in embedded clauses rather frequently. The more careful examination of *wa* and *ga* in embedded clauses by the RD measurement has showed that overall they conform to the basic principles

suggested in the analysis of *wa* and *ga* in matrix clauses. The results have also indicated that the principles apply to both *wa*-marked subjects and non-subjects in embedded clauses, as discussed in Section 3.4.1. I repeat the principles below as (17).

(17)

a.	A > B (*wa*)	A is more active than B.
b.	A < B (*ga*)	B is more active than A.
c.	A = B = active (*ga*)	A and B are equally active.
d.	A = B = nonactive (*ga*)	A and B are equally nonactive.

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

However, the RD measurement results have also shown that there is a difference between matrix clauses and embedded clauses. Although the use of *wa* in embedded clauses is in general identical to that in matrix clauses in terms of the principle above, embedded clauses differ from matrix clauses in the following two respects. First, there is an overall pattern that *ga* is the preferred choice over *wa* in embedded clauses. The results have indicated this overall trend in terms of the total number of tokens; *ga* is dominant over *wa* even in the small RD range where *wa* is dominant in matrix clauses. Second, there is a pattern in embedded clauses that *ga* may mark active elements and predicated by nonactive open proposition; namely, *ga* may occur in the environment in which *wa* occurs. The findings above may suggest that there is a trend that *ga* is used in the place of *wa*, at least in some cases, in embedded clauses, which has long been suggested in the previous studies. Among the claims as to the preference for *ga* over *wa* in embedded clauses, the strongest form is found in Kuno (1973: 56), where he states "the distinction between the thematic *wa* and the descriptive *ga* and the exhaustive-listing *ga* becomes neutralized

in subordinate clauses. All three are realized as *ga*, ... However, the contrastive *wa* can appear in subordinate clauses." Although Kuno claims that *wa* can appear in embedded clauses only when it has a contrastive sense, it is not clear at all in our database whether the *wa*-marked entities in embedded clauses exhibit a clear sense of contrast. Although the database contains cases in which *wa* seems to show contrast, as in (18.1), there are also many cases in which such clear sense of contrast is not present, as in (19.1).

(18)

1. [hokano osibai de wa yappa sooyuu koto wa hukanoona node]
   other play in as-expected that-kind thing impossible because

2. S-san yatteru aida dake demo sooyuu huu ni yatte miyoo to omotte
   Mr. S doing duration only at-least that-kind way in do try:HOR QT think
   'Because as expected that kind of thing is impossible to do in other plays, I think that I'd try doing that at least for the duration of Mr. S's play.'

(19)

1. de [otoosama wa saigonohoo no mono o ossharenaku natte kara]
   and father at-the-end GEN thing OBJ say:HON:NEG become after

2. anata wa otoosama ga osshatteru yoona koto ga owakarini natta
   you father saying:HON like thing understand:HON become

   kanji ga nasatta n desu tte
   feeling do:HON NOM COP QT
   '(I heard) after your father finally became unable to speak, you felt you understood what your father was saying.'

In the context preceding (18), the speaker has been talking about plays which she has participated in before, and she mentions that the play directed by Mr. S is exceptionally tough in that unlike other plays, Mr. S's play requires much more concentration and energy. In (18.1), the *wa*-marked elements, both *hokano osibai* 'other plays' and *sooyuu koto* 'that kind of thing' (which refers to the nature of Mr. S's play), exhibit clear sense of contrast, given the immediately preceding context.

In the context preceding (19), on the other hand, the speaker has been talking about the hearer's father who finally died in a hospital and the immediately preceding

context does not contain any competing referent for *otoosama* 'father' which is marked by *wa* in (19.1); therefore, it is not clear at all whether there is a sense of contrast on *wa*-marked element in (19.1).

3.4.5. Use of *Wa* and *Ga* and Embedded Clause Types

Given the general picture of *wa* and *ga* in embedded clauses, we need to further investigate the occurrence of *wa* and *ga* in embedded clauses in the database in order to examine whether there is a correlation between the token distribution and the types of embedded clauses. As discussed above, Sections 3.4.2 and 3.4.3 have revealed the overall tendency in embedded clauses that *ga* appears more frequently than *wa* in the environment in which *wa* is dominant over *ga* in matrix clauses. In order to see whether this overall tendency is due to particular embedded clause types, the token distribution of *wa* and *ga* will be presented for each embedded clause type below.

Table 20 shows the number of tokens for each embedded clause type in the database. As shown clearly in the table, there are two groups of embedded clause types: the ones in which both *wa* and *ga* occur frequently, i.e. (a) through (c), and the ones in which *ga* occurs but *wa* does not occur or occurs only once, i.e. (d) through (i).[33]

Table 20: # of tokens of *wa* and *ga* for each embedded clause type

Type of embedded clause	*wa* (non-subject)	*wa* (subject)	*ga*
a. clausal complement	10	24	51
b. kara/node 'because'	17	17	31
c. N-modifying clause	6	0	27
d. ba/to/tara/nara 'if/when'	0	2	16
e. toki 'when'	0	1	10
f. temo 'even though'	0	0	3
g. (te)kara 'after'	0	0	1
h. nagara/aida 'while'	1	0	1
i. made 'until'	0	0	1
Total	34	44	141

Given the overall token distribution patterns in Table 20, we may hypothesize that there is a tendency that *ga* is used in the place of *wa* at least in the embedded clause types (d) through (i) in Table 20, since the frequency of *wa* for those types is none or close to none, despite the relatively high frequency of *ga*. In order to investigate this point, I will examine the token distribution of *ga* in terms of the environment in which *ga* occurs.

As discussed throughout this chapter, *wa* and *ga* exhibit different activation patterns; *wa* is characterized by the pattern that *wa*-marked elements are more active than the open propositions which predicate the *wa*-marked elements, and *ga* exhibits the pattern that *ga*-marked elements are more active than the open propositions, or *ga*-marked elements and the open propositions are equally active or nonactive. Therefore, *wa* and *ga* are differentiated as in Table 21.[34]

Table 21: *Wa*/*ga* and their RD domains. White = *wa*, shaded = *ga*.

RD of *wa*/*ga*- marked elements	Predicate RD 1-3	Predicate RD 4-19	Predicate RD 20+	Predicate RD NPM
1 - 3				
4 -19				
20+				
NPM				

Table 21 shows two domains in terms of the RD of *wa*/*ga*-marked elements and the RD of predicate. Since *wa*-marked elements are more active than the predicate, tokens of *wa* fall on the unshaded area in the table. On the other hand, *ga*-marked elements are less active than the predicate, or *ga*-marked elements and predicates are equal in degrees of activation; therefore, *ga* falls on the shaded area in the table. For expository purposes, I will henceforth call the unshaded area in the table the *wa domain* and the shaded area *ga domain*.

In order to examine whether there are cases in which *ga* is being used in the place of *wa* in embedded clauses and which embedded clause type exhibits such pattern, I will present the token distribution of *ga* in terms of the two domains in Table 21. If *ga* occurs in the *ga* domain, it is the canonical use of *ga*; however, if *ga* occurs in the *wa* domain, i.e.

in the environment in which *wa* should occur, we may assume that *ga* is being used instead of *wa*. Table 22 illustrates the token distribution in terms of the *wa*/*ga* domains.

The token distribution in Table 22 appears to reflect a correlation between the frequency of *ga* in the *wa* domain and types of embedded clauses. In terms of the percentage, *ga* appears in the *wa* domain most frequently in noun-modifying clauses, i.e. 50% of the cases, while it appears least frequently in clausal complements, i.e. 32% of the cases, and 'because' clauses and 'if' clauses fall between these two extremes. Since there are only a small number of tokens for *toki* 'when', *temo* 'even though', *kara* 'after', *nagara*/*aida* 'while', and *made* 'until', we need more data for further discussion of those types.

Table 22: Token distribution of *ga* in terms of *wa*/*ga* domains

type of embedded clause	# of *ga* tokens in *wa* domain		# of *ga* tokens in *ga* domain		Total	
clausal complement	15	(32%)	32	(68%)	47	(100%)
kara/node 'because'	10	(35%)	19	(66%)	29	(100%)
ba/to/tara/nara 'if'	5	(39%)	8	(62%)	13	(100%)
N-modifying clause	10	(50%)	10	(50%)	20	(100%)
toki 'when'	1		3		4	
temo 'even though'	1		2		3	
kara 'after'	1		0		1	
nagara/aida 'while'	0		1		1	
made 'until'	0		1		1	
Total	43		75		118	

To further illustrate the overall pattern, Table 23 compares the number of *ga* tokens in the *wa* domain and the total number of subject-marking *wa* for each embedded clause type. For 'because' clauses and clausal complements, about 35% of the cases are marked by *ga* in the *wa* domain, which indicates that about 35% of the *ga* tokens for those clause types are used in the place of *wa*. For noun-modifying clauses and 'if' clauses, on the other hand, the frequency of *ga* in the *wa* domain is high; in the former 77% of the cases are marked by *ga* in the *wa* domain and in the latter all the five cases are marked by *ga* in the *wa* domain. Since there are only two tokens or less for *toki* 'when', *temo* 'even

though', *kara* 'after', *nagara*/*aida* 'while', *made* 'until', further discussion for those types requires more data.

Table 23: Token distribution of *ga* in the *wa* domain and *wa.*

type of embedded clause	# of *ga* tokens in *wa* domain		Total # of *wa* subject		Total	
kara/node 'because'	10	(33%)	20	(67%)	30	(100%)
clausal complement	15	(37%)	26	(63%)	41	(100%)
N-modifying clause	10	(77%)	3	(23%)	13	(100%)
ba/to/tara/nara 'if'	5	(100%)	0	(0%)	5	(100%)
toki 'when'	1		1		2	
temo 'even though'	1		0		1	
kara 'after'	1		0		1	
nagara/aida 'while'	0		1		1	
made 'until'	0		0		0	
Total	43		51		94	

Figure 1 summarizes the correlation between the frequency of *ga* in the *wa* domain and types of embedded clauses, which I discussed above. Of the two types of adverbial subordinate clause, the frequency of *ga* in the *wa* domain is higher in the 'if' clauses than the 'because' clauses, which indicates that the degree of *ga* substitution for *wa* is higher in the former than in the latter. Similarly, the frequency of *ga* in the *wa* domain is higher in the noun-modifying clauses than in the clausal complement clauses, which shows that *ga* is substituted for *wa* more frequently in the former than in the latter.

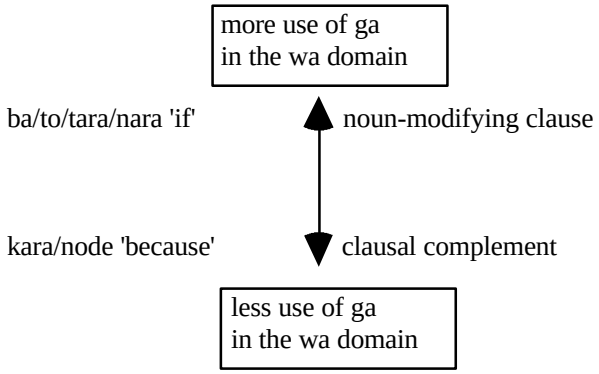


Figure 1: The use of *ga* in the *wa* domain and types of embedded clause

Given the RD measurement of *wa* and *ga* in embedded clauses, this section has demonstrated that *wa* and *ga* in embedded clauses are in general identical with those in

matrix clauses in terms of the principles proposed in this study; namely, *wa*-marked elements are more active than the predicate open proposition, and *ga*-marked elements are less active than the predicates or *ga*-marked elements and the predicates are equal in degrees of activation. However, embedded clauses are distinct from matrix clauses in that *ga* may be used in the environment in which *wa* appears, which in principle supports the long-noted claim in the literature that *ga* is substituted for *wa* in embedded clauses. However, this study could go beyond the claim by demonstrating that the degree of *ga* use in the place of *wa* in embedded clause varies depending on the embedded clause type in which they appear. The correlation between the use of *wa*/*ga* and the clause types found here will be further discussed in Chapter 5 in the framework of Role and Reference Grammar [RRG] (cf. Van Valin 1993) in connection with the global picture of focus/non-focus manifestation in Japanese.

3.5. *WA*-MARKED ELEMENTS OF RD 20+ AND NPM

Finally, this section examines *wa* which marks elements of large RD, i.e. 20+ and NPM, since those instances of *wa* do not conform to the principle that *wa* marks active elements and the RD measurement alone does not account for those cases of *wa*. Including *wa* in both matrix and embedded clauses, there are 75 *wa* tokens in total which mark elements of RD 20+ or NPM. The 75 cases of *wa* can be grouped into the following categories.

(20)

a. *Wa* marks elements which are active in the immediately preceding discourse context but not linguistically present in the preceding text.

b. *Wa* marks elements which are not active in the immediately preceding context but were activated a number of times in the preceding context; therefore those have been the on-going topic in the discourse, hence, are relatively easy to re-activate.

c. *Wa* marks elements which are made accessible to activation by an active element in the immediately preceding context.

d. *Wa* marks elements which are made accessible to activation by universally present element and knowledge shared by speaker and hearer.

e. None of the above

Although *wa*-marked elements in (20a) through (20d) are not captured by the RD measurement method, they are accounted for by the notion of activation, as will be discussed below. In terms of the token distribution, 88% of the total of 75 *wa*-marked tokens here fall on this category. The 9 tokens in (20e) will be discussed later from a structural point of view, rather than the notion of activation. Table 24 summarizes the token distribution for *wa*-marked elements of RD 20+ or NPM.

Table 24: Token distribution for *wa* of RD 20+ and NPM

Type	# of token	
17a	6	(8%)
17b	11	(15%)
17c	44	(59%)
17d	5	(7%)
17e	9	(12%)
Total	75	(100%)

3.5.1. Activation and Accessible to Activation

Of the six cases for (20a), two cases are found in the very beginning of the TV talk show, where the interviewer introduces the guest of the day to the viewer. In the immediately preceding context of the utterance, the viewers of the talk show know that there is a guest for the day and the host is going to introduce the guest at this point; therefore, the proposition *there is a guest today* is active, even though it is not explicitly mentioned in the preceding context. (21.1) is an example of this kind.

(21)
1. minasama konnichiwa
everyone hello
'Hello, everyone.'

2. kyoo no okyakusama wa eegahyooronka deirasshaimasu
today GEN guest movie-reviewer COP:HON
'Today's guest is a movie reviewer.'

The other cases in (20a) are found in the context in which the speaker refers to some entity or situation present in the situational context. For example, the speaker points at a part of the clothes which the hearer is wearing while mentioning the entity marked by *wa*. Therefore, the entities marked by *wa* in this category are not linguistically present in the preceding context; however, they are contextually active at the point of utterance.

The elements marked by *wa* in (20b) are not active in the immediately preceding context; however, they frequently appear, i.e. they became activated frequently, in the preceding discourse. Given the assumption of activation and decay, as discussed in Section 1.2, it would be reasonable to assume that the more active an element is, the slower it decays in activation. This leads us to a further assumption that if an element is slow in decay, it is relatively easy to reactivate it even without being mentioned in the immediately preceding context. For example, in one of the talk show programs in the database, the conversation is carried on between the host and a movie reviewer and the conversation starts out with the movie reviewer explaining how he teaches movie-related courses at a college and the referent category 'movie' keeps being reactivated throughout the following discourse. Given this previous context in which 'movie' is dominantly reactivated, the host reintroduces it by *wa* often after not mentioning it for more than 30 clauses.

The observation above may suggest the correlation between the frequency of mention in the discourse and likelihood of *wa*-marking for an element[35], which may in turn suggest the correlation between the degree of topicality and the speed of decay in activation. A referent which is highly frequently mentioned throughout the discourse carries high topicality and stays activated, while a referent which is rarely mentioned, e.g. mentioned only once throughout the discourse, carries low topicality and decays rapidly in activation after being mentioned. The former kind of referents may be seen as *global*

topic, while the latter as *non-topic*, and those in the middle of the hierarchy may be seen as *local topic*.

The use of *wa* in (20c) is similar to (20a) in that *wa* in (20c) marks an element which has never been mentioned in the preceding context; however, unlike those in (20a) *wa*-marked referents in this category are activated by a related active referent in the, ordinarily immediately, preceding context which forms a *cohesion* relation with the *wa*-marked referent.[36] Thus, due to a preceding related referent, the *wa*-marked elements in this category are made *accessible* to activation in Dryer's (1994) sense, as discussed in Section 1.2. (22) and (23) are examples of (20c).

(22)

1. tatoeba dantairyokoo dakara sanzyuunin gurai no
   for-example group-tour because 30-people about GEN
   'For example, because it was a group tour of about 30 people'

2. zenbu ano hoteru wa kimatteru si
   all um hotel being-chosen and
   '(they) chose the hotel (for us) and...'

'Hotel' in (22.2) is the first mention in the entire discourse; however, it is marked with *wa* due to the preceding referent 'group tour', which makes a cohesion pair with 'hotel' because of their semantic relation.

(23)

1. ima yatteru no wa ano A no tandai to soikara ano M no
   now doing NOM that GEN junior-college and then that GEN

   zyosidai no hizyookin no koosi desu ne
   women's-college GEN part-time GEN lecturer COP FP
   'What I'm doing now is A Junior College and the lecturer at M Women's University.'

2. naze ka otoko no daigaku wa yonde moraenai
   why Q man GEN college invite receive:NEG
   'I've never been invited to a men's university for some reason.'

'Men's university' in (23.2) has never been mentioned in the preceding context; however, it is marked with *wa* because of the preceding referent 'women's university', which in

obviously in a cohesion relation with 'men's university' due to the *oppositeness* between 'men' and 'women' and the *reiteration* of 'university' (Halliday and Hasan 1976). Table 25 lists some examples of such cohesion pairs, one of which primes the other which the speaker marks with *wa*, despite its first mention in the discourse.

Table 25: Examples of cohesion pairs

preceding referent	*wa*-marked referent
'group tour'	'hotel'
'women's university'	'men's university
'college'	'Dean'
'hotel'	'room'
'show'	'audience'
'(one's) youth'	'junior high school'
'birthplace'	'parents'
'radio'	'listener'

(20c) consists of another type, which is similar to the cases above in that the *wa*-marked element and the preceding element are in a cohesive relation due to their semantics but it is different in that this type manifests cohesion on the proposition level. Typically, the *wa*-marked proposition shares the same or synonymous lexical items with the preceding related proposition and the two propositions are synonymous. (24) is an example of this type.

(24)

1. demo wakai hito to isshoni yareta to yuu koto mo ne
   but young people with together do:POT:PST QT say NOM also FP
   'But the fact that (I) could do with the young people (in the company) also'

2. ano dandan tosi o toru totomoni tosiyori bakkari ga
   uh gradually age OBJ gain while old-people only

   katamarigachi desu yone
   likely-to-group-together COP FP
   'As (we) become older, the old people tend to stick to each other.'

3. dakedo ne wakai hito o genba demotte / kyonen haitta hito
   but FP young people OBJ work at last-year enter:PST people

   toka nee kotosi haitta hito toka sooyuu hito to

and FP this-year enter:PST people and such people with

isshoni sigoto ga dekita to yuu koto wa uresikatta desu nee
together work do:POT:PST QT say NOM pleasant:PST COP FP
'(I) had a pleasant time working with people such as those who entered (the company) last year and this year.'

In (24.3), the proposition *working with people such as those who entered the company last year and this year* makes a cohesion pair with the proposition *I could work with young people* in (24.1) due to the synonymous pairs 'do' and 'work', and 'young people' and 'those who entered the company last year and this year'.

The use of *wa* in (20a), (20b), and (20c) above can be accounted for in terms of activation, whether a referent or proposition is activated or at least made accessible to activation by non-linguistic cues or a related preceding referent or proposition which is linguistically present in the preceding context. In (20d), *wa* marks temporal expressions,[37] and this seems to require the notion *accessible to activation* in a rather broad sense. (20d) is distinct from (20a)-(20c) in that a *wa*-marked temporal expression does not have a preceding related referent. Those temporal expressions include 'last', 'first', 'summer', and 'Monday'.

(25) demo hontoni saigo *wa* awaredatta n desu ga
but really last pitiful:PST NOM COP but
'But it was really pitiful at the end.'

The temporal expressions only require the temporal concept 'now', i.e. the point of utterance, which is universally active when an utterance is made; in temporal representation the speaker's consciousness is focused on the 'now' of the interaction; therefore, 'time' which are linked to 'now' are also activated. Other 'time' which are not directly linked to 'now' are not active until it is mentioned; however, they are still accessible to activation due to the active point of reference 'now'.

3.5.2. Structural Consideration for *Wa*-marked Elements

The last category (20e) in Table 24 is not explained by the notion of activation. All the *wa* cases in this category share the one syntactic characteristic that *wa* marks a predicate. I find two structural types in this category as follows.

(i) *Wa* which accompanies the negative form of copula *de+wa+nai*:

(ii) *Wa* which marks (the *te* form of) a verb

(26.1) is an example of the first type, i.e., *wa* appearing in the negative form of the copula.

(26)

1. sooyuu atasi ga hizyooni koo pipitto wakaru taipu de wa nakatta node
such I very this-way quickly catch type COP NEG:PST because
'Because I was not the type of person who catches things quickly'

2. taihenni kurooitasimasite
really have-a-hard-time:HON
'(I) had a hard time (becoming a good assistant).'

In the second category, *wa* typically marks the *te* form of a verb, as shown in (27.1).

(27)

1. nani ka mono ga nai tte kiite wa ita n desu kedo
what Q thing exist:NEG QT hear be:PST NOM COP but
'(At least) I was hearing that there isn't anything (there), but'

2. hontoni mono ga nakute
really thing exist:NEG
'there was really nothing (there).'

In the database, *wa* which appears in the structural environments above typically marks elements of high RD. In order to illustrate this point, Table 26 summarizes the token distribution of *wa* for each structural type above in terms of RD, in comparison with the total number of *wa* in the whole database. The elements which precede the *wa* in the negative copula are all RD of 20+ or NPM. 60% of the *wa* tokens which mark the *te* form of verbs appear with elements of RD 20+ or NPM. The token distribution of *wa* for the two structural types above suggests that there is correlation between the syntactic

environment in which *wa* occurs and the activation status of the *wa*-marked elements.[38] *Wa*-marked predicates will be discussed in terms of the Role and Reference Grammar in Chapter 5, in order to examine the correlation in terms of focus between *wa*-marked elements and their constituent structure.

Table 26: Token distribution of *wa* for the three structure types in terms of RD of *wa*-marked elements

Type	RD 1-3		RD 4-19		RD 20+/ NPM		Total	
(i)	0		0		6	(100%)	6	(100%)
(ii)	1	(20%)	1	(20%)	3	(60%)	5	(100%)
Wa total	145	(50%)	72	(25%)	75	(26%)	292	(100%)

3.6. CONCLUSION

In this chapter, I have demonstrated that the functional contrast between *wa* and *ga* which I claimed in Chapter 2 is reflected in the conversational Japanese data. I repeat the contrast between *wa* and *ga* in terms of the notion of activation in (28) and in terms of the notion of focus in Figure 2.

(28)

WA: A *wa*-marked referent is more mutually active than the open proposition with which the referent combines.

GA: A *ga*-marked referent is less mutually active than the open proposition with which the referent combines or they are equally mutually active or nonactive.

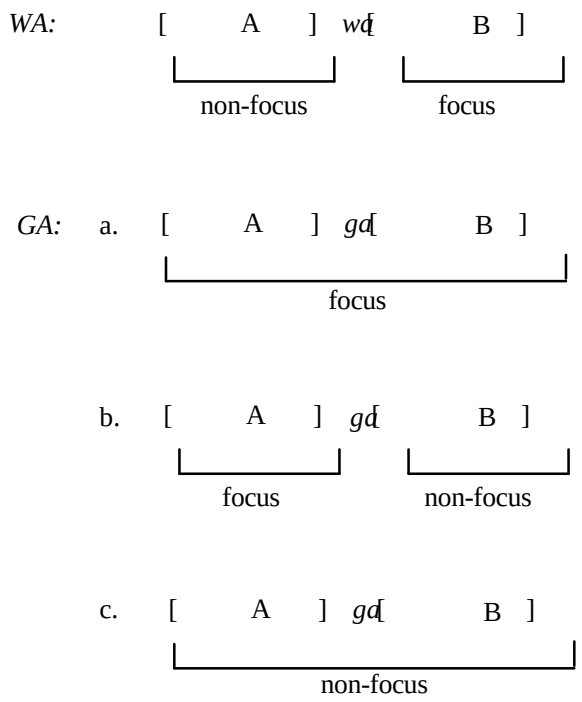


Figure 2: Focus patterns of *wa*/*ga* sentences

The functions of *wa* and *ga* discussed throughout suggest a functional contrast in terms of the informational flow in discourse; *wa* exhibits a special function of marking an element which serves as a basis for a new, i.e. nonactive, proposition, while it is not the case with *ga*; *ga* is used in the *elsewhere* cases.[39] Furthermore, the examination of the conversational Japanese data has not shown the two separate functions of *wa thematic wa* and *contrastive wa* in terms of the activation patterns; in both cases, the entity marked by *wa* is already active and it simply makes a new proposition relevant at the point at which the new proposition is introduced.

In addition to the two types of *ga* sentences which are commonly discussed in the literature, in (a) and (b) in Figure 2, the present study has revealed that conversational Japanese exhibits a third type where the *ga*-marked referent and the open proposition are both active, as in (c) in Figure 2. Although (a) and (b) are the two dominant types of *ga* sentences in terms of the frequency, the sentences in (c) exhibit the unique characteristic of

ga that the focus pattern may be neutral with *ga* in terms of the RD measurement, leaving the assignment of focus up to the subtle cognitive state of the speaker and the addressee, as to among active referents, which are in the center of attention and which are not.

The results from the RD measurement have also suggested that the use of *wa* is more locally motivated than has been assumed in the literature, and in this sense this study agrees with the claims by Clancy and Downing (1987) that *wa* is a *local cohesive device*, i.e. the use of *wa* is largely locally motivated instead of marking thematicity at the global discourse level.[40] We have found a dominant tendency that the use of *wa* is motivated by activation structure at the local level rather than global topicality, though this study did not examine whether it is a unique aspect of spoken Japanese. Finally, the examination of the database revealed that there are instances of *wa*, though they constitute only a small portion of the database, in which the use of *wa* needs a structural account rather than an account by the notion of activation.

CHAPTER 4

POSTPOSING: ACTIVATION AND FOCUS

4.1. INTRODUCTION

The word order in Japanese has long been labeled as SOV (e.g. Greenberg 1963); however, it has been pointed out in a number of studies (e.g. Hinds 1983, Saito 1985, Shibamoto 1985, Simon 1989, Fujii 1991, Hudson 1993) that its word order is by no means rigid and its verb-final characteristic is often violated.[41] The non-canonical word order in Japanese can be classified into two types: word order variation in the preverbal string and word order variation which involves postverbal elements. The former is often labeled as *scrambling* and the latter as *postposing*.

The purpose of this chapter is to offer an analysis of the postposing construction in Japanese on the basis of the notion of activation. More specifically, this chapter will investigate the system of postposing in terms of the various activation patterns of sentences and attempt to provide a clue along with the findings in the previous chapters as to the use of *wa* and *ga* to capturing the interplay between morphosyntax and information dynamics in Japanese.

The discussion of this chapter will proceed as follows. Section 4.2 lays out word order variation in Japanese and characteristics of postposing construction. The previous studies of postposing are discussed in Section 4.3 in terms of two angles: syntactic constraints on postposing and functions of postposing. Section 4.4 examines postposing in mini-discourse, i.e. primarily question-answer pairs, in order to capture the system in terms of different activation patterns. In Section 4.5, the analysis is extended to the database of conversational Japanese on the basis of the same quantitative methods as used in Chapter 3. Finally, Section 4.6 summarizes the discussion of this chapter.

4.2. WORD ORDER VARIATION AND POSTPOSING IN JAPANESE

In spite of its typological characteristic SOV, Japanese exhibits rather flexible word order within the same clausal unit. In preverbal word order variation, the word order is flexible to a considerable extent. For example, Shibatani (1990: 260) lists the following sentence variants of scrambling for a simple ditransitive sentence.

(1)

a. Taro ga Hanako ni sono hon o yatta
DAT that book OBJ give:PST
'Taro gave that book to Hanako.'

b. Hanako ni Taro ga sono hon o yatta

c. sono hon o Taro ga Hanako ni yatta

d. ? Hanako ni sono hon o Taro ga yatta

e. ? sono hon o Hanako ni Taro ga yatta

f. Taro ga sono hon o Hanako ni yatta

As Shibatani notes, however, if there is more than one constituent fronted, as in (1d) and (1e), the sentence becomes awkward, though they are not totally ungrammatical. This ordering flexibility is also the case within an embedded clause, though the same awkwardness results if more than one constituent is fronted, as in (2b) and (2f). However, an element cannot be scrambled out of an embedded clause, as in (2g), (2h), and (2i).[42]

(2)

a. Taro ga [Hanako ga toshokan de Ken kara moratta] hon o yonda
library in from receive:PST book OBJ read:PST
'Taro read the book which Hanako received from Ken in the library.'

b. ? Taro ga [Ken kara toshokan de Hanako ga moratta] hon o yonda

c. Taro ga [Ken kara Hanako ga toshokan de moratta] hon o yonda

d. Taro ga [Hanako ga toshokan de Ken kara moratta] hon o yonda

e. Taro ga [toshokan de Hanako ga Ken kara moratta] hon o yonda

f. ? Taro ga [toshokan de Ken kara Hanako ga moratta] hon o yonda

g. * Hanako ga Taro ga [___ Ken kara toshokan de moratta] hon o yonda[43]

h. * toshokan de Taro ga [Hanako ga ___ Ken kara moratta] hon o yonda

i. * Ken kara Taro ga [Hanako ga toshokan de ___ moratta] hon o yonda

As in scrambling, postposing too exhibits flexibility to a considerable extent. As to this point, Simon (1989: 6) notes as follows.

> One noteworthy feature of postposed sentences is that virtually any constituent can appear in postverbal position; for example, NP, PP, AP, AdvP, demonstrative and conjunction, and a combination thereof. A postposed NP may be of any argument type, and it may be and often is, a genitive phrase. Likewise, a postposed PP may be of any semantic type. Another feature is that it is not unusual for two or three elements to appear here, although it is rare with four or more.

Below, I list some examples of postposing to illustrate the point above.

First, postposing is possible for any number of constituents, as in (3).

(3)

a. Hanako ga toshokan de Ken ni ageta yo hon o
library in DAT give:PST FP book OBJ
'Hanako gave a book to Ken in the library.'

b. Hanako ga toshokan de ageta yo Ken ni hon o[44]

c. Hanako ga ageta yo toshokan de Ken ni hon o

d. ageta yo Hanako ga toshokan de Ken ni hon o

(3a) illustrates postposing of the object NP, and indeed, postposing is possible for any type of constituent: for example, the *ga*-marked subject NP as in (3e), the indirect object NP as in (3f), the locative PP as in (3g).

(3)

e. toshokan de Ken ni hon o ageta yo Hanako ga

f. Hanako ga toshokan de hon o ageta yo Ken ni

g. Hanako ga Ken ni hon o ageta yo toshokan de

Also, postposing is possible for both *wa*-marked subject and non-subject, as in (4) and (5) respectively.

(4) uchi ni kaetta yo Hanako wa
home to return:PST FP

'Hanako went home.'

(5) France ni iku yo ryokoo wa
to go FP trip
'As for the trip, (I) will go to France.'

It is also the case with noun-modifying phrases or elements. For example, genitive phrases as in (6), demonstratives as in (7), and adjectival phrases as in (8).

(6) Hanako ga hon o yonda yo Ken no
book OBJ read:PST FP GEN
'Hanako read Ken's book.'

(7) kinoo tsuini eega o mita yo ano
yesterday finally movie OBJ see:PST FP that
'(I) finally saw that movie.'

(8) Hanako ga kuruma o katta yo sugoku ookii
car OBJ buy:PST FP very big
'Hanako bought a very big car.'

Postposing is possible with adverbials and PPs in general, as in (9), (10), and (11).

(9) nan no terebi miteru no ima
what GEN TV watching FP now
'What TV program are you watching now?'

(10) Hanako ga Ken o mita tte siken no toki ni
OBJ see:PST QT exam GEN time at
'(I heard) Hanako saw Ken during the exam.'

(11) kinoo daigaku ni kita yo kuruma de
yesterday university to come:PST FP car by
'(I) came to the university by car.'

Postposing is widely possible with complex sentences as well. For example, with coordinate clauses as in (12), adverbial subordinate clauses as in (13), and complement subordinate clauses as in (14).

(12) Hanako wa toshokan ni ita yo Ken wa uchi ni kaetta kedo
library in be:PST FP home to return:PST but
'Ken went home, but Hanako was in the library.'

(13) Ken wa terebi o miteta yo Hanako ga kita toki
TV OBJ watching FP come:PST when
'When Hanako came, Ken was watching TV.'

(14) Hanako wa itteta yo asita Ken to hanasu tte
saying:PST FP tomorrow with talk QT
'Hanako was saying that (she) will talk with Ken tomorrow.'

Postposing is also allowed with embedded clauses within complex NPs: noun complement clauses and relative clauses, as in (15) and (16) respectively.

(15) uwasa kiita yo Hanako ga kekkonsuru tteyuu
rumor hear:PST FP marry QT
'(I) heard the rumor that Hanako will get married.'

(16) Ken ga shasin o miseteta yo nihon de totta
photograph OBJ showing FP Japan in take:PST
'Ken was showing the photograph which he took in Japan.'

The discussion above does not provide possible postposing constructions exhaustively; however, given the examples above, it should be clear that any type of phrasal or clausal unit is in principle eligible to appear in the postverbal position.[45]

The status of postverbal elements is often controversial since it is not always clear as to the syntactic link between the preceding clause and the postverbal element. For example, some consider postverbal elements as *afterthoughts*, which are simply appended to the end of sentences (e.g. Kuno 1978b, Shibatani 1990). However, it has also been observed (e.g. Clancy 1982, Simon 1989) that there are two types of postposing. In one type, there is an intervening pause between the postverbal elements and the preceding string and there are two sentence-final intonations occurring first with the preceding clause and then with the postposed element. In the other type, there is no such intervening pause and the sentence including the postverbal unit is produced in one breath group with only one sentence-final intonation. Given the observation above, the postverbal unit in the former is considered to be afterthought, i.e. separate from the preceding unit, while the postverbal unit in the latter type is considered to be a part of the preceding unit. Claiming the necessity of separating the two types of postposing, Simon (1989: 43) states the following.

> An "afterthought" analysis may be well-motivated in cases in which a considerable pause intervenes between the verb and the postverbal element so that there is sufficient time for the speaker to reflect on the statement he/she originally makes or to monitor the hearer's reaction. It does not seem valid, however, in many other cases when the postverbal elements are added following little or no pause.

Following Simon (1989), the present study distinguishes the two types of postposing, and for the purpose of the present study a postverbal unit is considered to be part of the preceding utterance only if there is no intervening pause between the two, and the analyses of postposing in mini-discourse and conversational Japanese in the following sections will focus only on postposing with no intervening pause, hence the examples of postposing construction for discussion are all intended to be postposing of non-pause type, unless noted otherwise.

Finally, the flexible word order, including postposing as well as scrambling, is one of the striking characteristics of spoken Japanese, especially informal speech. In this point, Clancy (1982: 67-68) states as follows.

> "In the written narratives, SOV order was consistently maintained, but the word order in the oral narratives was much more flexible, with subjects, objects, and many other constituents appearing after the verb. Similarly, the writers always placed relative clauses and other modifiers before head nouns, whereas speakers sometimes produced a noun first, and then added one or more modifiers. The postposed word orders which appeared in this sample of oral narratives are typical of conversational Japanese, which rather frequently fails to exhibit canonical verb-final word order."

Supporting this greater flexibility of word order in spoken Japanese, Peng (1977) notes that there were 288 cases of postposing, which came to 9.2% of all utterances in the data elicited from interviews, Simon (1989) reports that there were 207 cases of postposing in her 125-minute conversation data, and Fujii (1991) observes 215 cases of postposing in her 6-hour conversation data from TV talk shows.[46] The 140-minute conversation data for the present study contains 119 cases of postposing in which there is no intervening pause between the postverbal unit and the preceding unit.

Regarding the types of postverbal units, the 119 cases of postposing in the database consist of the following categories.[47]

Table 1: Postverbal elements by category and token distribution

Types of postverbal unit	# of tokens
subject	37
adverbial	25
PP	20
adverbial clause	16
coordinate clause	8
object	7
topic NP	3
genitive	2
noun-modifying clause	1
Total	119

The types of postverbal elements and the token distribution patterns here roughly correspond with those in Peng (1977), Shibamoto (1985), and Simon (1989) in that subjects, adverbials, and adverbial clauses are among those which appear in the postverbal position most frequently.

4.3. PREVIOUS APPROACHES TO POSTPOSING IN JAPANESE

In general, there are two types of studies of postposing construction in Japanese in terms of their goals: (i) the structural aspects of postposing, including its formation and constraints, and (ii) the discourse functions of postposing. This section will summarize major previous works of both types; first, the syntactic issues involved in postposing, primarily in terms of constraints on postposing, and next, functions of postposing in discourse.

4.3.1. Constraints on Postposing

The non-canonical word order in Japanese has been studied rather extensively in terms of its formation and constraints on variation by both formalist theories (e.g. Inoue 1978, Kuroda 1980, Saito 1985, Simon 1989) and functionalist approach (e.g. Kuno

1978a, 1978b, 1980, Simon 1989, Hudson 1993). As discussed in the previous subsection, Japanese exhibits postverbal word order variation to a considerable extent. However, the flexible word order does not mean that any element of a sentence can appear anywhere. One obvious, but important constraint is that postverbal elements must be lexical items or items which can appear independently. Namely, non-independent elements, including so-called particles and the copula, cannot be postposed apart from their 'head' elements. For example, the object marker cannot be separated from the object, as shown in (17b) and (17c).

(17)
a. kinoo okane o moratta
yesterday money OBJ receive:PST
'(I) received the money yesterday.'

b. * kinoo okane ___ moratta o

c. * kinoo o okane ___ moratta

Another constraint on postposing is what Simon (1989: 102), a recent comprehensive study of postposing, explicitly states in the Government and Binding [GB] framework. Proposing a movement analysis for postposing, she states that only a maximal projection can be postposed; a non-maximal projection, including a head, cannot be moved. Simon observes that this constraint applies to any level of maximal projection; NP can be postposed but not N alone, S' but not S alone, AP but not A alone, AdvP but not Adv alone.[48] Simon (1989) cites examples in (18) and (19) to illustrate this point in terms of NP.

(18) [$_{NP1}$ ___ [$_{N}$ shoosetsu]] o yonda no [$_{NP2}$ furansugo no]
novel OBJ read:PST FP French GEN
'(I) read a French novel.'

(19) * [$_{NP1}$ [$_{AP}$ mizikai] [$_{NP2}$ furansugo no] ___] yonda no [$_{N}$ shoosetsu] o

short French GEN read:PST FP novel OBJ
'(I) read a short French novel.'

According to Simon, postposing is possible in (18) since the postposed element is the whole genitive NP, which is a maximal projection, while the postposed element in (19) is the head noun alone; therefore, it is ungrammatical.[49]

The studies of the syntactic properties of postposing share the central issues with those of scrambling in that they both face a necessity of accounting for the possible types of non-canonical word order. In analysis of postposing, Simon (1989) bases her theoretical assumption on Saito's (1985) movement analysis of scrambling in the GB theory. Saito, in turn, supports Harada's (1977) transformational analysis of scrambling and argues that scrambling is a case of a Move-alpha.[50] Characterizing scrambling, Saito claims that scrambling can adjoin phrases to any node, move any maximal projection, and occur successive-cyclically and its bounding properties are straightforwardly accounted for by the Subjacency condition (e.g. Chomsky 1973). As already noted in Harada (1977), *long-distance* scrambling is subject to the basic constraints of subjacency. For example, scrambling out of a relative clause and a nominal complement clause causes ungrammaticality or at least awkwardness, as illustrated in (20) and (21) from Saito (1985: 246).

(20)?* ano hon o [$_S$ John ga [$_{NP}$ [$_S$ ___ katta hito]] o sagasiteiru rasii]
that book OBJ buy:PST person OBJ looking-for seem
'It seems that John is looking for the person who bought that book.'

(21) ? Bill o [John ga [$_{NP}$ [$_S$ Mary ga ___ saketeiru toyuu] uwasa] o kiita]
OBJ avoiding QT rumor OBJ hear:PST
'John heard a rumor that Mary is avoiding Bill.'

For her analysis of postposing, Simon (1989) proposes a similar line of argument by suggesting a movement rule with the subjacency constraint.[51] According to Simon, NP and S' are bounding nodes for subjacency in Japanese; therefore, no element can move across more than one bounding node in a single movement for postposing. Simon (1989)

discusses the following points to support the movement analysis for postposing. (Examples below are from Simon (1989).)

A sentence is ungrammatical or awkward if an element is postposed out of a relative clause (crossing S' and NP).

(22) * [$_{NP}$[$_{S'}$ kinoo ___ katta] dress] o mita no Mari ga
yesterday buy:PST OBJ see:PST FP
'(I) saw the dress Mari bought yesterday.'

A sentence is ungrammatical or awkward if an element is postposed out of a nominal complement clause (crossing S' and NP).

(23)?? [$_{NP}$[$_{S'}$ Ken ga ___ kekkonsuru tteyuu] uwasa] kiita yo Mari to
marry QT rumor hear:PST FP with
'(I) heard the rumor that Ken is marrying Mari.'

A sentence is ungrammatical or awkward if an element is postposed out of a coordinate NP (crossing two NPs).

(24) * [$_{NP1}$ [$_{PP}$ [$_{NP3}$ France no [$_{N}$ sizin]] to] [$_{NP2}$ ___ [$_{N}$ isha]]] ga shoo o moratta
GEN poet and doctor award OBJ
receive:PST

no yo America no
NOM FP GEN
'A French poet and an American doctor received the awards.'

Despite the basic predication by the subjacency constraint above, Simon (1989) admits that there are cases which subjacency fails to account for. First, although a sentence is grammatical if an element is postposed out of a sentence-final clause, whether it is a matrix or adverbial subordinate clause, a sentence is ungrammatical if an element is postposed out of a non-final matrix or adverbial subordinate clause, i.e. the element crosses only one S', as shown in (25).

(25)
a. * [$_{S1}$ Ken ga ___ hataraite]-te, [$_{S2}$ okusan ga uti ni iru no yo] kaisha de
work and wife home in be NOM FP office in
'Ken works in the office and (his) wife stays home.'

b. * [$_{PP}$ [$_{S'}$ ___ mita] atode] piano o renshuusuru yo terebi o
see:PST after OBJ practice FP TV OBJ
'(I) will practice piano after (I) watch TV.'

Secondly, a sentence is grammatical even though an element is postposed out of a predicative nominal of copulative sentences, i.e. it crosses S' and NP.

(26) ano hito wa [$_{NP}$ [$_{S'}$ ___ tonari ni hikkosite kita] isha] da yo kyonen
that person next-door in move come:PST doctor COP FP last-year
'That person is a doctor who moved in next door last year.'

As mentioned above, the properties of postposing have been examined from a functional perspective as well. I will discuss previous studies on discourse functions of postposing in the next section; therefore, here I will briefly discuss the previous functional approach particularly to the structural constraints on postposing.

Although Simon (1989) provides a formalist analysis of postposing, a part of the study is solely devoted to discussion of constraints on postposing from a semantic point of view in order to complement the formalist analysis, i.e. in order to account for the cases which the notion of subjacency fails to predict. Here I cite basic concepts from Hudson (1993), which is in principle an extension of the semantic analysis in Simon (1989), primarily utilizing the same types of notions.

Hudson (1993) uses three types of notions: *transparency of verbs*, *right association*, and *linguistic distance*. Here the notion of transparency is used in Kuno's (1976) sense, i.e. the degree of genericness of an element. On the basis of this notion, Hudson (1993: 10) states "we can say that the copula is the most transparent, that existential verbs are the second most transparent, and that specific action verbs and adjectives are the least so." Those different degrees of transparency of predicates are used to account for the gradient acceptability of postposing in the following, which Hudson (1993: 9) cites, where the subjacency constraint fails to predict, i.e. they are all postposing out of a relative clause.

(27) kore wa [$_{NP}$[$_{S'}$ ___ kinoo mottekita] wain] da yo Mari ga

this yesterday bring:PST wine COP FP
'This is the wine which Mari brought yesterday.'

(28)?* [$_{NP}$[$_{S'}$___ kinoo mottekita] wain] ga aru yo Mari ga
yesterday bring:PST wine exist FP
'There is wine which Mari brought yesterday.'

(29) * [$_{NP}$[$_{S'}$___ kinoo mottekita] wain] ga oisii yo Mari ga
yesterday bring:PST wine delicious FP
The wine which Mari brought yesterday is delicious.'

(30) * [$_{NP}$[$_{S'}$___ kinoo mottekita] wain] o nomu yo Mari ga
yesterday bring:PST wine OBJ drink FP
(I) will drink the wine which Mari brought yesterday.'

Postposing in (27) is completely acceptable, while that in (29) and (30) is completely unacceptable, and (28) is marginally permissible in this regard. Assuming the different degrees of transparency of predicates, Hudson (1993: 10) claims that "more transparent the semantic content of the main clause is, the easier it is to postpose from a complex NP."[52]

Independent of the notion of transparency, Hudson (1993) proposes processing constraints on postposing, in terms of the *principle of right association* in Kimball (1975: 159), i.e. "the tendency in a parse is always to associate the terminal immediately inputted as a constituent of the lowest, rightmost non-terminal." In order to account for the general pattern that postposing is possible only out of the right-most clause, regardless of the clause type, Hudson (1993: 8) explains that it is difficult to postpose an element out of a non-initial clause due to the difficulty of association between the postposed unit and the clause which the gap because the clause with the gap is "shipped off to semantics before the rest of the sentence is parsed." The following are some of the examples which Hudson (1993: 7-8) cites to support the claim.

(31) [$_{S1}$ Taro-chan wa soto de asondeta] kedo
outside at playing:PST but

[$_{S2}$ Jiro-chan wa ___ ita no] uchi ni
be:PST FP home at

'Taro was playing outside, but Jiro was at home.'

(32) * [$_{S1}$ Taro-chan wa ___ asondeta] kedo
playing:PST but

[$_{S2}$ Jiro-chan wa uchi ni ita no] soto de
home at be:PST FP outside at
'Taro was playing outside, but Jiro was at home.'

Postposing is possible out of the final clause as in (31), while it is not out of the non-final clause, as in (32).

The same pattern is found in postposing out of an embedded clause. Despite the subjacency constraint that postposing is not allowed out of a relative clause, as shown in (34), postposing is possible out of a relative clause which is predicated by copula alone, as in (33), and sentences are not as bad as those as in (34) if the whole embedded clause is postposed, as in (35).

(33) kore wa [$_{NP}$[$_{S'}$ ___ kinoo kabutteta] boosi] da yo Ken ga
this yesterday wearing:PST hat COP FP
'This is the hat which Ken was wearing yesterday.'

(34) * [$_{NP}$[$_{S'}$ ___ kinoo kabutteta] boosi] wa kore da yo Ken ga
yesterday wearing:PST hat this COP FP
'The hat which Ken was wearing yesterday is this one.'

(35) ? ___$_i$ kore da yo [$_{NP}$[$_{S'}$ ___$_j$ kinoo kabutteta] boosi] wa$_i$ Ken ga$_j$
this COP FP yesterday wearing:PST hat
'The hat which Ken was wearing yesterday is this one.'

Summarizing her observation, Hudson (1993: 8) states "we cannot make a blanket statement to the effect that postposing is barred out of a relative clause or that Subjacency is at work, since it is allowed when the clause appears in rightmost position, directly preceding the postposed element."

One other notion which Hudson (1993) appeals to is *linguistic distance* (Haiman 1983). Hudson uses this notion to account for the following observation. Postposing is normally possible out of sentential complements; however, the longer the distance

between the complement clause and the end of the sentence, the less acceptable the sentence becomes. The following are from Hudson (1993: 11).

(36) [$_{S'}$[$_{S}$ ___ akeru] no] o mita no kinko o
open NOM OBJ see:PST FP safe OBJ
'(I) saw (him/her) opening the safe.'

(37) ? [$_{S'}$[$_{S}$ ___ akeru] no] o tasikani hakkiri mita no kinko o
open NOM OBJ surely clearly see:PST FP safe OBJ
'(I) surely clearly saw (him/her) opening the safe.'

(38)?? [$_{S'}$[$_{S}$ ___ akeru] no] o tasikani kono me de hakkiri mita no kinko o
open NOM OBJ surely this eye with clearly see:PST FP safe OBJ
'(I) surely clearly saw (him/her) opening the safe with my own eyes.'

The assumption behind the notion of linguistic distance is that the more linguistic distance there is, the more difficult it is to associate the postposed unit and the original unit with a gap. However, Hudson (1993: 12) adds "the distance factor is perhaps nothing more than a tendency, however. It seems to come into play only when postposing is already allowed in other domains. ... even when the distance is relatively short, postposing out of a complex subject NP, for example, results in an unacceptable sentence."

As briefly touched on in this section, the issues of constraints on postposing have been investigated from different angles, such as the structural, semantic, and cognitive perspectives, and it has been suggested (e.g. Simon 1989) that it is only by those different perspectives together that the properties of postposing can be fully captured. In the following section, I will discuss some of the major previous works on discourse functions of postposing.

4.3.2. Discourse Functions of Postposing

Study of discourse functions of non-canonical word order has been done much more extensively for postposing constructions than scrambling, i.e. word order variation within the preverbal string. To my knowledge, Kuno (1978a) is the earliest to claim the

correlation between the word order and the information structure in Japanese. On the basis of his observation of question-answer pairs, Kuno (1978a: 68) states that the postverbal elements are either (i) the elements which the speaker ellipsed in the preverbal string because they are recoverable by the preceding context but adds at the end of the utterance to confirm the message conveyed, or (ii) supplementary information. Paraphrasing the principle, Kuno adds that utterance with postverbal elements should convey the intended message without the postverbal elements.

On the basis of oral narratives, Clancy (1982: 69) makes the similar line of argument to Kuno's (1978a) notion of *recoverable* by stating "postposing was apparently used to defocus either familiar or easily deducible information which was in some way semantically subordinate to the material preceding the main verb."

Contrary to the position that the postverbal elements are recoverable or deducible; therefore, they are in some way "redundant", Hinds (1982) suggests that postverbal elements are the "indispensable" part of an utterance by proposing the following functions of postposing: (i) resolution of ambiguity, i.e. elements are inserted in the postverbal position because the speaker realizes that the hearer may not be able to comprehend the intended message without mentioning the referent expressed by the postverbal elements, or (ii) emphasis, i.e. postverbal elements are obvious from context but the speaker inserts the elements to place some sort of emphasis on the utterance.

Maynard's (1989: 36-37) analysis is similar to Hinds' above in that she states that the speaker adds presupposed information in the postverbal position for the hearer to fully comprehend the information. In addition to this, however, Maynard adds two functions of postposing: The speaker places less salient information in the postverbal position, and the speaker places some expressions of doubt or hesitation in the postverbal slot so that the utterance becomes pleasing for the hearer to receive.[53] Overall, Maynard uses two different notions to account for postposing: *presupposition* and *saliency*. The notion of presupposition is similar to Kuno's (1978a) notion of recoverable discussed above. As for

the notion of saliency, Maynard (1989: 35) states "when a speaker introduces two pieces of totally new information which are neither familiar nor easily deducible, one may be chosen to be postposed simply because the piece of information the postposed element bears is not considered as important or relevant as the other." In this regard, postposing is considered to have a function of "backgrounding the postposed pieces of information and foregrounding the other" (Maynard 1989: 36).

The combination of deducibility and importance is also found in Simon (1989). For postposing with no preceding pause before the postverbal elements, Simon (1989: 189) proposes a principle labeled as *Important Information First*; namely, "these sentences [with postverbal elements] are simply results of important or urgent information coming to the speaker's mind first and thus being vocalized first, especially under time pressure."[54] Simon (1989: 193) points out that the principle is best illustrated in question-answer pairs, illustrating with the following example.

(39)
A: anmari itta koto nai no buffet style
much go:PST NOM NEG FP
'Haven't you been to buffet style much?'

B: nai wa sonna no
NEG FP that-kind NOM
'That, I haven't.'

In (39B), the speaker places the predicate first, which answers the question, and therefore more important than the other, and places the secondary information in the postverbal position. The notion of saliency or importance as found in Maynard (1989) and Simon (1989) seems plausible, given the fact that there are cases in which the postverbal elements are *new*, hence not deducible or recoverable, where the defining postverbal information as *old* fails. After investigating the conversational Japanese in terms of the notion of activation, I will come back to the notion of saliency/importance in order to see how and to what extent the notion of activation and saliency/importance complement each other.

Before going on to the notion of activation in postposing, I discuss one other recent study of postposing because its data base is similar to that of the present study. On the basis of conversational Japanese, Fujii (1991) investigated discourse factors for postposing in terms of a number of different measurements. Fujii concludes that there are two types of postposing: one in which the postverbal elements are eligible for ellipsis and one in which the postverbal elements are not eligible for ellipsis. In the former, the speaker simply adds information which the speaker wants to confirm or get hearer's attention to. The latter is further classified into two types in terms of the place of *focus*: one in which the preverbal string is more focused than the postverbal elements and one in which the postverbal elements are more focused than the preceding elements. In the former, the postposing results from asserting first the focused elements, and in the latter the postverbal elements serve to rectify the ambiguity caused by the speaker's mistakenly dropping the elements in the preverbal string.

To define the notion of *focus*, Fujii uses Payne's (1985, 1990) notion of *pragmatic markedness*. Namely, elements have focus if they exhibit characteristics such as *single focus contrast* (assertion of missing or incorrectly assumed information), *multiple foci of contrast* (assertion of a correct match-up between two or more pairs of items), *counter expectation* (assertion of information counter to culturally, situationally, or textually expected presuppositions), *restatement* (restatement of previously mentioned information), *added detail restatement* (assertion of additional information which amplify the background assumption), *questions and answers to questions*, and *negation*.

Although it is noteworthy that Fujii (1991) attempts to give a systematic discourse analysis of postposing constructions on the basis of conversational Japanese, there are some points which need further clarification. First, it is not clear how Fujii defines the postposing construction in terms of the prosodic characteristics. As discussed earlier, there are at least two types of constructions which involve postverbal elements: postverbal elements preceded by a pause and those without a preceding pause (cf. Clancy 1982,

Simon 1989). Although functional distinctions between them are not entirely clear, postverbal elements preceded by a pause seems to be distinct from those without a preceding pause in that in the construction with a considerable intervening pause between the predicate and the postverbal elements, the speaker has sufficient time to reflect on the utterance which the speaker has made; therefore, whatever elements added to the utterance in the postverbal position tend to be the results from the speakers reflection of the original utterance, hence, some sort of modification or addition of information to the original utterance. It may be the case that one of the postposing types in Fujii (1991) which repairs the speaker's mistaken ellipsis is typically the pause type.

The second problem may be related to the point above. Although Fujii distinguishes the two types of postposing among those whose postverbal elements cannot be ellipsed, her characterization of one type as the speaker's *repair* strategy is not entirely convincing. According to Fujii, unlike the postposing with the focus on the preverbal string, the postposing with the focus on the postverbal elements is characterized by the speaker's correction or modification of the utterance which has been made. However, the identification of the two types is in principle based on the location of focus, i.e. either on the preceding elements or postverbal elements, and a solid criterion to identify the speaker's correction or modification of the utterance is not given in the analysis; therefore, the analysis does not give a clear-cut explanation as to why the focus falls on the postverbal elements in the particular postposing type.

Finally, although Fujii assumes correlation between the reversed word order and the speaker's and hearer's cognitive status, it is only in the judgment of eligibility for ellipsis of elements that Fujii uses the cognitive notion, i.e. notion of saliency, and her notion of focus is entirely based on the semantics of utterance elements, i.e. Payne's (1985, 1990) notion of pragmatic markedness.

The functions of postposing in Japanese have been one of the central issues in discourse and conversation analysis; however, the studies of postposing have not reached

a consensus due to different theoretical backgrounds. Despite the diversity of frameworks, the previous studies are in general centered on the following points.

(i) The postverbal elements represent *old* information which is *presupposed*, *deducible*, *recoverable*, etc.[55]

(ii) The postverbal elements represent less *important* or *relevant* information than the information represented by the preceding preverbal elements.

(iii) The postverbal elements represent information which plays a role to disambiguate or modify the information represented by the preceding elements.

The present study investigates postposing constructions in Japanese in terms of a different framework. Unlike the previous studies which solely examine the semantics of postverbal elements, the present study investigates information represented by postposing constructions in terms of the notion of activation, as in the study of *wa* and *ga* in the previous chapters, and examine how and to what extent the notion captures the system of postposing.

4.4. POSTPOSING IN MINI-DISCOURSE

Before discussing the quantitative analysis of postposing in conversational Japanese, this section examines postposing on the mini-discourse level, mostly question-answer pairs, and lays out the basic principle of postposing on the basis of the notion of activation and the basic assumption that non-focus involves activated information.

Examples which most clearly illustrate the condition for postposing are question-answer pairs, as in (40).

(40)

A: kinoo no kaigi wa dare ga kimasita ka
yesterday GEN meeting who come:PST Q
'Who came to yesterday's meeting?'

B: Ken ga kimasita yo
come:PST FP

'Ken came.'

B': # kimasita yo Ken ga
'Ken came.'

'Ken' in (40B) is not mentioned in the previous context; so it is nonactive information in the immediately preceding context, while the open proposition *X came* is active by being mentioned. In (40B'), the sentence with *Ken* in the postverbal position is inappropriate, where the active information is followed by the nonactive information in the postverbal position. On the other hand, if there is preceding context which makes 'Ken' active, the sentence with the postverbal element becomes acceptable, as in (41).

(41)
A: Ken osoi nee
late FP
'Ken is late.'

B: a Ken ga kimasita yo
ah come:PST FP
'Ah, there Ken comes.'

B': a kimasita yo Ken ga
'Ah, there Ken comes.'

'Ken' is mentioned in (41A), hence it is activated in the consciousness of speaker B; therefore, *Ken* may be either in the preverbal canonical position or in the postverbal position, as shown in (41B) and (41B'). The open proposition *X came* in (41B) is not active in the preceding context; therefore, in the postposing construction, i.e. (41B'), the nonactive information precedes the active information.

On the other hand, the postposing construction becomes less natural if none of the elements of the sentence, including the postverbal element, is active in the preceding context, which is shown in (42).

(42) [speaking to someone who is not expecting anyone to come]

a. Ken ga kimasita yo
come:PST FP
'Ken has come.'

b. ? kimasita yo Ken ga

In the immediately preceding context of (42), neither 'Ken' nor the open proposition *X came* is active and in such context the postposing construction becomes awkward. (43) illustrates this point with another example. (43B') is awkward with the postverbal element and the preceding element equally nonactive.

(43)
A: doo sita no?
how did FP
'What's the matter?'

B: ame ga hutteru yo
rain falling FP
'It's raining.'

B': ? hutteru yo ame ga
'It's raining.'

Given the observation above, I lay out the degrees of acceptability of postposing constructions in terms of three different focus structures as in Lambrecht (1986, 1987, 1988, 1994) as follows. (See Chapter 6 for discussion of focus structure in Lambrecht.) In so-called *predicate-focus* constructions, where the predicate is the focus of the sentence, as in (41B), the active information may be in the postverbal position. In postposing construction of this type, nonactive information precedes postverbal active information. In so-called *argument-focus* structure, where a particular argument is the focus of the sentence, as in (40B), the focused argument cannot be in the postverbal position, following the nonfocus part of the sentence. In so-called *sentence-focus* structure, where the whole sentence is the focus, as in (42) and (43B), the postposing construction is not appropriate, or at least the canonical order without postverbal elements is more natural than the postposing construction. So, postposing in predicate-focus sentence is more acceptable than that in sentence-focus sentence, and postposing in sentence-focus is in

turn more acceptable than in argument-focus sentence. The different degrees of acceptability of postposing in terms of the three focus types are summarized as follows.

(44) *Predicate Focus*

A: Ken osoi nee
late FP
'Ken is late.'

B: kimasita yo Ken ga
come:PST FP
'Ken has come. (There Ken comes.)'

(45) *Sentence Focus*

A: doo sita no?
how do:PST FP
'What's the matter?'

B: ? kimasita yo Ken ga
come:PST FP
'Ken came.'

(46) *Argument Focus*

A: kinoo dare ga kimasita ka
yesterday who come:PST Q
'Who came yesterday?'

B: # kimasita yo Ken ga
come:PST FP
'Ken came.'

The different degrees of acceptability of postposing above is accounted for by the relative degrees of activation of the postverbal elements and the preceding elements. Postposing is most acceptable when the postverbal active information follows nonactive information, as in the predicate focus-construction, and it is least acceptable when the

postverbal nonactive information follows active information, as in the argument-focus construction, and the sentence-focus construction falls in the middle, where the postverbal information and the preceding information are equally nonactive. The acceptability hierarchy above implies the basic principle that elements in the postverbal position should be more active than those in the preceding string, and postposing in argument-focus is exactly the opposite of the ideal pattern stated in the principle, hence it is the least acceptable case. Figure 1 illustrates this point.

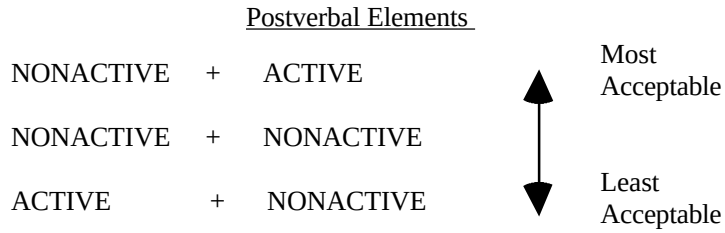


Figure 1: Acceptability hierarchy for postposing in terms of focus types

In the hierarchy above, there is another logical possibility of the active-nonactive combination, namely, a case in which postverbal active information follows another active information, where the sentence repeats a proposition which is mentioned in the preceding discourse or the sentence represents a new proposition by combining active propositions in the preceding context. (47) illustrates discourse context in which a postposing construction repeats a proposition active in the preceding context.

(47)
A1: kyoo Ken ga kuru tte
today come QT
'(I heard) Ken will come today.'

B1: a soo
oh so
'Is that so?'

[sound of doorbell]

[looking through the window]
B2: a Ken ga kita yo

ah come:PST FP
'Ah, Ken has come. (There Ken comes.)'

B2': a kita yo Ken ga
'Ah, Ken has come. (There Ken comes.)'

In (47B1), both 'Ken' and the open proposition *X comes* are active, being mentioned in (A1), and the postposing construction, i.e. (B2'), is as good as the canonical sentence, i.e. (B2). The observation above implies an independent principle in addition to the hierarchy in Figure 1 that the information represented by the postverbal elements should be active, and I add the forth possibility to the hierarchy as in Figure 2. The hierarchy in Figure 2 indicates two separate principles for the optimal conditions for postposing. First, the information represented in the postverbal position should be active in the immediately preceding context, which accounts for the two less acceptable types in which nonactive information is followed by postverbal nonactive information and active information is followed by postverbal nonactive information. Second, the information represented by postverbal elements should be more active than that represented by preceding elements. This principle accounts for the observation that the nonactive+active case is better than the other two which are lower in the hierarchy and the nonactive+nonactive case, in which the two are equal in activation, is better than the active+nonactive case, which is a mirror image of the optimal pattern.

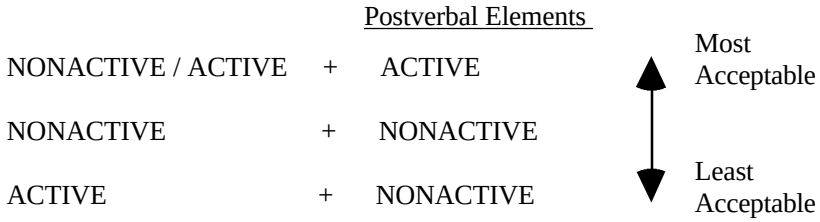


Figure 2: Acceptability hierarchy for postposing in terms of focus types (Revised)

Given the hierarchy above, I further examine postposing constructions in connection with sentence-final particles. Japanese has a rich inventory of sentence-final

particles, which is known as *shuujoshi*, and it has been observed (e.g. Tokieda 1950, Chafe 1982, Clancy 1982, Maynard 1989) that spoken Japanese is characterized by the frequent use of sentence-final particles. The sentence-final particles have a wide variety of discourse functions and are used to express the speaker's judgment about the utterance. For example, Martin (1975: 914) states that the final particles are used to "impart some additional hint of the speaker's attitude toward what he is saying - doubt, conviction, caution, inquiry, confirmation or request for confirmation, recollection, etc." The final particles *ne(e)* and *na(a)*, for example, reflect "the speaker's consideration of the addressee, and the addressee feels more participation in the conversation with mutual understanding. Thus, these particles may be called particles of rapport" (Uyeno 1971: 132).

The sentence-final particles of our interest here are those which place emphasis on the elements preceding the particles. We have seen above that postposing of the sentence-focus type normally results in an unnatural sentence, even though it is not as bad as postposing of the argument-focus type. However, the acceptability of the sentence varies according to particular final particles attached to the predicate. The postposing in a sentence-focus construction in (48B) illustrates this point.

(48)
A: doo sita no?
how do:PST FP
'What's the matter?'

B: kita {no/?yo/?ne/?na} Ken ga
come:PST FP
'Ken came.'

In (48B), the postposing construction is acceptable with the final particle *no*, while it is not with particles such as *yo*, *ne*, and *na*. Martin 1975 observes that the final particle *no* has the function of placing exclamations on the sentence, while *yo*, *ne*, *na* do not.[56] The use of *no* in (48B) gives the effect of emphasizing the proposition expressed by the

predicate, such as expression of speaker's surprise, anger, joy, etc. caused by Ken's coming. It should be noted that prosodic emphasis on the predicate also makes the postposing construction more acceptable even though it is the sentence-focus type.

A similar effect of sentence-final elements on the acceptability of postposing is observed in the use of a combination of the nominalizer and the copula *no da* or *n da*, which is so-called *explanatory* ending.[57]

(49)
A: doo sita no?
how do:PST FP
'What's the matter?'

B: kita n desu Ken ga
come:PST NOM COP
'(It's that) Ken came.'

The sentence-final construction *no da*/*n da* has an effect of placing emphasis on the preceding elements (Maynard 1990: 237), and the use of this particular construction in (49B) gives the same effect as in (48B).

The observation above leads us to the assumption that the degree of emphasis on sentence elements does play a role in determining the acceptability of postposing construction, and this assumption is in favor of the analysis of postposing in terms of importance or saliency, as found in previous studies, since the speaker's emphasis on a particular element reflects the speaker's assumption as to the degree of importance on the element; namely, the speaker places emphasis on elements because the speaker assumes those elements to be important. However, it should be noted that the effect of emphasis by final particles and the *no da*/*n da* constructions on the acceptability of postposing is found only in utterances of the sentence-focus type.

(50)
A: kinoo dare ga kita no?
yesterday who come:PST FP
'Who came yesterday?'

B: kita {#no/#yo/#ne/#na/ #n desu} Ken ga
come:PST FP NOM COP
'Ken came.'

The postposing construction in (50B) is an argument-focus type, in which the open proposition *X came* is active and the element *X* is nonactive in the preceding context. Here the postposing is unacceptable regardless of the type of predicate-final construction, including the particle *ne* and *na*, which "soften a statement and invite confirmation" (Martin 1975: 916) rather than place an emphasis.

Having examined the role of sentence-final particles in postposing, it should be clear that the notion of importance does play a role in determining the acceptability of postposing. However, it is not only the use of final particles that reflects the degree of importance of particular sentence elements. In fact, the acceptability hierarchy for postposing proposed in Figure 2 does reflect the acceptability of postposing in terms of the relative degree of importance of postverbal and preceding elements. Defining the notion of importance in light of the goal of the utterance, i.e. what is important is important information relative to the goal of the utterance in the surrounding discourse, we may assume that what is nonactive is always more important than what is active, and furthermore, the less active a referent or a proposition is, the more important the information represented by the referent or proposition is. Given this background, the acceptability hierarchy for postposing in Figure 2 is summarized with one general principle stated in (51).

(51) The postposing construction is acceptable if and only if the element in postverbal position is less important than the preceding elements.

In the nonactive+active cases in the hierarchy, the preceding elements are obviously more important than the postverbal elements; therefore, it is perfectly acceptable. In the active+nonactive cases, on the other hand, the postverbal elements are more important than the preceding elements, which is the complete violation of the principle; therefore, it is the least acceptable type. In the nonactive+nonactive cases, the

activation pattern does not differentiate the postverbal and preceding elements in terms of degree of importance; therefore, postposing of this type is in general not perfectly acceptable. However, postposing of this type becomes acceptable if the preceding elements are marked more important than the postverbal elements either morphologically or prosodically, or both. The active+active cases are similar to the nonactive+nonactive ones in that the notion of activation does not play a role in determining the relative degree of importance. However, the preceding elements in the active+active type may be considered more important than the postverbal elements, relative to the goal of the utterance, given the fact that the postverbal elements of this postposing type are normally eligible for ellipsis, while the preceding elements are not. (47') shows an example to illustrate this point. It should be noted that neither the postverbal elements nor the preceding elements are eligible for ellipsis in the nonactive+nonactive type, which is shown in (48').

(47')

A1: kyoo Ken ga kuru tte
today come QT
'(I heard) Ken will come today.'

B1: a soo
oh so
'Is that so?'

[sound of doorbell]

[looking through the window]

B2: a kita yo {<u>Ken ga</u> / }
ah come:PST FP
'Ah, Ken has come. (There Ken comes.)'

B2': a {kita yo / # } <u>Ken ga</u>
'Ah, Ken has come. (There Ken comes.)'

(48')
A: doo sita no?
how do:PST FP
'What's the matter?'

B: kita no {<u>Ken ga</u> / # }
come:PST FP
'Ken came.'

B': {kita no / # } <u>Ken ga</u>
'Ken came.'

I should note that the principle stated in (51) is not an original one; as discussed in Section 4.3.2, it has been observed in previous studies (e.g. Maynard 1989, Simon 1989) that the postverbal elements represent less important or relevant information than the information represented by the preceding elements.[58] However, the present study breaks down the general principle into the four activation patterns which are shown by the hierarchy in Figure 2, and the classification in terms of the four activation types becomes useful in order to investigate the postposing constructions in the conversational database, which will be discussed in the following section.

Finally, the present study favors the analysis of postposing in terms of the notion of importance, in connection with the notion of activation, over the analysis of postposing in terms of *new*/*old* distinction, i.e. the postverbal elements represent old information which is presupposed, deducible, recoverable, etc. For example, Kuno (1978a) suggests that importance of information is defined by the notion of recoverability; namely, information is important if it is not recoverable by the preceding context. In this view, Kuno argues that information which is not recoverable cannot appear in the postverbal position. However, there are cases in which the postverbal elements represent unrecoverable information, as in (52.2).

(52)
1. kinoo Hanako to Ken to Taro ga uchi ni kita yo
yesterday and and home to come:PST FP
'Hanako, Ken, and Taro came to my house yesterday.'

2. sorede hon o takusan mottekita yo Ken to Taro ga
and book OBJ many bring:PST FP
'And Ken and Taro brought many books.'

In (52.2), *Ken and Taro* in the postverbal position is *new* in that the particular set [Ken, Taro] is not recoverable by the preceding context, even though each individual of the set is *old*, being mentioned in (52.1). I should note that the fact that the postposing construction is acceptable in (52.2) is accounted for straightforwardly by the notion of importance in terms of activation. The particular set [Ken, Taro] is not the center of attention since the set is not present in the preceding context; however, each member of the set is active since they are mentioned in (52.1), hence, the set [Ken, Taro] is at least more active than the proposition *X brought many books*, which is not activated at all in the preceding context, and in this sense, the former is less important than the latter.

4.5. POSTPOSING: QUANTITATIVE ANALYSIS OF CONVERSATIONAL JAPANESE

As discussed in the previous section, the principle in (51) predicts the acceptability of the postposing construction from the functional perspective, primarily on the basis of the notion of activation. The purpose of this section is to see whether the principles drawn from the mini-discourse data are valid in the database of conversational Japanese.

The discussion for this section will proceed as follows. Section 4.5.1 describes the quantitative methodology for the analysis. Section 4.5.2 discusses the results from the RD measurement. First, the RD for the postverbal elements is tabulated to examine the degrees of activation of postverbal elements. Next, the analysis will be extended to the comparison between postverbal elements and the preceding elements in terms of degree of activation. Section 4.5.3 discusses cases in which the postverbal elements exhibit large RD, i.e. RD NPM, in light of the notion of activation, and finally Section 4.5.4 lays out the

measurement results in connection with the acceptability hierarchy for postposing proposed in Figure 2.

4.5.1. Quantitative Analysis Methodology

The analysis in this chapter is based on the same spoken Japanese data as used in the analysis of *wa* and *ga* in Chapter 3. In the RD measurement of this section, the *postposed elements* include adverbial subordinate clauses placed following the matrix clause as well as phrasal units placed following the predicate within the same clausal unit. The transcribed text was prepared by identifying the RD measurement units, i.e. clausal and often phrasal units, following the same procedure as in Section 3.2. The postposed elements were identified by the criterion of "pause"; namely, postverbal elements were considered as postposed elements if there is no intervening pause between the units, and if there is an intervening pause, postverbal elements were treated as an independent phrase or clause token and therefore not included as postposed elements.

In counting the number of clauses back to the preceding elements, the clause immediately preceding the postposed elements was ignored since the clause was considered as the utterance unit which the postposed elements belong to. The RD was measured for both preceding elements and postposed elements in the postposing constructions. In the case of measuring RD for propositions, the same method as used in Section 3.3.3 was applied; namely, the propositions in the preceding discourse context were considered as a preceding coreferential proposition only if they represent for the whole proposition expressed by the postposed elements in question.

Finally, the RD measurement for this section did not include the following postverbal elements: (i) postverbal adverbials modifying the utterance as a whole, which express the speaker's evaluative attitude toward the statement, such as *yappari* 'as expected', *ikuranandemo* 'at least', *hakkiriitte* 'frankly speaking', and *kitto* 'certainly', (ii) postverbal adverbials which serve as introduction of a statement, such as *tsumari* 'in other

words' and *sooieba* 'speaking of', and (iii) postverbal adverbials which serve to fill the gap in conversation, such as *kangaetemitara* 'thinking of' and *ima omoeba* 'thinking of (that) now'. The adverbials such as the above do not refer to a particular proposition in the previous discourse. Instead, they are uttered to signal the speaker's evaluation about the proposition expressed by the statement as in (i), or they are uttered as a part of the turn taking strategies, such as claiming the conversation turn as in (ii) and filling the gaps between turns as in (iii).[59] Since the adverbials of the above kind do not refer to a proposition in the previous discourse, the RD measurement was not applied to them; therefore, they are not included in the discussion in the following section.

4.5.2. Measurement Results

In order to present the results from the RD measurement, this section will discuss the RD of postverbal elements first and the token distribution in terms of RD of both the postverbal and preceding elements next.

Table 2 shows the token distribution of elements in the postverbal position. There are two overall patterns in the results: one pool of tokens falls on the range of small RD, i.e. RD 1 through 5, and the other on the range of large RD, i.e. RD 20+ and NPM. In terms of the total number of tokens, the former trend is more dominant than the latter. There are 57 tokens, 58% of the total, in the range of RD 1 through 5, while there are 30 tokens, 31% of the total, in the range of RD 20+ and NPM. The results here indicate that there is a fairly clear pattern that the postverbal elements are active; however, at the same time there is an opposite pattern that the postverbal elements may be nonactive.

Table 2: Token distribution of postverbal elements in terms of RD

RD of postverbal element	# of tokens	
1	34	(35%)
2-5	23	(24%)

6-10	5	(5%)
11-15	2	(2%)
16-19	4	(4%)
20+	10	(10%)
NPM	20	(20%)
Total	98	(100%)

In order to further examine the data, Table 3 presents the token distribution in terms of both the RD of postverbal elements and the RD of the propositions expressed by the preverbal elements, including the predicates. As illustrated in Table 3, there is an overall tendency that the RD of postverbal elements is smaller than that of the preceding elements of the clausal unit. Namely, there are 74 cases, i.e. 76% of the total of 98 tokens, in which the RD of postverbal elements is smaller than that of the preceding elements, which is indicated by the boxed cells in bold face in Table 3. This result indicates an overall trend that the entities in the postverbal position are more active than the proposition expressed by the preceding elements in the clause. Conversely, the opposite situation is rare; there are only two cases, 2% of the total, in which the RD of postverbal elements is greater than that of the preceding elements.

In spite of the general tendency that the postverbal elements are more active than the preceding elements, the results show a pattern which deviates from the overall tendency. Namely, there are 20 cases, i.e. 20% of the total, in which the postverbal entities and the preceding elements are equally nonactive, in which the RD of both postverbal elements and the preceding elements is RD 19-20 or NPM. The 20 cases in which the RD of the postverbal elements is NPM will be discussed in the next section to further investigate the activation status of the postverbal elements.

Table 3: Token distribution of postverbal elements in terms of their RD and RD of sentence minus postverbal elements

RD of sentence	RD of postverbal element									
minus postverbal	1	2-3	4-6	7-9	10-12	13-15	16-18	19-20	NPM	Total

element										
1	2	0	0	0	0	0	0	0	0	2
2-3	2	0	1	0	0	0	0	0	0	3
4-6	1	2	1	0	0	0	0	0	1	5
7-9	0	0	1	0	0	0	0	0	0	1
10-12	0	0	0	0	0	0	0	0	0	0
13-15	1	0	0	0	0	0	0	0	0	1
16-18	0	0	0	0	0	0	0	0	0	0
19-20	2	0	0	0	0	0	1	1	0	4
NPM	26	13	6	3	2	1	3	9	19	82
Total	34	15	9	3	2	1	4	10	20	98

4.5.3. Postverbal Elements of RD NPM

As discussed in the analysis of *wa* and *ga* on the basis of the RD measurement, it is not necessarily the case that RD directly represents the degree of activation of every element in the database. There is a limit to the measurement in that RD captures the degree of activation only if an element is activated by an identical element which is linguistically present in the preceding discourse context. However, as discussed earlier, an element becomes activated not only by the preceding co-referential expression but also by other factors as well, such as related concepts to the element, non-verbal expressions, etc. Given this limitation of the RD measurement, this section further examines the activation status of the 20 postverbal elements whose RD is NPM.

The 20 tokens of RD NPM can be classified into the three types as follows and the token distribution is shown in Table 4.

(53)

a. The postverbal elements represent information which is activated not by a co-referential expression in the preceding discourse but by a non-verbal cue in the discourse.

b. The postverbal elements represent information which is made accessible to activation by a related element in the preceding context.

c. None of the above

Table 4: Distribution of tokens of RD NPM in terms of the three types

Type	# of tokens	
(53a)	2	(10%)
(53b)	5	(25%)
(53c)	13	(65%)
Total	20	(100%)

(53a) indicates the situation in which an element becomes activated by non-linguistic cues, such as the speaker's or the hearer's gesture which draws attention to particular elements in discourse. (54) illustrates one of the two such cases in the database.

(54)

A1: migigawa ga okaasama ne
right-side mother FP
'(The one) on the right side is your mother.'

A2: sorede T-san de
then and
'Then, (next one is) Mr. T.'

A3: anata wa dokoni irassharu n desu
you where be:HON NOM COP
'Where are you?'

B: [pointing at a photograph]
atasi ichiban hidari ni koo kidotte hitori imasu kedo sita no hoo ni
I first left in this-way pose alone be but bottom GEN area at
'I am at the bottom by myself, posing at the left-most (position).'

In (54), they are looking at a photograph of B's family, and in (54B) the speaker explains where she is by pointing at the photograph. Although *sita no hoo* is the first mention in the discourse, it is situationally activated by B's nonverbal cue of pointing at the photograph.

(53b) represents another type of cases which the RD measurement does not capture. An element becomes accessible to activation by a related element in the preceding discourse context. (55) shows an example of this type.

(55)

1. hotondo ima made mitakotomonai yoona hitotachi ga ne yattekimasita
almost now until have-never-seen like people FP come:PST
'There came people whom I had never seen before.'

2. sono hitotachi ga M toka S toka T toka ne sooyuu hitotachi desu
that people and and and FP such people COP
'Those were people such as M, S, and T.'

3. sooyuu hitotachi ga ne kaita n desu saku o ne
such people FP write:PST NOM COP story OBJ FP
'Those people wrote stories.'

In the preceding context of (55), they have been talking about story-writers who were invited to the broadcasting cooperation which the speaker worked for. In the immediately preceding context of (55.3), *those people*, i.e. the story-writers, are active, being mentioned many times; therefore, *saku* 'story' in the postverbal position in (55.3) is made accessible to activation by the active related element *story-writers*.

Finally, (53c) indicate cases in which the postverbal elements are neither situationally activated nor made accessible to activation by a preceding related element. In other words, the postverbal elements in those cases are nonactive, and therefore, the postverbal elements and the preceding elements are equally nonactive. (56) shows an example of this type.

(56)
1. nihonzin no atasitachi nisitemireba tottemo fuziyuuna koto ga ookute ne
Japanese GEN we for very inconvenient thing many FP
'There were many inconvenient things for us Japanese.'

2. tatoeba dantairyokoo da kara sanzyuunin kurai no
for-example group-tour COP because 30-people about GEN
'For example, because (it was) a group tour of about 30 people.'

In the preceding context of (56), they have been talking about the speaker's trip to a foreign country and inconvenience which she had; however, the postverbal element *about 30 people* in (56.2) is the first mention in the discourse. (57) is another example of this type from the database.

(57)
1. ano sibai wa tochuu de kaettekuru toki arimasu kedo
uh play middle in come-back time be but
'(I) sometimes leave in the middle of a play.'

2. eega wa tochuu de kaetta koto nai desu ne
movie middle in leave:PST thing be:NEG COP FP

donnani tsumaranai to omottemo
however boring QT think
'(I) have never left in the middle of a movie no matter how boring (I) think it is.'

In the preceding context of (57), they have been talking about movies, and in (57) the speaker compares movies with plays. The proposition represented by the postverbal elements in (57.2) is the first mention in the discourse.

As in (56) and (57), all cases of (53c) in the database are the sentence-focus type; i.e. the whole sentence including the postverbal elements represents new information. Although it is not totally clear, due to lack of clear indication of the speaker's intention, it seems to be plausible to assume that the postposing construction of this type has resulted from placing the more *important* elements before the less *important* elements. In (56.2), for example, the speaker is explaining why the trip was inconvenient, and the discourse context would suggest that the fact that it was a group tour is more important than the fact that there were about 30 people in the tour in that the former serves better than the latter as the reason why the trip was inconvenient. (57.2) is probably a clearer case, due to a contrast between the two entities *play* and *movie*. In (57.2), the preceding elements can be viewed as more important than the postverbal elements in that the former brings a clear contrastive sense with (57.1); namely, given the proposition *I sometimes leave in the middle of a play*, the speaker needs the proposition *I have never left in the middle of a movie* in order to make a clear contrast, rather than the proposition *no matter how boring I think it is*.

4.5.4. Discussion

Given the quantitative analysis of the conversational Japanese data in the previous section, this section summarizes the results from the RD measurement in connection with the acceptability hierarchy for postposing proposed in Section 4.5.2.

The acceptability hierarchy, which is repeated in Figure 3, predicts that the postposing construction is most acceptable when the postverbal element is active, and if the postverbal element is nonactive, it is more acceptable with nonactive preceding elements than with active preceding elements.

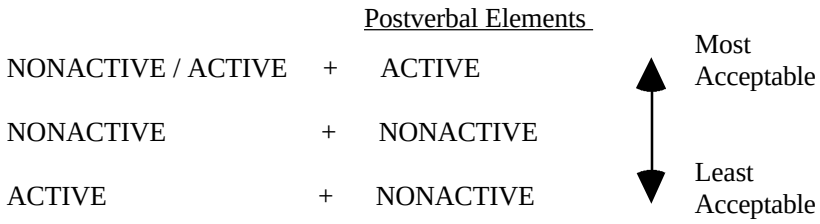


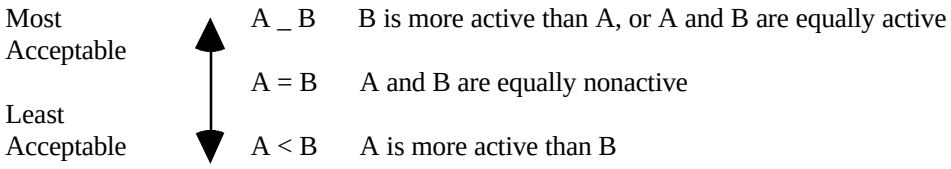
Figure 3: Acceptability hierarchy for postposing in terms of focus types

The results from the quantitative analysis of the database support the claims stated above in that the predicted patterns are reflected in the token distribution. Table 5 shows the correlation between the token distribution in Table 3 and the acceptability hierarchy for postposing in Figure 3. There is a fairly clear tendency that the majority of the tokens, i.e. 85% of the total, fall on the nonactive/active+active construction, which is the most acceptable type. The 83 tokens here include the seven cases in which the postverbal elements are accessible to activation or situationally activated, as discussed in Section 4.5.3. On the other hand, there are only two tokens, i.e. 2% of the total, for the active+nonactive construction, which is the least acceptable type. There are 13 tokens, i.e. 13% of the total, for the nonactive+nonactive type; there are more tokens for this type than the least acceptable type but less tokens than the most acceptable type.

Table 5: Token distribution in terms of the acceptability hierarchy

Acceptability hierarchy		# of tokens	
(most acceptable)	nonactive/active+active	83	(85%)
	nonactive+nonactive	13	(13%)
(least acceptable)	active+nonactive	2	(2%)
Total		98	(100%)

The empirical findings in the quantitative analysis suggest that we should define the three focus types for the hierarchy in terms of the relative degrees of activation, instead of the binary distinction active vs nonactive. In the nonactive+active type in Table 5, the postverbal elements range from RD 1 through RD 20+, and the preceding elements from RD 2 through RD NPM; however, the crucial point is that the postverbal elements are more active than the preceding elements in this focus type. Therefore, I modify the acceptability hierarchy for postposing as in Figure 4 so that it captures the relative degrees of activation of the postverbal and preceding elements.



A: proposition expressed by the preverbal elements, including the predicate
B: elements in the postverbal position

Figure 4: Acceptability hierarchy for postposing

4.5.5. *Wa* and *Ga* in Postposing Construction

Finally, I discuss the use of *wa* and *ga* in the postposing construction in terms of the activation patterns. Among the 98 cases of postposing examined in the previous sections, there are 23 tokens whose postverbal elements are marked by *wa*, including 16 *wa*-marked subjects and 7 *wa*-marked non-subjects, and 9 tokens whose postverbal elements are marked by *ga*. The dominance of *wa* in the postposing constructions would be plausible given the general principle that the postverbal elements are active. In order to further examine the use of *wa* and *ga* in postposing construction, Tables 6 and 7 show the token distribution of *wa*/*ga*-marked postverbal elements respectively in terms of the RD of postverbal elements and the preceding elements.

Table 6: Distribution of *wa*-marked postverbal elements and % of the total

RD of sentence minus postverbal element	RD of postverbal element		
	1-5	6-20	NPM
1-5	2 [9%]	0	0
6-20	0	0	0
NPM	18 [78%]	3 [13%]	0

Table 7: Distribution of *ga*-marked postverbal elements and % of the total

RD of sentence minus postverbal element	RD of postverbal element		
	1-5	6-20	NPM
1-5	5 [56%]	0	0
6-20	0	1 [11%]	0
NPM	2 [22%]	0	1 [11%]

Among the 23 postverbal elements marked by *wa*, there are 21 cases, i.e. 91% of the total, in which the postverbal elements are more active than the preceding elements. On the other hand, this pattern is not found among the *ga*-marked postverbal elements; a total of 56% of the tokens are the cases in which the postverbal and preceding elements are equally active, and there are only 22% of tokens for the case in which the postverbal elements are more active than the preceding elements, where the *wa* tokens cluster. Furthermore, the postverbal elements which are subjects, i.e. eligible for *ga* marking, in the range of RD 1 through 5 for both postverbal and preceding elements are typically marked by *ga*. The five tokens of *ga* for RD 1 through 5 come to 83% of the total of 6 tokens in the RD range.

The empirical finding above leads to the link between the focus types of postposing construction and the focus types for *wa* and *ga*. As discussed in Section 3.3.4, the focus types for *wa* and *ga* consist of four types, which I repeat as (58) below.

(58)

a. A > B (*wa*) — A is more active than B.

b. A < B (*ga*) — B is more active than A.

c. A = B = active — A and B are equally active.
ga

d. A = B = nonactive — A and B are equally nonactive.
ga

A: referent of noun phrase marked by *wa*/*ga*
B: open proposition with which A combines

For *wa* sentences, the *wa*-marked elements are more active than the rest of the sentence, and this coincides with the token distribution of the *wa*-marked postverbal elements in Table 6; namely, 91% of the total fall on the boxed cells in bold face in the table. For *ga* sentences, there are three possibilities in terms of the focus patterns: (i) the *ga*-marked elements and the predicates are equally active, (ii) they are equally nonactive, and (iii) the predicates are more active than the *ga*-marked elements. However, the last focus pattern of *ga* above is the least likely case for postverbal elements since it is the least acceptable type in the acceptability hierarchy in Figure 4. Accordingly, there is no *ga* token for this type, as shown in Table 7. The other two focus types of *ga* are the only possible types in the postposing construction; namely, the *ga*-marked postposed elements and the preceding elements, i.e. the predicates, are equally active or nonactive, which fall on the boxed cells in bold face in Table 7. There are 78% of the total tokens, i.e. 7 out of the total of 9, for those two types.

In summary, the acceptability hierarchy for postposing in Figure 4 and the four focus types of *wa*/*ga* sentences stated in (58) complement each other in order to predict the token distribution patterns in the database in the following way. Due to the four focus types of *wa* and *ga*, the overall token distribution patterns of *wa*-marked postverbal elements and the *ga*-marked postverbal elements do not overlap; namely, *wa* is dominant for postposing sentences whose *wa*/*ga*-marked elements is more active than the preceding elements, and *ga* is dominant for other focus types. On the other hand, the acceptability hierarchy for postposing rules out the focus pattern in which the predicate is more active

than the *ga*-marked element since the focus pattern goes against the general principle of postposing.

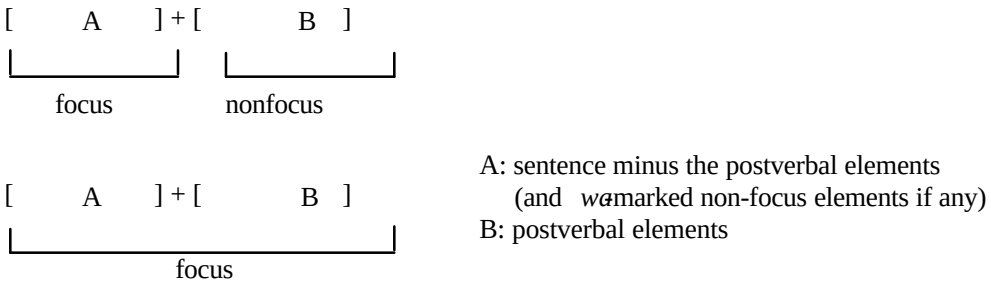


Figure 5: Focus structure of postposing construction

Figure 5 illustrates the focus structure of the postposing construction in Japanese. There are two focus types: one in which the postverbal unit is nonfocus and the preceding unit is focus and the other in which the whole sentence is focus. The former is exemplified by the cases in which the postverbal elements are active and the preceding elements are nonactive, or they are both active. In this focus type, the postverbal unit may be marked by either *wa* or *ga*; but the preceding elements are clearly more important than the postverbal elements. The second type in Figure 5 is exemplified by the cases in which the postverbal elements and the preceding elements are equally nonactive, and the postverbal unit is typically marked by *ga* if it is eligible for *ga* marking, i.e. it is the subject. In this focus type, although the whole sentence is within the focus, the preceding elements are more focused than the postverbal elements in that the former is more important than the latter. The cases in which the postverbal and preceding units are both active deserve further comments. They are different from those in which the postverbal and preceding units are both nonactive in the following ways. Against the background that the postverbal units may contain any syntactic category, whether it is phrasal or clausal, the first cases above typically include only phrasal units in the postverbal position, and in this type of postposing the postverbal and preceding units together make a new proposition by combining an active open proposition in the preceding unit and an active element in the

postverbal position, as suggested by the dominant use of *ga* in this type. (59B2) illustrates this case.

(59)

A1: nihon mo mono ga nakatta n desu ne
Japan also thing exist:PST NOM COP FP
'There was lack of materials in Japan too.'

A2: zyoorikuyooshuutee desu ka tsukutteta
boat-for-landing COP Q making:PST
'(The things called) landing boats? (You were making (them).'

B1: watakusi saigo ni sono sensoo no ano hoosoo ne ano tennooheeka no ano
I last at that war GEN that broadcast FP that emperor GEN that

hoosoo o iwayuru gyokuonhoosoo to yuu no o kiita toki wa
broadcast OBJ so-called QT say NOM OBJ hear:PST when

ano Sapporo koogai no beniyaitakoozyoo datte
that suburb GEN plywood-factory COP
'When I heard that broadcast by the emperor, which is so-called Gyokuon-hoosoo, (I) was in the plywood factory in the suburb of Sapporo.'

B2: zyoorikuyooshuutee no yoosuruni ano ita nan desu yo ne beniyaita ga
boat-for-landing GEN in-short that board NOM COP FP FP plywood
'That plywood was the board for the landing boats.'

In (B2) above, there are two active elements: *plywood* and the open proposition *X is board for the landing boats*, and the two are combined in (B2).[60]

Finally, if a postposing sentence contains a *wa*-marked unit in the preceding string, the *wa*-marked unit is not included in the focus since *wa* always marks nonfocus, as discussed in the previous chapters. Therefore, the presence of a *wa*-marked element preceding the focus gives us another focus pattern; namely, nonfocus is followed by focus, and the focus is in turn followed by postverbal nonfocus. This focus type is exemplified by a case in which a sentence contains a *wa*-marked, i.e. active, unit at the beginning as well as postverbal active elements, which is illustrated by (60B).

(60)

A: kinoo Ken to Hanako ga sensee no uchi ni itta tte

yesterday and teacher GEN home to go:PST QT
'(I heard) Ken and Hanako went to the professor's house.'

B: soo Ken wa nihon no uta o utatta n da yo sensee no uchi de
right Japan GEN song OBJ sing:PST NOM COP FP teacher GEN home at
'(That's) right, Ken sang a Japanese song at the professor's house.'

The sentence in (60B) has three focus components: *Ken wa, nihon no uta o uttatanda yo, sensee no uchi de*. The first and the last are active, being mentioned in (62A), while the second is nonactive. Therefore, the focus structure is represented by the sequence [nonfocus]*wa* + [focus] + [nonfocus].

4.6. CONCLUSION

In this chapter, I investigated the postposing construction in Japanese in the same framework as used in the analysis of *wa* and *ga* in the previous chapters, i.e. on the basis the theory of focus tied to activation: nonfocus is associated with active referents or propositions. Both the use of *wa* and *ga* and the use of postposing construction can be captured by the notions of activation and focus. Against the background of previous studies which often separate the issues of constraints on postposing and its discourse functions, the present study was an attempt to unify the two by proposing fundamental principles behind postposing. 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, or it may represent focus secondary to the focus involved in the preceding unit. The fundamental function of postposing which highlights the focus of the sentence accounts for the constraint on having nonfocus in the postverbal unit. Despite the difference: morphology and syntax, the fundamental function of postverbal units as nonfocus, or at most secondary focus, is analogous to the basic function of *wa* in that both postverbal units and *wa*-marked elements serve as a link between the preceding discourse and the primary focus of the sentence, i.e. they serve as a

basis for a new proposition by making the new proposition relevant at that point of discourse.

CHAPTER 5

WA AND *GA*, AND FOCUS STRUCTURE IN ROLE AND REFERENCE GRAMMAR

5.1. INTRODUCTION

The purpose of the present chapter is to apply the theory of information structure in RRG to the findings in the previous chapters regarding *wa* and *ga*. The analysis in this chapter utilizes the constituent structure in RRG and its theory of information structure (cf. Van Valin 1993a, 1993b). As Van Valin (1993a: 2) puts it, "RRG takes language to be a system of communicative social action, and accordingly, analyzing the communicative functions of grammatical structures plays a vital role in grammatical description and theory from this perspective." Given the emphasis on the communicative functions, RRG provides an excellent framework for describing the system of focus structure in Japanese on the basis of the findings in the previous chapters.

The analysis in this chapter will proceed as follows. Section 5.1 discusses the basic concepts in RRG and Section 5.2 extends the discussion to its theory of information structure along with evaluation of the theory as to the definitions of topic and focus and the notion of focus domain. Section 5.4 discusses a previous study of *wa* and *ga* in RRG and extends the analysis with a variety of cases of *wa*. The discussion is also given to lay out the focus structure of *wa*/*ga* sentences and to investigate the use of *wa* and *ga* in complex sentences in connection with a concept of information unit in subordinate clauses.

5.2. BASIC CONCEPTS IN RRG

Among major syntactic theories, RRG is characterized as a structural-functionalist theory of grammar (cf. Foley and Van Valin 1984, Van Valin 1993a). Unlike the Chomskyan view of language, RRG considers the communicative functions of grammatical structures as crucial; it claims that grammatical structure be captured in

connection with its semantic and pragmatic functions. Unlike a radical functionalist view of language, such as Hopper (1987), RRG is characterized by its claim that grammar is a system in the traditional structuralist sense.

Given the interplay of syntax, semantics and pragmatics in grammar, clause structure is represented in terms of four dimensions in RRG: constituent projection, operator projection, logical structure projection, and focus structure projection, as shown in Figure 1.

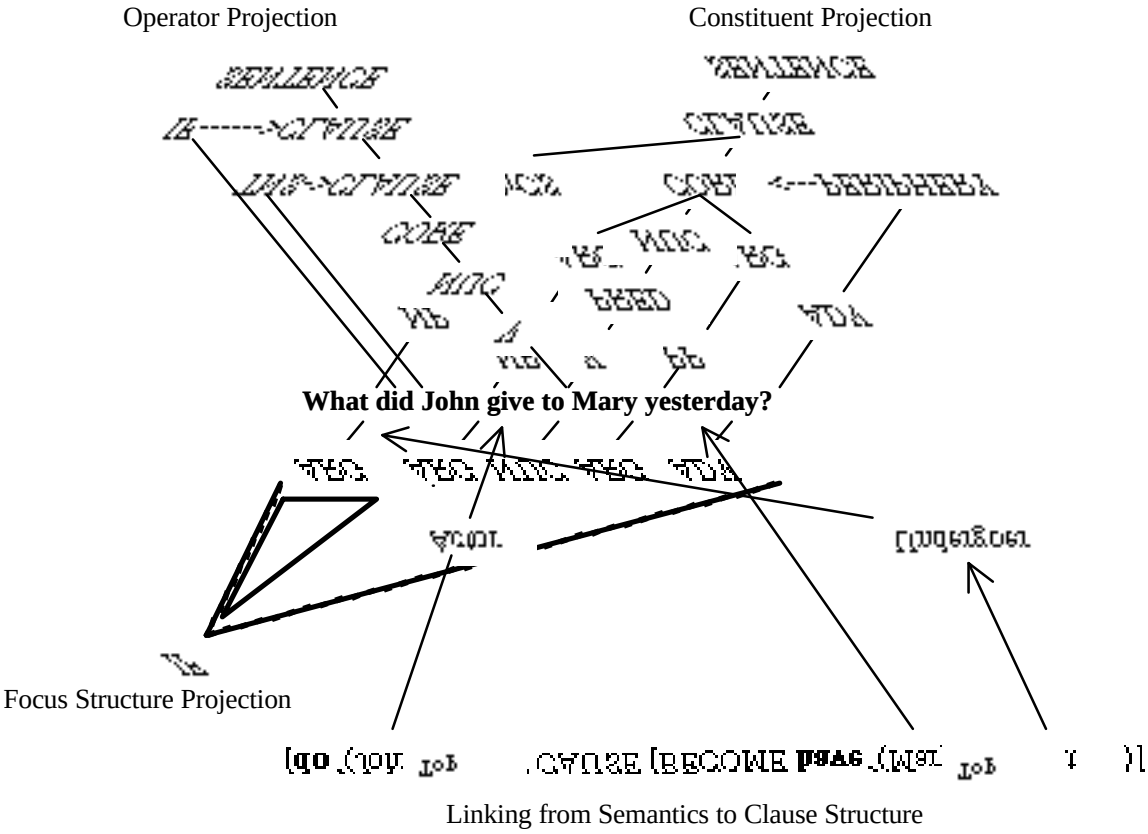


Figure 1: The four projections in RRG

The constituent projection represents the layered structure of the clause [LSC]. The LSC does not represent grammatical relations since they are not considered as basic nor universal in RRG, unlike Relational Grammar (cf. Perlmutter 1983). RRG posits only one level of syntactic representation and there are no syntactic rules similar to those in traditional transformational theories. The LSC defines sentence units semantically and is

considered to be universal, regardless of the syntactic characteristics of a given language, e.g. configurational or non-configurational, head-marking or dependent-marking. In the LSC, the primary constituent units are the *nucleus*. The nucleus contains the predicate, usually a verb. The next layer is the *core*, which contain the nucleus and the arguments of the predicate. The *periphery* is an adjunct to the core and contains non-arguments of the predicate.

There are two additional elements in LSC. The *precore slot* [PrCS] is within the clause but outside the core. The PrCS is the position for focused elements, e.g. question words in languages in which they do not occur *in situ*, or fronted elements as in an English sentence like *That guy I hate*. The left-detached position [LDP] is within the sentence but outside the clause. The LDP contains sentence initial elements, most commonly adverbials and topical elements, which are separate from the clause by a pause, as in *Tomorrow, I will go to the football game*. Figure 2 illustrates an English sentence containing all the LSC elements discussed above.[61]

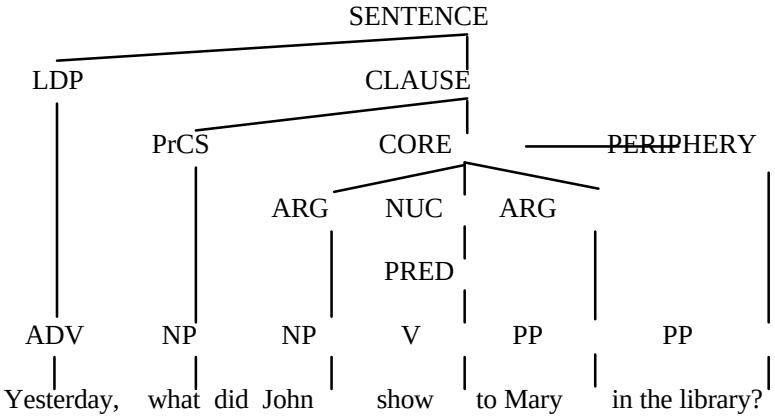


Figure 2: Layered structure of the clause in RRG

Morphologically-realized grammatical categories, such as aspect, tense, and modality, are not part of the LSC; they are represented in the operator projection. Those morphologically-realized grammatical categories are treated as *operators* which modify different layers of the clause. For example, the realization of tense in *did* in Figure 2 is

treated as a tense operator and it modifies the clause, which indicates a tense operator modifies not only the nuclear nor the core but the clause as a whole. Table 1 summarizes examples of operators found in a variety of languages and the layers which they modify.

Table 1: The layers of the clause and types of operators

Nuclear operators	Aspect Directionals (only those modifying orientation of action or event without reference to participants)
Core operators	Directionals (only those expressing the orientation or motion of one participant with reference to another participant or to the speaker) Modality (root modals, e.g. ability, permission, obligation) Internal (narrow scope) negation
Clausal operators	Status (epistemic modals, external negation) Tense Evidentials Illocutionary Force

The role of the lexicon plays an important role in RRG in that it employs a rich system of lexical representation, which is called *logical structure* [LS]. As shown in Figure 1, the LSs are linked to the syntactic representations according to semantic macroroles, i.e. *actor* and *undergoer*, determined in terms of the argument positions in the LS representations by the Actor-Undergoer Hierarchy.[62] The LSs are based on the scheme for lexical decomposition proposed in Dowty (1979), which modifies Vendler's (1967) classification of verbs: states, activities, achievements and accomplishments. LS is determined by the class membership of a verb, which is in turn determined by a set of syntactic and semantic tests proposed by Dowty (see Van Valin 1993a: 35). The set of distinctions among states, activities, achievements and accomplishments is considered to be one of the universal principles in human languages. (1) lists examples of verbs in English in terms of the class membership.

(1)

STATES	ACHIEVEMENTS	ACCOMPLISHMENTS	ACTIVITIES
have	receive	give	walk
know	learn	teach	talk
believe	realize	convince	think (about)
be dead	die	kill	watch

The LSs are presented formally in the schemata in Table 2. According to Dowty (1979), the stative predicates are primitive (they are in bold face in LS), and other classes

are derived from the stative predicates. Achievements are inchoative; therefore, they are represented as states plus a BECOME operator. Accomplishments are causative in nature; therefore, they have the structure in which an activity predicate and an achievement predicate are linked together by a CAUSE operator. (2) lists some examples of English verbs with their LS.

Table 2: Verb classes and logical structures

Verb Class	Logical Structure
STATE	**predicate**' (x) or (x,y)
ACHIEVEMENT	BECOME **predicate**' (x) or (x,y)
ACTIVITY (Agentive)	(DO (x)) **do**' (x, [**predicate**' (x) or (x,ι)])[63]
ACCOMPLISHMENT	ϕ CAUSE ψ, where ϕ normally an activity predicate and ψ an achievement predicate.

(2)

STATES

The watch is broken.	**broken**' (the watch)
The magazine is on the desk.	**be-on**' (the desk, the magazine)

ACHIEVEMENTS

The watch broke.	BECOME **broken**' (the watch)
Max arrived at the office.	BECOME **be-at**' (the office, Max)

ACTIVITIES

The ball rolled.	**do**' (the ball, [**roll**' (the ball)])
The door squeaks.	**do**' (the door, [**squeak**' (the door)])

ACCOMPLISHMENTS

Joan tossed the journal on the desk.	[**toss**' (Joan, the journal)] CAUSE [BECOME **be-on**' (the desk, the journal)]
Max ran to the office.	[**run**' (Max)] CAUSE [BECOME **be-at**' (the office, Max)]

As in simple sentences, the structure of complex sentences is captured in terms of the three units of LSC. The linkage of units is possible at any layer of the clause; therefore, there are three possible *juncture* levels: clausal, core, and nuclear. In terms of the relations between the units in complex sentences, RRG posits three *nexus* relations: *coordination*, *subordination*, and *cosubordination*. If the two units are independent of each other, the nexus type is coordination. If the one unit is dependent on the other, the nexus type is

either subordination or cosubordination. If the dependency is embedding, i.e. the linked unit is a modifier or an argument of the other unit, the nexus type is subordination. If the linked unit is dependent on the other only in terms of the operators of a particular juncture level, the nexus type is cosubordination. Figure 3 summarizes the three nexus types.

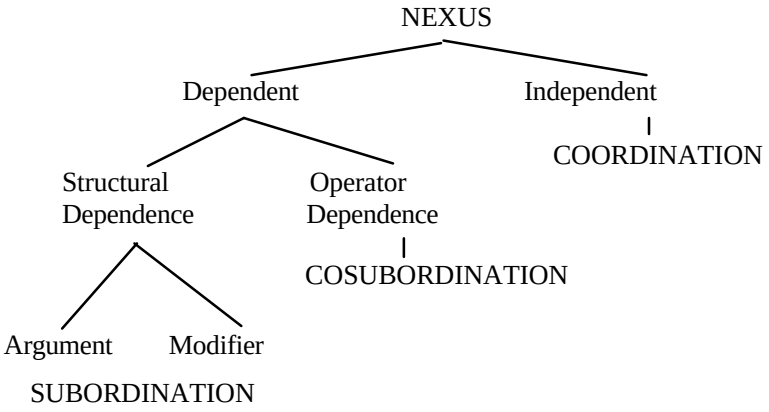


Figure 3: Features defining the nexus types (Van Valin and LaPolla, in press)

Given the three levels of juncture and three possible nexus relations, there are nine possible juncture-nexus types in universal grammar. All nine juncture-nexus types need not be realized in a particular language. English lacks nuclear coordination and subordination (cf. Van Valin and LaPolla, in press), and Japanese lacks nuclear cosubordination (Hasegawa 1992). The nine juncture-nexus types are ranked in terms of tightness of the syntactic bond involved in the linkage, nuclear cosubordination being the tightest and clausal coordination being the loosest. While these juncture-nexus types are purely syntactic, each of those linkage relations can express a certain range of semantic relations between the linked units. Although the relationship between the syntactic and semantic relations is not a clear-cut one-to-one mapping, the semantic relations can be ranked in terms of how closely related the propositions in the linkage are. Closely-related propositions in the linkage are coded as aspects of a single action to a greater extent than loosely-related propositions. The two hierarchies, i.e. syntactic relations and semantic relations, gives universal regularities that the closer the semantic relation between the

linked propositions is, the stronger the syntactic link involved in them. The two hierarchies are presented as the Interclausal Relations Hierarchy [IRH] in Figure 4.[64]

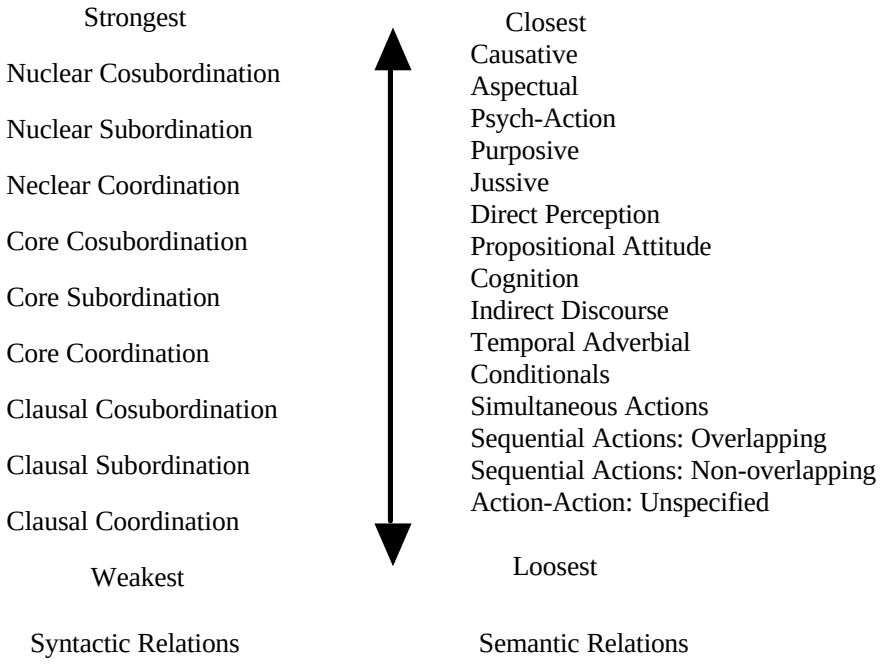


Figure 4: Interclausal Relations Hierarchy in RRG

5.3. INFORMATION STRUCTURE AND FOCUS DOMAIN

This section focuses particularly on the RRG theory of information structure since the primary purpose for this and the following chapters is to examine how the interplay between the focus structure and the system of grammar in Japanese is captured by the theory.

The theory of focus structure in RRG is based on Lambrecht's (1986, 1987, 1988, 1994) notions of topic and focus, which are defined as follows.

Topic:

> "A referent is interpreted as the topic of a proposition if in a given situation the proposition is constructed as being about this referent, i.e. as

expressing information which is relevant to and which increases the addressee's knowledge of this referent (Lambrecht 1994: 131)"

Focus:

"The semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition (Lambrecht 1994: 213)"

Although Lambrecht (1994) uses the notion of pragmatic presupposition for defining focus, he makes it in an ambiguous way by defining focus in terms of both shared beliefs and activated propositions, as he states "what counts for this notion [pragmatic presupposition] are not only the speaker's assumptions about the hearer's state of mind but also the speaker's assumptions about the hearer's assumptions about the speaker's state of mind (Lambrecht 1994: 60)."

The primary concern as to information distribution in RRG is the grammatical system which serves to indicate the scope of focus and non-focus in sentences, which is termed the focus structure by Lambrecht (1994: 222). The core of Lambrecht's theory is the taxonomy of focus types; he proposes three focus-structure categories: *predicate-focus structure*, *argument-focus structure*, and *sentence-focus structure*. The predicate-focus structure is the unmarked subject-predicate sentence type, where the predicate is the focus and the subject is in the presupposition. In the argument-focus structure, the focus identifies the missing argument in an open proposition which is in the presupposition. The sentence-focus structure is so-called event-reporting or presentational sentence, in which both the subject and the predicate are the focus. Lambrecht (1994: 223) illustrates the three focus-structure categories with the following examples.

(3) PREDICATE-FOCUS STRUCTURE

What happened to your car?

a.	My car/It broke DOWN.	English
b.	(La mia macchina) si ROTTA.	Italian
c.	(Ma voiture) elle est en PANNE.	French
d.	(Kuruma wa) KOSHOO-shi-ta.	Japanese

(4) ARGUMENT-FOCUS STRUCTURE

I heard your motorcycle broke down?
a. My CAR broke down.
b. Si rotta la mia MACCHINA./E la mia MACCHINA che si rotta.
c. C'est ma VOITURE qui est en panne.
d. KURUMA ga koshoo-shi-ta.

(5) SENTENCE-FOCUS STRUCTURE

What happened?
a. My CAR broke down.
b. Mi si rotta (ROTTA) la MACCHINA.
c. J'ai ma VOITURE qui est en PANNE.
d. KURUMA ga KOSHOO-shi-ta.

Different languages code focus and non-focus in different ways in terms of their syntactic structures, and therefore, information structure in RRG is concerned with the syntactic range in which the focus may occur in an utterance and the part of the range in which focus actually occurs in the utterance. The former is called the *potential focus domain* [PFD] and the latter the *actual focus domain* [AFD] (Van Valin 1993a: 29-30). In order to integrate focus structure into the projection grammar representation of clause structure, RRG posits a focus structure projection, which is separate from both constituent and operator projections but related to both. In terms of the constituent projection, the basic information units in focus structure (the minimal focus domains) are predicates, arguments, and peripheral PPs, i.e. the nucleus, a core argument or a peripheral PP, and in terms of the operator projection, the illocutionary force [IF] operator specifies the type of speech act which the sentence is in; so the PFD must be within the scope of the IF operator. Each focus structure projection represents both PFD and AFD within the scope of the IF operator, which is illustrated in Figure 1.

Given the theory of focus discussed in Chapter 1 and the analyses of focus structures involving *wa*/*ga* and postposing construction in the previous chapters, the present study will reject the assumption that non-focus is associated with pragmatic

presupposition. Since I utilize the notion of activation to define focus and non-focus, I modify the RRG definitions of focus for the purpose of the present study as follows.

Focus:

> A part of an utterance is focus if the speaker assumes that information expressed by the part is least active in the hearer's mind in the immediately preceding context of the utterance.[65]

The definition of focus above tells us that least active elements of utterance are always in the focus, and therefore within the focus domain of the utterance. On the other hand, domain of non-focus, i.e. outside the focus domain, always contains active elements of the utterance. The characterization of focus and non-focus above implies that active elements may be either focus or non-focus of the utterance. Therefore, there is implicational relationship between active/nonactive and focus/non-focus as in (6).

(6) If an element is nonactive, then it is focus of the sentence.
If an element is active, then it is either non-focus or focus of the sentence.

However, as discussed throughout this study, the degree of activation is not dichotomous. i.e. active and nonactive; rather, what determines the focus structure of the utterance is the relative degree of activation of elements. Therefore, the relations between the two tiers active/nonactive and focus/nonfocus is captured in the hierarchy in Figure 5.

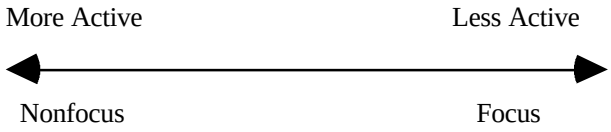


Figure 5: Focus-nonfocus hierarchy in terms of activation

The hierarchy tells us that an element which appears leftmost on the cline is nonfocus and an element which appears rightmost on it is focus. The degree of difference in activation among elements varies, depending on the preceding context. For example, a sentence with *wa* typically exhibits the focus pattern with two parts: one element which is active and the other which is nonactive. However, there are often cases in which all elements of a

sentence fall on the active side of the cline of the hierarchy above. There are also cases in which all elements of a sentence fall on the nonactive side of the cline. In those cases, the hierarchy predicts that among those elements, even more active or even less active elements will be nonfocus or focus of the sentence respectively.

Given the basic concepts of constituent structure in RRG and the theory of focus above, I will investigate how this basic assumption of focus-activation interface fits in the focus patterns manifested by sentences with *wa* and *ga* in the following sections.

5.4. *WA* AND *GA* IN RRG

Despite the long history of *wa*/*ga* analysis in Japanese linguistics, it is quite recent that linguists began analyzing the two particles within the RRG framework. To my knowledge, Hasegawa (1992) gives the pioneer analysis of *wa* and *ga* in RRG, which I will summarize below.

5.4.1. Hasegawa's Analysis

On the basis of Lambrecht (1988), Hasegawa (1992: 38-42) classifies focus patterns of *wa*/*ga* sentences in terms of the distinction between *predicate focus* and *narrow focus*, i.e. *argument focus* in Lambrecht (1994). Hasegawa cites the following to illustrate this point.

(7) Predicate Focus

a. kare wa/#ga sinda
he die:PST
'He died.'

b watasi wa/#ga tegami o kaita
I letter OBJ write:PST
'I wrote the/a letter.'

(8) Narrow Focus

a. kare GA/#wa sinda[66]
he die:PST
'HE died / It was he who died.'

b WATASI GA/#wa tegami o kaita
I letter OBJ write:PST
'I wrote the letter / It was I who wrote the letter.'

The examples in (7) shows that in the predicate focus construction, where the subject is anaphoric and the predicate serves as focus, the subject NP cannot be marked by *ga*. On the other hand, in (8), where the subject NP serves as focus and the predicate is the nonfocus, *wa* is not appropriate.

In addition to *wa* in predicate-focus, Hasegawa (1992) claims another function of *wa*, i.e. *wa* marking contrast, which follows the traditional definitions of *wa* (cf. Kuno 1972, 1973). The dual function of *wa* is illustrated by (9).

(9)
a. kare wa sinda (Topic)
he die:PST
'He died.'

b KARE WA sinda (Contrast: Narrow Focus)[67]
he die:PST
'HE died (but someone else survived).'

Furthermore, Hasegawa (1992: 40) cites two other points as follows in order to support the claim that contrastive *wa* is distinct from topic *wa*. First, *wa*-marked contrastive NPs can appear in a subordinate clause, while *wa*-marked topic NPs cannot, which is shown in (10).

(10)
a. Hiro ga/#wa kuru node shokuzi o yooi sita
come because meal OBJ preparation do:PST
'Because Hiro will come, (I) prepared a meal.'

b HIRO WA/#ga kuru node shokuzi o yooi sita
come because meal OBJ preparation do:PST
'(Although the others won't come,) Hiro will come, so (I) prepared a meal.'

Secondly, *wa* cannot mark a WH-phrase unless the *wa*-marked phrase is interpreted as contrastive, as illustrated in (11).

(11)

a. dare ga/#wa kimasu ka
who come Q
'Who will come?'

b DARE WA/#ga kimasu ka
come Q
'(I know that some won't come, but) who will come?'

In terms of the constituent projection of RRG, Hasegawa (1993) posits three different structures for the use of *wa* and *ga* discussed above. (12) through (14) are examples of a sentence with *ga*, both narrow focus NP and non-topical argument NP, a sentence with *wa*-marked topic NP, and a sentence with *wa*-marked contrastive NP respectively, and those three are represented in terms of the layered structure as in Figures 6 through 8 respectively (Hasegawa 1993: 41-42).

(12) ZYOON GA atama ga ii
head good
'JOAN (and only Joan in the current universe of discourse) is smart.'

(13) zyoon wa akarui (Topic NP-*wa*)
cheerful
'Joan is cheerful.'

(14) ZYOON WA kita (Contrastive NP-*wa*)
come:PST
'JOAN came (but someone else didn't).'

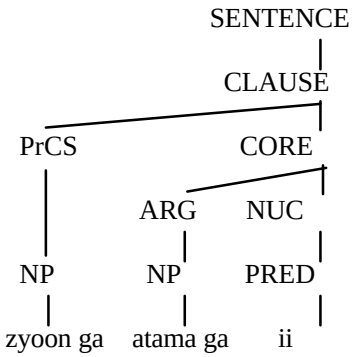


Figure 6: Narrow focus and non-topical argument NPs

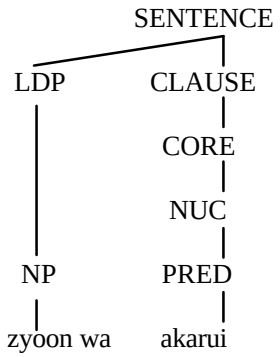


Figure 7: Topic NP-*wa*[68]

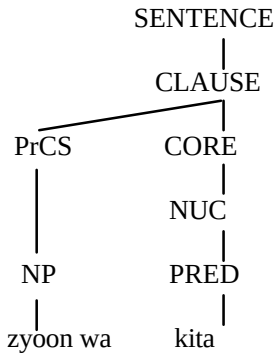


Figure 8: Contrastive NP-*wa*

The sentence in Figure 6 shows the argument focus structure with the leftmost NP as the focus, which is in PrCS. The second *ga* NP *atama ga* is the argument NP which is not the topic nor the focused NP as *zyoon ga* in PrCS. Therefore, there are two possibilities for *ga* NPs in terms of the layered structure; the *ga* NP may be either in PrCS if it is the focus as in argument focus construction or in ARG if it is not such focus of the sentence.

As for *wa* NPs, Hasegawa (1992) differentiates topic *wa* NP and contrastive *wa* NP by placing the former in LDP and the latter in PrCS. Regarding *wa*-marked contrastive NPs in PrCS, Hasegawa (1992: 42) states that "[c]ontrastive NP-wa's and narrow focus NP-ga's are cognitively similar: both convey the idea 'THIS entity, but not something else.' Therefore, it is plausible to associate a contrastive NP-*wa* with the PrCS, as in the case of a narrow focus NP-*ga*."

5.4.2. Scope of the IF Operator and Constituent Projection

As discussed in the previous chapters, the focus theory employed by the present study does not assure two functions of *wa*, topic and contrastive; the unmarked use of *wa*, i.e. the majority of the cases of *wa* in the conversation database, is to mark active elements in the discourse, whether the *wa*-marked NP exhibits contrastive sense or not. In addition to this empirical finding, the crucial evidence for this basic function of *wa* comes from the

following observation. Whether a *wa*-marked NP is contrastive or not, it is outside of the scope of the question, i.e. the IF operator.

(15)
A: Taro wa kaigi ni ikimasita ka
meeting to go:PST Q
'Did Taro go to the meeting?'

B: ikimasendesita
go:NEG:PST
'(No, he) didn't go.'

B': # Hanako ga/wa ikimasita
go:PST
'(No,) Hanako went.'

(15B) negates the predicate 'to go', while (15B') negates the *wa*-marked NP, which results in awkwardness, whether the NP in the answer is marked by *wa* or *ga*. The same pattern is found with *wa*-marked NP which has a clear contrastive sense.

(16)
A: Ken wa kaigi ni ikimasita kedo Taro wa ikimasita ka
meeting to go:PST but go:PST Q
'Ken went to the meeting, but did Taro go?'

B: ikimasendesita
go:NEG:PST
'(No, he) didn't go.'

B': # Hanako ga/wa ikimasita
go:PST
'(No,) Hanako went.'

If a sentence has more than one *wa*-marked NP, at least the second *wa*-marked NP is said to be contrastive (Kuno 1973: 48). In such sentences, the *wa*-marked NPs, whether it is the first or the second, are outside of the scope of the question. As shown in (17B') and (17B"), the answer cannot negate the direct object NP nor the subject NP.

(17)
A: Taro wa eego wa hanasimasu ka
speak Q
'Does Taro speak English?'

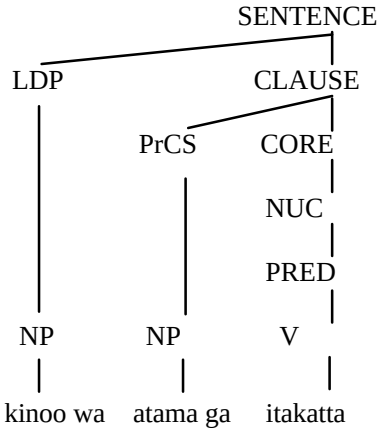
B: hanasimasen
speak:NEG:PST
'(No, he) doesn't speak.'

B': # doitsugo o/wa hanasimasu
German OBJ speak
'(No, he) speaks German.'

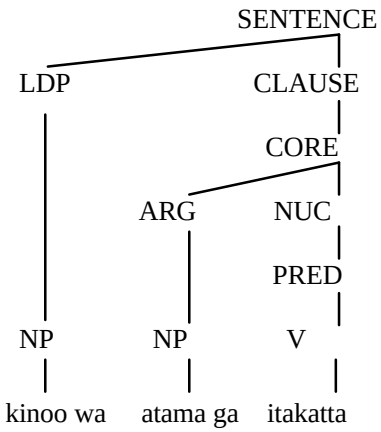
B": # Hanako ga/wa hanasimasu
speak
'(No,) Hanako speaks.'

Given the observation above, the present study posits only LDP for *wa*-marked NP and does not distinguish *wa*-marked topic NP and contrastive NP in terms of the constituent projection. Summarizing the discussion above, *wa*-marked NPs are in LDP, whether it is contrastive or topical, and regardless of the number of *wa*-marked NPs, while *ga*-marked NPs are either in PrCS if the NP is the focus of argument-focus or under ARG if the NP has no such focus; it is simply within the focus of sentence-focus. The sentence in (18) is ambiguous in terms of the focus assignment; therefore, there are two patterns in terms of the layered structure. Figures 9 and 10 show the constituent projection of the sentence with the *ga* NP in the PrCS and the sentence with the *ga* NP under the ARG respectively.

(18) kinoo wa atama ga itakatta
yesterday head pain:PST
'(I) had a headache yesterday.'



'(I) had a HEADACHE yesterday.'



'(I) HAD A HEADACHE yesterday.'

Figure 9 Figure 10

As discussed in Section 2.2, it is not only NPs but also almost any sentence elements that can be marked by *wa*. Below, I will examine those *wa*-marked non-NPs in terms of the scope of the question operator. (19) is an example of an indirect object marked by *ni* and *wa*, which shows the *wa*-marked indirect object is not in the scope of the question operator.

(19)
A: Taro wa Hanako ni wa sono hon o ageta no?
DAT that book OBJ give:PST Q
'Did Taro give the book to Hanako?'

B: agenakatta
give:NEG:PST
'(Taro) didn't give.'

B': # Ken ni ageta
DAT give:PST
'(Taro) gave (it) to Ken.'

The same pattern is found with *wa*-marked PPs.

(20)
A: Tokyo de wa sushi o tabeta no?
in OBJ eat:PST Q
'(Did you) eat sushi in Tokyo?'

B: tabenakatta
eat:NEG:PST
'(I) didn't eat.'

B': tempura o tabeta
OBJ eat:PST
'(I) ate tempura.'

B": # Osaka de tabeta
in eat:PST
'(I) ate (sushi) in Osaka.'

In (20), while the verb and the object NP can be negated by the answer, the *wa*-marked locative PP cannot.

Wa-marked clausal complements show the same pattern.

(21)
A: Ken ga nihon ni iku koto wa sitteta no?
Japan to go NOM knowing:PST FP
'Did (you) know that Ken was going to Japan?'

B: siranakatta
know:NEG:PST
'(I) didn't know.'

B': ? Hanako ga iku koto wa sitteta
go NOM knowing:PST
'(I) knew that Hanako was going (to Japan).'

In (21), while the answer can negate the verb, awkwardness results if a part of the complement clause is negated as in (B'), or at least it is not a cooperative answer to the question.

Another piece of evidence for the claim that *wa*-marked elements in the examples above are not in the scope of question comes from the incompatibility of WH phrase with *wa*. As discussed in Section 2.4.2, *wa* is in principle not compatible with WH phrase, and this is the case with the examples above.

(22) * dare wa kaigi ni ikimasita ka
who meeting to go:PST Q
'Who went to the meeting?'

This is also the case with the second *wa* NP, which is said to be contrastive (Kuno 1973).

(23) * Taro wa nani wa kaimasita ka
what buy:PST Q
'What does Taro buy?'

Furthermore, this principle applies to *wa*-marked indirect objects and other PPs.

(24) * Taro wa dare ni wa sono hon o ageta no?
who DAT that book OBJ give:PST Q
'Who did Taro give the book to?'

(25) * doko de wa sushi o tabeta no?
where in OBJ eat:PST Q
'Where did (you) eat sushi?'

As discussed in Section 2.4.2, there are cases in which *wa* is acceptable with WH phrases, i.e. *wa* is acceptable with WH phrases if there is an identifiable set of individuals and every member of the set is exhaustively represented in the WH-*wa* phrases (Miyagawa 1987: 188). For example:

(26)
A: kinoo Ken to Hanako to Taro ga kuru yotee datta kedo,
yesterday and and come plan COP:PST but

hutari sika konakatta
two only come:NEG:PST
'Ken, Hanako, and Taro were going to come yesterday, but only two came.'

B: dare ga/wa kite, dare ga/wa konakatta no?
who come who come:NEG:PST FP
'Who came, and who didn't?'

Those *wa*-marked WH phrases are within the scope of the question, i.e. the IF operator. As I discussed, however, the *wa*-marked WH phrases are the special cases in that those *wa*-marked WH phrases only partially satisfy the principle that *wa* marks active elements and combines with nonactive open propositions; the WH-*wa* phrases follow the principle in that they collectively represent an active set of individuals; however, at the same time it violates the principle since each *wa*-marked WH phrase represents only an unspecified subset of the active whole set, and also the open proposition which the *wa*-marked element fills is active. This partial violation of the principle accounts for the native speakers' judgment regarding the choice of *wa*/*ga* in contexts such as (26): some native speakers do not accept *wa*, while *ga* is invariably the acceptable choice.

The general principle that *wa* cannot mark WH phrases also applies to *wa*-marked clausal elements. It has long been pointed out (e.g. Muraki 1970, Inoue 1978) that

Japanese does not observe the complex NP constraint (Ross 1967); unlike English it is widely possible to form a WH question out of an embedded clause in Japanese.

(27) Ken wa [dare ga nihon ni iku koto] o sitteta no?
who Japan to go NOM OBJ knowing FP
'Who did Ken know that ___ was going to Japan?'

(28) Ken wa [dare ga katta] kuruma o untensita no?
who buy:PST car OBJ drive:PST FP
'Who did Ken drive a car which ___ bought?'

(29) Ken wa [dare ga okane o nusunda toiu] uwasa o sinziteru no?
who money OBJ steal:PST QT rumor OBJ believe FP
'Who does Ken believe a rumor that ___ stole the money?'

However, the sentences are unacceptable if *wa* marks the clausal complement or the complex NP which contains the WH phrase.

(27')* Ken wa [dare ga nihon ni iku koto] wa sitteta no?
who Japan to go NOM knowing FP
'Who did Ken know that ___ was going to Japan?'

(28')* Ken wa [dare ga katta] kuruma wa untensita no?
who buy:PST car drive:PST FP
'Who did Ken drive the car which ___ bought?'

(29')* Ken wa [dare ga okane o nusunda toiu] uwasa wa sinziteru no?
who money OBJ steal:PST QT rumor believe FP
'Who does Ken believe the rumor that ___ stole the money?'

Given the observation above, it is clear that it is not only WH phrases but also embedded clauses which contain a WH phrase that *wa* cannot mark.[69]

The tests in terms of question-answer pairs and the WH question formation above lead us to a claim that *wa*-marked NPs, PPs, and embedded clauses are outside the scope of the IF operator. However, there are cases in which *wa*-marked elements fall within the scope of the IF operator.

(30)
A: sonnani hayaku wa Toronto ni tsukenai n deshoo?
that fast in arrive:POT:NEG NOM COP:TNT
'(You) cannot arrive in Toronto that fast, can you?'

B: tsukemasu
arrive:POT
'(You) can arrive.'

B': yoru osokuni tsukemasu
night late arrive:POT
'(You) can arrive (there) late at night.'

B": # Toronto ni tsukemasu
in arrive:POT
'(You) can arrive in Toronto.'

In (30A), *wa* marks an adverb. The answer to (30A) can repeat the verb or the verb along with the information as to how long it takes to arrive in Toronto; however, the answer cannot simply repeat the verb and the PP as in (B"). This illustrates that the *wa*-marked ADV is actually the scope of the question.

Similarly, *wa*-marked predicates too are in the scope of the question. (31) is an example of the *wa*-marked copula.

(31)
A: Ken wa kyonen sensee de wa nakatta n deshoo?
last-year teacher COP NEG:PST NOM COP:TNT
'Ken was not a teacher last year, was he?'

B: gakusee desita
student COP:PST
'(He) was a student.'

B': # ototosi sensee de wa nakatta
the-year-before teacher COP NEG:PST
'(Ken) was not a teacher the year before.'

The answer to (31A) cannot replace *kyonen*, while it can replace the *wa*-marked predicate *sensee* with a different element, which indicates that the *wa*-marked predicate is the scope of the question. (32) shows the same pattern with a verb marked by *wa*.

(32)
A: baa ni itta kedo biiru o nomi wa sinakatta n deshoo?
bar to go:PST but beer OBJ drink do:NEG:PST NOM COP:TNT

'(You) went to a bar, but (you) didn't drink beer, did you?'

B: nonda yo
drink:PST FP
'(I) drank.'

B': # uiskii o nomi wa sinakatta
whisky OBJ drink do:NEG:PST
'(I) didn't drink whisky.'

(32) shows that the scope of the question is the *wa*-marked verb; the answer cannot replace the noun *biiru*, as in (B'), but it is perfectly acceptable to repeat the verb to give an affirmative answer, as in (B). The same pattern is found with the *te* form of verbs marked by *wa*.

(33)
A: Hanako ni denwasite wa mita n deshoo?
to call see:PST NOM COP:TNT
'(You) tried calling Hanako, didn't you?'

B: sita yo
do:PST FP
'(I) did.'

B': # Taro ni sita yo
to do:PST FP
'(I) called Taro.'

In (33), the repetition of the verb is perfectly appropriate to give an affirmative answer, while replacing elements other than the verb, e.g. *Hanako*, results in awkwardness. Again, this shows that the *wa*-marked verb is the scope of the question.

Furthermore, the same pattern is found with *wa*-marked adjectives.

(34)
A: sono mondai wa Hanako ni muzukasiku wa nakatta n deshoo?
the question for difficult NEG:PST NOM COP:TNT
'The question wasn't difficult for Hanako, was it?'

B: kantandatta yo
easy:PST FP
'(It) was easy.'

B': # Taro ni muzukasiku wa nakatta
for difficult NEG:PST
'(It) wasn't difficult for Taro.'

The awkwardness results if the answer replaces the element *Hanako*, while it is perfectly acceptable to replace the adjective, which indicates that the scope of the question in (34A) is the *wa*-marked adjective.

Interestingly, the observation above indicates that there are two extreme functions of *wa*: *wa* marking elements outside the scope of the IF operator and *wa* marking elements which is the scope of the IF operator. The *wa*-marked NPs, PPs and complements fall on the former, while *wa*-marked adverbials and predicates fall on the latter. In addition to the difference in the scope of questions, there is another kind of evidence for the two types of *wa*; *wa*-marked elements which are outside the scope of questions in general appear at the beginning of sentences, along with other *wa*-marked NPs if any, while *wa*-marked elements which are the scope of questions are not subject to such ordering restriction.

(35) *Wa*-marked subject

a. Taro wa kaigi ni ikimasita ka
meeting to go:PST Q
'Did Taro go to the meeting'

b. * kaigi ni Taro wa ikimasita ka[70]

(36) *Wa*-marked object

a. sono hon wa tomodachi kara moraimasita
that book friend from receive:PST
'(I) received the book from a friend.'

b. * tomodachi kara sono hon wa moraimasita

(37) *Wa*-marked indirect object

a. tomodachi ni wa sono hon o agemasita
DAT that book OBJ give:PST

'(I) gave the book to a friend.'

b. * sono hon o tomodachi ni wa agemasita

(38) *Wa*-marked PP

a. Tokyo de wa sushi o tabemasita
in OBJ eat:PST
'(I) ate sushi in Tokyo.'

b. * sushi o Tokyo de wa tabemasita

(39) *Wa*-marked clausal complement

a. Hanako wa [Ken ga nihon ni iku koto] wa tomodachi kara kiite sitteita
Japan to go NOM friend from hear knowing:PST
'Hanako knew that Ken is going to Japan since (she) heard (it) from a friend.'

b. ? Hanako wa tomodachi kara kiite [Ken ga nihon ni iku koto] wa sitteita

For the cases of *wa* above, the sentence is more acceptable if the *wa*-marked elements appear at the beginning of the sentence, along with the other *wa*-marked NPs if any.

On the other hand, the ordering restriction of preposing *wa*-phrases such as above does not apply to *wa*-marked ADVs, as follows. Furthermore, there are cases in which preposing ADV-*wa* phrases results in awkwardness, as in (41), which is exactly the opposite pattern to the *wa*-marked elements as seen above.

(40)
a. sonnani hayaku wa Toronto ni tsukenai
that fast in arrive:POT:NEG
'(You) cannot arrive in Toronto that fast.'

b. Toronto ni sonnani hayaku wa tsukenai

(41)
a. kono hikooki de sonnani takaku wa tobenai
this airplane by that high fly:POT:NEG
'(You) cannot fly that high by this airplane.'

b. ? sonnani takaku wa kono hikooki de tobenai[71]

In *wa*-marked predicates, *wa* connects elements within the CORE, and preposing the *wa*-marked element alone results in ungrammaticality. For example, with the copula:

(42)
a. Ken wa kyonen sensee de wa nakatta n deshoo?
last-year teacher COP NEG:PST NOM COP:TNT
'Ken was not a teacher last year, was he?'

b. * sensee de wa Ken wa kyonen nakatta n deshoo?

With a verb:

(43)
a. baa ni itta kedo biiru o nomi wa sinakatta n deshoo?
bar to go:PST but beer OBJ drink do:NEG:PST NOM COP:TNT
'(You) went to a bar, but (you) didn't drink beer, did you?'

b. * baa ni itta kedo nomi wa biiru o sinakatta n deshoo?

With an adjective:

(44)
a. sono mondai wa Hanako ni muzukasiku wa nakatta n deshoo?
the question for difficult NEG:PST NOM COP:TNT
'The question wasn't difficult for Hanako, was it?'

b. * sono mondai wa muzukasiku wa Hanako ni nakatta n deshoo?

Finally, *wa*-marked WH phrases are subject to the same ordering restriction. The contrast between (45B) and (45B') shows this point; the sentence is awkward with the preposed *wa*-marked WH phrase in (45B').

(45)
A: Ken to Hanako to Taro ga sensee no uchi no paatii ni iku tte itta kedo
and and teacher GEN house GEN party to go QT say:PST but

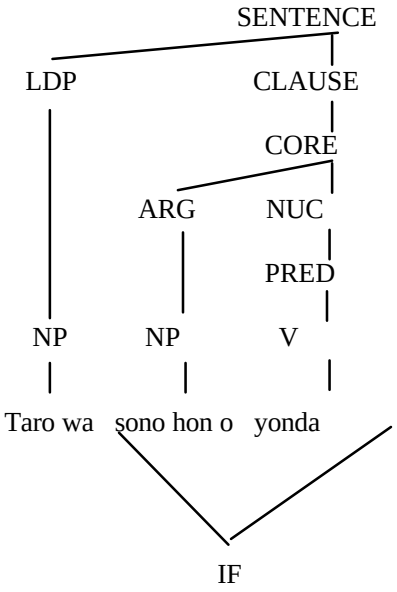
kekkyoku hutari sika konakatta tte
after-all two-people only come:NEG:PST QT
'Ken, Hanako, and Taro said that they would go to the party at the teacher's

house, but (I heard) only two people came after all.'

B: hontoo? sono paatii ni dare wa kite dare wa konakatta no?
really that party to who come who come:NEG:PST FP
'Really? Who came and who didn't come to the party?'

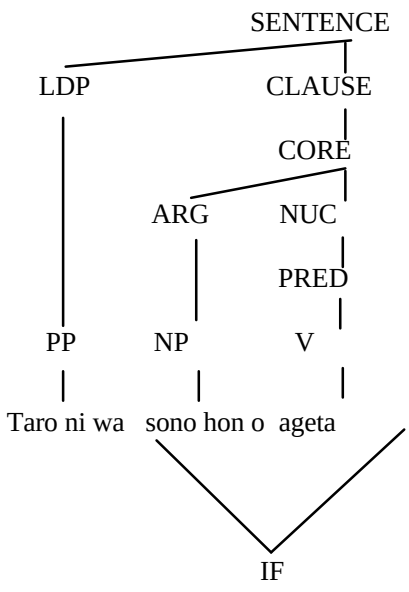
B': ? hontoo? dare wa sono paatii ni kite dare wa konakatta no?
really who that party to come who come:NEG:PST FP
'Really? Who came and who didn't come to the party?'

The two kinds of criteria discussed above, i.e. the scope of the question operator and the eligibility of preposing, suggest two types of *wa* in terms of the constituent projection: one type in the LDP, which is outside of the scope of the IF operator, and the other type within the CLAUSE, which is within the scope of the IF operator. The former type includes *wa*-marked NPs, PPs, and complement clauses, while the latter includes *wa*-marked ADV, predicates, and WH phrases. Figures 11-16 show the constituent projection along with the scope of the IF operator of an example for each of the six above.



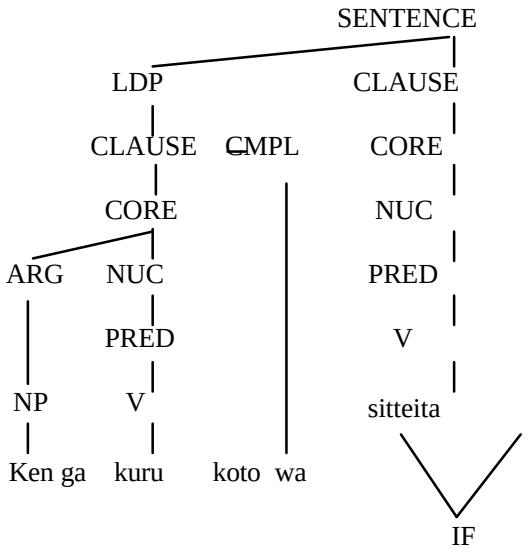
Taro wa sono hon o yonda
that book OBJ read:PST
'Taro read the book.'

Figure 11: *Wa*-marked NP



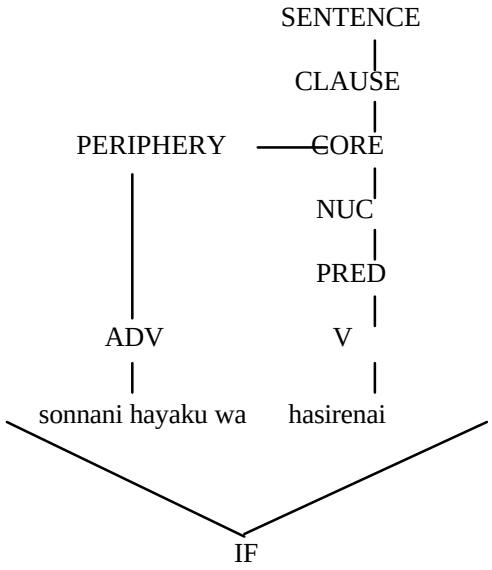
Taro ni wa sono hon o ageta
to that book OBJ give:PST
'(I) gave the book to Taro.'

Figure 12: *Wa*-marked PP



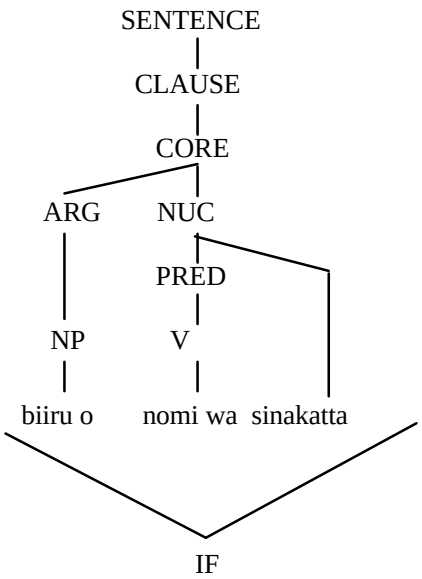
Ken ga kuru koto wa sitteita
come NOM knowing:PST
'(I) knew that Ken was coming.'

Figure 13: *Wa*-marked clausal complement



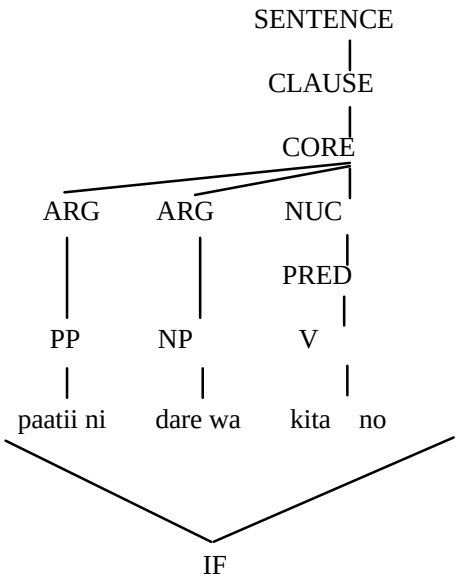
sonnani hayaku wa hasirenai
that fast run:POT:NEG
'(I) cannot run that fast.'

Figure 14: *Wa*-marked adverbial



biiru o nomi wa sinakatta[72]
beer OBJ drink do:NEG:PST
'(I) didn't drink beer.'

Figure 15: *Wa*-marked verb



paatii ni dare wa kita no?[73]
party to who come:PST FP
'Who came to the party (and who didn't)?

Figure 16: *Wa*-marked WH phrase

The claim made above that *wa* consists of the two types, one outside the scope of the IF operator and the other within the scope, is also supported by the empirical finding discussed in Section 3.5.2. The distribution of *wa* tokens in the conversation database shows that *wa*-marked predicates are typically high in RD, i.e. RD of 20+ and NPM. There are 11 cases in which *wa* marks the PRED (6 tokens marking the copula and 5 tokens regular verbs), and 9 out of the 11 tokens, i.e. 82% of the total, fall on the range of RD 20+ and NPM. This pattern contrasts strikingly with the token distribution of other types of *wa*, i.e. *wa*-marked NPs, PPs, and complement clauses, where *wa*-marked elements are in principle low in RD. I repeat the RD measurement results in Table 3 to summarize the discussion above.

Table 3: Token distribution of *wa* in terms of the structural types

Type	RD 1-19		RD 20+/NPM		Total	
Copula+*wa*	0		6	(100%)	6	(100%)
Verb+*wa*	2	(40%)	3	(60%)	5	(100%)
Other *wa* tokens	215	(77%)	66	(24%)	281	(100%)

5.4.3. Focus Domain

As discussed earlier, RRG posits the focus structure projection to indicate the domain of focus for particular sentences. On the basis of the analysis in the previous chapters and in the previous section, I will discuss the focus domain of *wa*/*ga* sentences below.

In Chapters 2 and 3, I made a claim that *wa* marks an element of nonfocus which combines with an open proposition of focus, while *ga* is used elsewhere. I repeat the schema of focus structure of *wa*/*ga* sentences below, in Figure 17.

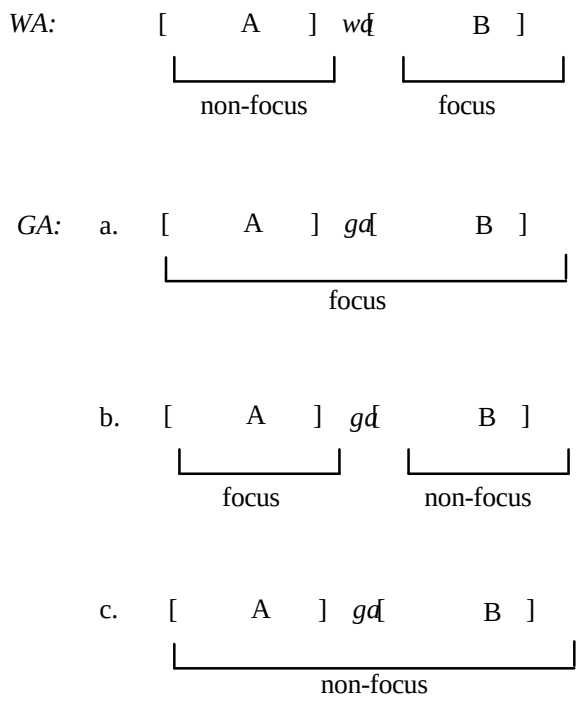


Figure 17: Focus structure of *wa*/*ga* clauses

Given the discussion in the previous section that all *wa*-marked elements are not nonfocus, the schema in Figure 17 needs to be modified as in Figure 18 so that it captures *wa*-marked predicates and adverbs within the focus of a sentence.

NONFOCUS	FOCUS
NP *wa*	PREDICATE *wa*
PP *wa*	ADV *wa*
CLAUSE *wa*	WH *wa*
NP *ga*	NP *ga*

Figure 18: *Wa*/*ga*-marked elements in terms of focus

In terms of the RRG representation, NP *wa*, PP *wa*, and CLAUSE *wa* appear in the LDP, which is outside the potential focus domain [PFD], while PREDICATE *wa*, ADV *wa*, and WH *wa* are within the actual focus domain [AFD]. *Ga*-marked NPs may appear either in the outside of the AFD but within the PFD or within the AFD. The *ga*-marked elements in sentence focus and argument focus, i.e. (a) and (b) in Figure 17

respectively, are within the AFD. On the other hand, the *ga*-marked elements in neutral focus, i.e. (c) in Figure 17, may be either within the AFD or outside the AFD but within the PFD. (46B) shows an example of neutral focus to illustrate this point.

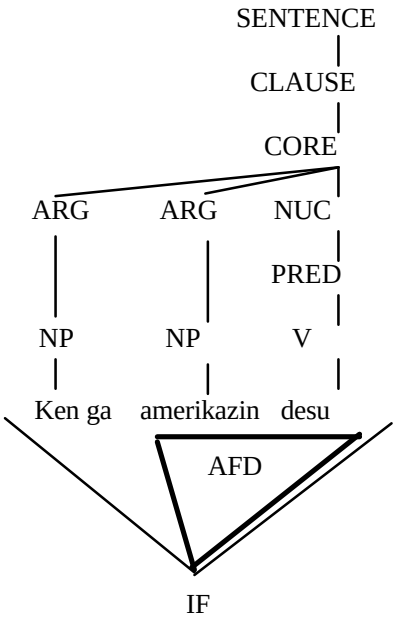
(46)

A: Sally ga amerikazin de Ken ga igirisuzin na no?
Sally NOM American and Ken NOM British NOM FP
'Is Sally American and Ken British?'

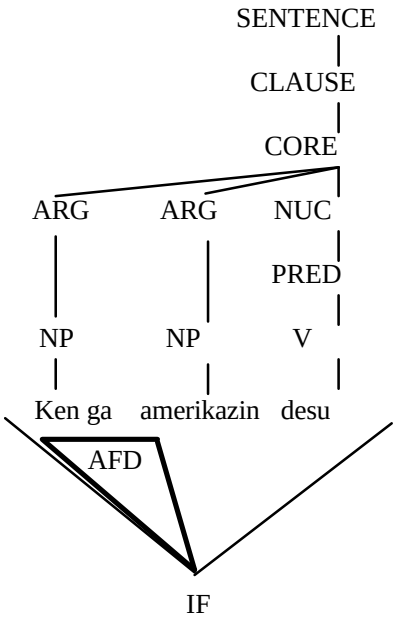
B: ie Sally ga igirisuzin de Ken ga amerikazin desu
no Sally NOM British and Ken NOM American COP
'No, Sally is British and Ken is American.'

In Section 2.4.1, I discussed that *wa* and *ga* are interchangeable in either clause in (46B); the choice of *wa*/*ga* is determined by the speaker's assumption as to the focused referent or proposition in the addressee's mind. *Wa* is used if the speaker assumes that *wa*/*ga*-marked referent is the center of attention in the addressee's mind, while *ga* is used if the speaker assumes that *wa*/*ga*-marked referent is not the center of attention, though it is active, in the addressee's mind.

The same type of principle applies to the use of *ga* in neutral focus, regarding the assignment of the AFD. In (46B), for example, the *ga*-marked element is within the AFD if the speaker assumes that the *ga*-marked element is less active than the open proposition in the addressee's mind, while the *ga*-marked element is outside the AFD (but within the PFD) if she assumes the *ga*-marked element is more active, but not active enough to mark it with *wa*, than the open proposition in the addressee's mind. Figures 19 and 20 illustrate the two patterns above respectively in terms of the focus structure projection. (The figures show only the second clause of (46B) to simplify the figure.)



'Ken is American.' (=46B)
('Ken' is more active than *X is American*.)
Figure 19



'Ken is American.' (=46B)
('Ken' is less active than *X is American*.)
Figure 20

The sentence in Figure 19 is analogous to sentences with a *wa*-marked non-focus element, e.g. *wa*-marked NPs, in that the *wa*/*ga*-marked elements are not the actual focus of the sentence; however, *wa*-marked elements of non-focus are always outside the PFD, i.e. the scope of the IF operator, while *ga*-marked NPs are always within the PFD. On the other hand, the sentence in Figure 20 is identical with *ga*-sentences of argument focus in terms of the focus structure projection since the AFD consists of the *ga*-marked element in both cases. However, there is a contrast in terms of activation that the *ga*-marked element in Figure 20 is active, though not the center of attention, while the *ga*-marked elements in argument focus are nonactive.

5.4.4. *Wa*/*ga* and Focus Structure in Complex Sentences

In terms of the PFD in simple sentences, Japanese is analogous to English in that they both have the option of the LDP, which is outside of the focus domain, and the rest of the sentence falls in the PFD. However, Japanese is differentiated from English in that

the range of PFD in complex sentences is much greater in the former than in the latter. In English, subordinate clauses are often outside the scope of the IF operator, which is suggested by the Complex NP constraint (Ross 1967) or the more general principle Subjacency (e.g. Chomsky 1973).

(47)

a. Ken believes that Hanako bought a car.

b. What does Ken believe that Hanako bought ___ ?

c. Ken believes the rumor that Hanako bought a car.

d. * What does Ken believe the rumor that Hanako bought ___ ?

In (47b), the WH word is the direct object of the verb in the subordinate clause, while in (47d), the WH word is the direct object of the verb in a clause which is a complement to a nominal head. The Complex NP constraint states that no element can move across more than one bounding node, i.e. NP and S in English, in a single movement.

A principle such as subjacency is impossible in RRG since it posits only one level of representation and have no transformation rules. Instead, RRG posits the following two general principles to account for the phenomena above (Van Valin 1993a).

(48) General restriction on question formation: The element questioned (the WH-word in a simple, direct WH-question or the focal NP in a simple, direct yes-no question) must function in a clause which is within the PFD of the sentence.[74]

(49) A subordinate clause may be within the PFD if and only if it is a direct daughter of (a direct daughter of) the clause node which is modified by the IF operator.

The principle in (49) differentiates the clausal complement and the noun complement in (47) in terms of the PFD as in Figures 21 and 22.

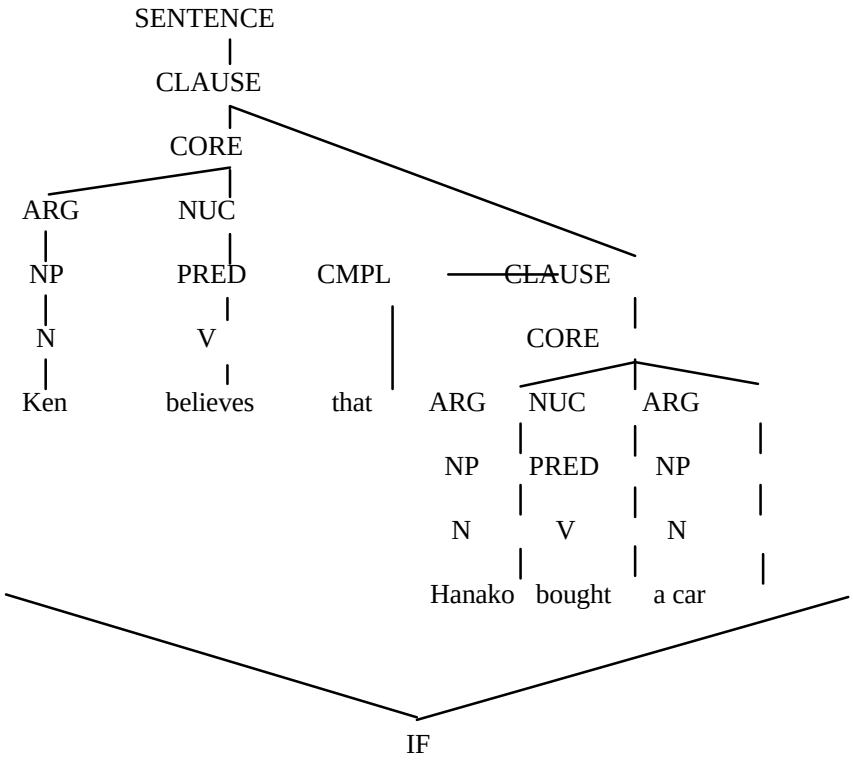


Figure 21: Clausal complement and the focus domain[75]

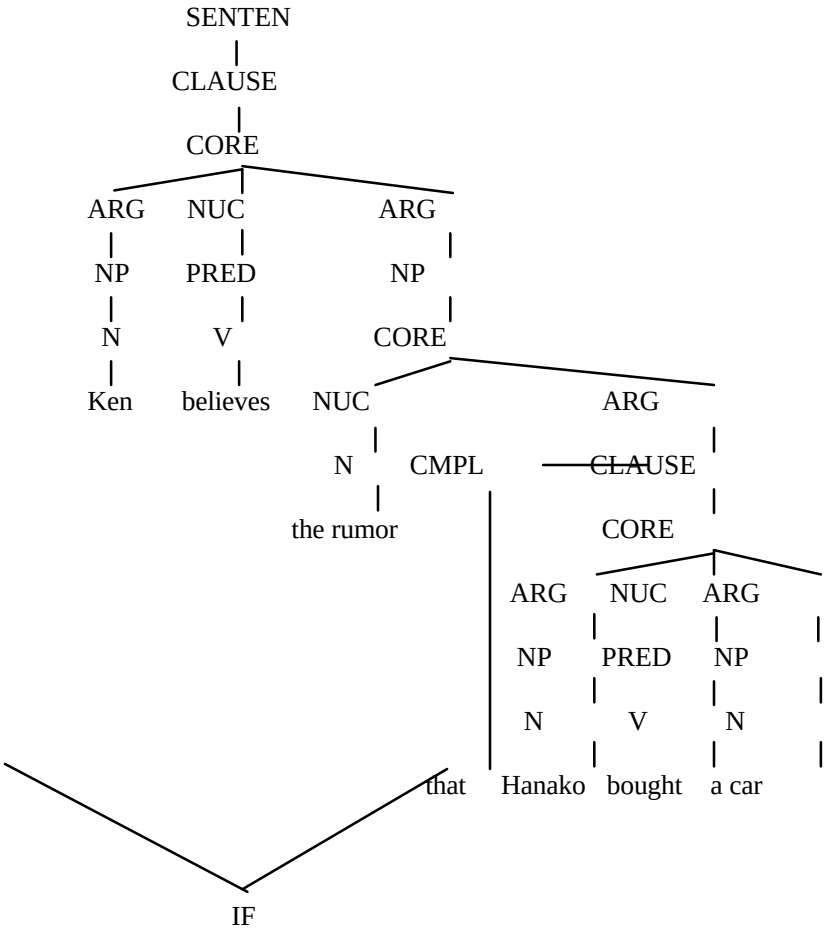


Figure 22: Noun complement and the focus domain

In Figure 21, the subordinate clause is a direct daughter of the clause which is directly dominated by the sentence node; the subordinate clause *Hanako bought a car* is within the PFD, while in Figure 22, the subordinate clause is the core argument of the NP, which in turn is the core argument of the matrix argument; therefore, the subordinate clause is outside the PFD. The difference in the range of the PFD accounts for the fact that the WH question formation out of the subordinate clause is allowed in the former, while it is not in the latter.

Unlike English, the structural principle in (49) does not apply to Japanese since Japanese allows WH question formation out of an embedded clause rather freely, as discussed earlier.

(50) Clausal complement

Ken wa [dare ga nihon ni iku koto] o sitteta no?
who Japan to go NOM OBJ knowing FP
'Who did Ken know that ___ was going to Japan?'

(51) Relative clause

Ken wa [dare ga katta] kuruma o untensita no?
who buy:PST car OBJ drive:PST FP
'Who did Ken drive a car which ___ bought?'

(52) Noun complement

Ken wa [dare ga okane o nusunda toiu] uwasa o sinziteru no?
who money OBJ steal:PST QT rumor OBJ believe FP
'Who does Ken believe a rumor that ___ stole the money?'

(53) Clausal adverbial

Hanako wa [dare ga kita toki] neteita no?
who come:PST when sleeping:PST FP
'Who was Hanako sleeping when ___ came?'

The examples above indicate that subordinate clauses in Japanese are within the PFD since a WH question can be formed out of those clauses, which strikingly contrasts with English, where there is considerable constraint in this regard. As Van Valin (1993a: 147) notes, languages vary regarding the acceptability of WH question formation out of subordinate clauses. Languages such as Russian and German are more strict than English in that they does not allow WH question formation out of any type of subordinate clause. English is analogous to languages such as Lakhota and Danish in that they allow WH question formation out of some subordinate clauses. Obviously, Japanese falls on the other end of this scale in terms of the principle in (49) since it in principle allows WH question out of any structural type of subordinate clauses.

Despite the difference among languages in terms of the principle in (49), however, languages do seem to be alike in terms of the correlation between the extractability, i.e. question formation, out of an embedded clause and the focus structure of the sentence. English, for example, follows the structural principle in (49), as discussed above; however, there are cases in English in which sentences that meet the principle but question formation results in ungrammaticality. Van Valin (1993a, b) observes the relative acceptability for WH-question formation in English, as shown by the contrast in the following.

(54)

a. What did Fred say that Mary had bought?

b. * What did Fred murmur/chortle/lisp that Mary had bought?

Despite the structural parallel of the two, whose structure meets the principle (49), there is a contrast between (54a) and (54b) in terms of the acceptability. Van Valin (1993b: 14) notes that the verbs in (54b) are informationally distinctive or rich, while the verb in (54a) is not, and "when the verb is informationally rich, it naturally draws the focus"; therefore, the focus on the verb blocks the extraction out of the clausal complement, which is the non-focus part of the sentence. Similarly, in Japanese the extraction out of a *wa*-marked clausal complement and an embedded clause within a *wa*-marked NP, i.e. nonfocus of the sentence, is impossible, as discussed in the previous section. Although English contrasts with Japanese in terms of the structural principle in (49), they both exhibit the correlation between the focus domain and the extractability, which is stated in (48).

As discussed above, in Japanese the focus domain includes embedded clauses unless the embedded clause or the head noun is marked with *wa*. However, this does not mean that all types of embedded clauses within the focus domain bear the same relationship with the matrix clause in terms of the information structure. At this point, it is useful to discuss Hasegawa's (1989) analysis of relative clauses containing a WH phrase and their possible answers.

Hasegawa (1989) claims that in relative clauses containing a WH phrase, what is questioned is necessarily what is to be identified in English, while the former is not necessarily the latter in Japanese. In essence, this claim is based on the following observation: (i) the WH question formation out of a relative clause is not allowed in English, while it is allowed in Japanese, and (ii) it is always possible in the reply to a WH question to supply only the value of the WH phrase in English, while it is not always possible to do so in the reply to a question with the WH phrase in the relative clause in Japanese. Hasegawa (1989: 142) cites the following example from Nishigauchi (1984).

(55)

A: dono kyoozyu ga suisensiteiru hito ga saiyoosare soo desu ka
which prof. recommending person employ:PAS likely COP Q
'(lit.) A person that which professor is recommending is likely to be employed?'

B: Suzuki kyoozyu ga suisensiteiru hito desu
prof. recommending person COP
'(It) is the person that Prof. Suzuki is recommending.'

B': # Suzuki kyoozyu desu
'(It) is Prof. Suzuki.'

In (55), the answer must repeat the entire NP including the subordinate clause, and supplying only the value of the WH phrase is inappropriate. Given the observation of sentences such as above, Nishigauchi (1984) claims that in answers to WH questions with relative clauses containing the WH phrase the truncated answer as in (B') is in principle not allowed in Japanese, and on the basis of this assumption he claims that, although there is no movement at S-structure in Japanese, the Subjacency constraint does exist at LF (which represents information relevant to the semantic interpretation of a sentence in question); moving the entire NP [$_{NP}$ [$_{S'}$ *dono kyoozyu ga suisen siteiru*] *hito*] is the only way to observe subjacency, assuming bounding nodes of NP and S'/S.

As Hasegawa (1989) points out, however, there are cases in which the truncated answers as in (55B') are acceptable for questions such as above.[76]

(56)

A: nani too o sizisuru hito ga ichiban ooi desu ka
which political-party OBJ support people most many COP Q
'(lit.) People who support which political party are most numerous?'

B: zimintoo o sizisiteiru hito desu
L-D Party OBJ supporting people COP
'(It) is the people who support the Liberal-Democratic Party.'

B': zimintoo desu
'(It) is the Liberal-Democratic Party.'

(57)
A: dono koohosha o sizisuru gakusee ga ooi desu ka
which candidate OBJ support student many COP Q
'(lit.) Students who support which candidate are numerous?'

B: Jeff Chang o sizisuru gakusee desu
OBJ support students COP
(It) is the students who support Jeff Chang.'

B': Jeff Chang desu
'(It) is Jeff Chang.'

Hasegawa (1989) accounts for the difference between (55) and (56)/(57) in terms of two distinct concepts: what is questioned vs what is to be identified; (55) is a case of *broad reading*, i.e. what is questioned is not what is to be identified, while the latter cases are *narrow reading*, i.e. what is questioned is what is to be identified. The two different readings are schematically shown in Figure 23.

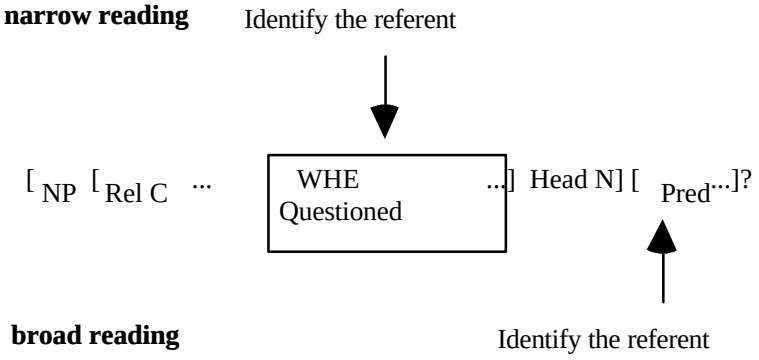


Figure 23: Two readings of relative clause construction (Hasegawa 1989)
(WHE = WH-expression)

In (55) what is questioned is the value of the WH phrase *dono kyoozyu*, while what is to be identified is the referent of the head noun *hito* which identification is made possible by the relative clause modifying the head noun; therefore, supplying only the value of the WH phrase for the answer is not acceptable. On the other hand, in (56) and (57), the referent of the WH phrase is both questioned and to be identified; therefore, it is acceptable, or even preferred, to give only the value of the WH phrase for the answer.

Hasegawa (1989) claims that the choice between the narrow reading and the broad reading of relative clauses containing a WH phrase is determined by context. More specifically, she uses the following two criteria: (i) the referentiality of the head noun, and (ii) the inherent topic-worthiness. The first criterion tells us that if the head noun is interpreted as referential, the broad reading is preferred. The head noun *hito* in (55) is intended to be referential since it refers to a particular person, while the head nouns in (56) and (57) are interpreted to be nonreferential, which is suggested by the use of the adjective *ooi* 'many'. Regarding the second criterion above, Hasegawa (1989: 147) states that "[c]ertain entities (e.g. humans) have inherently higher probability to be selected as a topic than others (e.g. animals) when contextual support is minimal. When the value of WHE [WH-expression] is high in topic-worthiness, and the sentence can be interpreted as expressing its property, a narrow-reading is possible even when the head N is referential." Hasegawa (1989) cites (58) to illustrate this point.

(58)

A: dono chookyoosi ga kunrensita inu ga yuushoosimasita ka
which trainer train:PST dog win:PST Q
'(lit.) A dog that which trainer trained won?'

B: Suzuki chookyoosi desu
trainer COP
'(It) is Trainer Suzuki.'

The head noun in (58A) is intended to be referential; therefore, (58B) should be inappropriate, according to the 'referability' criterion. However, (58B) is acceptable for at least some speakers, and Hasegawa notes that this is due to the two possible readings for (58A); the broad reading is possible since the head noun is intended to be referential, while the narrow reading is possible as well, because the value of the WH phrase can be high in topic-worthiness; the value is human and especially "producing prize-winning champions is an important property of trainers but not of "ordinary" people" (Hasegawa 1989: 47).

Although Hasegawa's (1989) analysis above is solely on the relative clause construction, the distinction between what is questioned and what is to be identified in WH questions is useful for investigation of the information structure in complex sentences in general. As seen above, what is questioned in WH questions is the value of the WH phrase, while what is to be identified is the core of the information unit in question. In the case of the narrow reading, what is questioned is what is to be identified; therefore the relative clause itself forms its own information unit. In the case of the broad reading, on the other hand, what is to be identified is the referent of the head noun and what is questioned is merely a part of the information which helps the questioner to identify the referent of the head noun; therefore, in this case the relative clause and the head noun together form the information unit. Given this assumption, we can extend the concepts of narrow reading and broad reading to the difference in *informational dependency* of the subordinate clause on the head noun, i.e. the matrix clause. If a sentence has a broad-reading relative clause, the relative clause is informationally dependent on the matrix clause; the relative clause is only supplementary to the matrix clause since the information conveyed by the relative clause is only to help the questioner to identify a part of the matrix clause. On the other hand, if a sentence has a narrow-reading relative clause, the relative clause is informationally independent of the matrix clause in that the relative clause itself contains the item which is to be identified by the question, though the relative clause is syntactically dependent on the matrix clause. This informational dependency suggests

the degree of *matrixhood* of the subordinate clause. In terms of the degree of matrixhood, relative clauses of narrow-reading are high; therefore, they are like matrix clauses informationally. On the other hand, relative clauses of broad-reading are low in the degree of matrixhood; therefore, they are subordinate to the matrix clause both syntactically and informationally.

Given the concept of matrixhood discussed above, I further examine complex sentences of Japanese by measuring the matrixhood of different types of subordinate clause. As discussed above, it is possible to form a WH question out of noun complement clauses. In an answer to the question, however, it is preferred to repeat the head noun along with the value of the WH phrase, as shown in (59).

(59)
A: Ken wa [dare ga okane o nusunda toiu] uwasa o sinziteru no?
who money OBJ steal:PST QT rumor OBJ believe FP
'Who does Ken believe the rumor that ___ stole the money?'

B: # Taro da yo
COP FP
'(It) is Taro.'

B': Taro ga nusunda toiu uwasa da yo
steal:PST QT rumor COP FP
'(It) is the rumor that Taro stole (the money).'

Obviously the broad reading is preferred in (59). Note that the value of the WH phrase is human. In spite of Hasegawa's (1989) criterion of inherent topic-worthiness, the broad reading is preferred here. On the other hand, the noun complement does exhibit the same tendency as in relative clauses in that if the head noun is non-referential, the truncated answer is acceptable.

(60)
A: [kaigairyokoo de nani o nusumareta toiu] todoke ga ichiban ooi desu ka
overseas-trip on what OBJ steal:PAS:PST QT report most many COP Q
'(lit.) Reports that what was stolen on the overseas trip are most numerous?'

B: saihu desu
wallet COP

'(It) is wallet.'

B': saihu o nusumareta toiu todoke desu
wallet OBJ steal:PAS:PST QT report COP
(It) is the report that wallets are stolen.'

Given the observation above, it is not very clear as to which of relative clauses and noun complement clauses exhibits greater matrixhood than the other, since with both simply giving the value of the WH phrase is not acceptable in some cases, while it is acceptable in the other cases. However, the noun-modifying clauses, including both relative clause and noun complement clause, show a striking contrast with clausal complement clauses in terms of the degree of matrixhood. With clausal complements, it is invariably not only acceptable but preferred to give only the value of the WH phrase in the answer.

(61)
A: Ken wa [dare ga okane o nusunda tte] sinziteru no?
who money OBJ steal:PST QT believe FP
'(lit.) Does Ken believe that who stole the money?'

B: Taro da yo
COP FP
'(It) is Taro.'

B': Taro ga nusunda tte sinziteru n da yo
steal:PST QT believe NOM COP FP
'(Ken) believes that Taro stole (the money).'

Although (61B) and (61B') are both acceptable, it is even the case that the former is preferred over the latter as the answer; (61B') sounds redundant, repeating the elements which are too obvious by the preceding context. The observation here suggests that clausal complements exhibit greater matrixhood than noun-modifying clauses.

The more crucial point is found in answers to WH questions as follows. In answers to a WH question containing a noun-modifying clause, it is not possible to give an elliptical answer without the head noun.

(62)
A: dono kyoozyu ga suisensiteiru hito ga saiyoosare soo desu ka

which prof. recommending person employ:PAS likely COP Q
'(lit.) A person that which professor is recommending is likely to be employed?'

B: # Suzuki kyoozyu ga suisensiteiru
prof. recommending
'(It is the person that) Prof. Suzuki is recommending.'

(63)
A: Ken wa [dare ga okane o nusunda toiu] uwasa o sinziteru no?
who money OBJ steal:PST QT rumor OBJ believe FP
'Who does Ken believe the rumor that ___ stole the money?'

B: # Taro ga nusunda toiu
steal:PST QT
'(It is the rumor that) Taro stole.'

In the case of clausal complement, it is best to give only the value of the WH phrase as in (61B); however, repeating the whole complement clause as in (64B) is not as bad as repeating a noun-modifying clause without the head noun as in (62B) and (63B).

(64)
A: Ken wa [dare ga okane o nusunda tte] sinziteru no?
who money OBJ steal:PST QT believe FP
'(lit.) Does Ken believe that who stole the money?'

B: Taro ga nusunda tte
steal:PST QT
'(Ken believes that) Taro stole.'

Furthermore, noun-modifying clauses contrast with clausal complements as to the acceptability in placing a matrix NP between the subordinate clause and the following matrix element. (65) and (66) show that the sentence becomes ungrammatical if the matrix subject intervenes the embedded clause and the head noun. On the other hand, the matrix subject can appear between the clausal complement and the matrix predicate, as shown in (67).

(65)
a. Hanako ga [Ken ga katta] hon o karita no?
buy:PST book OBJ borrow:PST FP
'Did Hanako borrowed the book which Ken bought?'

b. * [Ken ga katta] Hanako ga hon o karita no?

'Did Hanako borrowed the book which Ken bought?'

(66)
a. Ken ga [Taro ga okane o nusunda toiu] uwasa o sinziteru no?
money OBJ steal:PST QT rumor OBJ believe FP
'Does Ken believe the rumor that Taro stole the money?'

b. * [Taro ga okane o nusunda toiu] Ken ga uwasa o sinziteru no?
'Does Ken believe the rumor that Taro stole the money?'

(67)
a. Ken ga [dare ga okane o nusunda tte] sinziteru no?
who money OBJ steal:PST QT believe FP
'(lit.) Does Ken believe that who stole the money?'

b. [dare ga okane o nusunda tte] Ken ga sinziteru no?
'(lit.) Does Ken believe that who stole the money?'

Although WH question formation is possible out of both clausal complements and noun-modifying clauses, close examination in terms of the eligibility for an elliptical answer to a WH question and the intervenability of a matrix element between the embedded clause and the matrix element suggests that clausal complements exhibit greater matrixhood than noun-modifying clauses; the former is more informationally independent than the latter.

The different degrees of matrixhood is also found with different subordinate clauses of the same syntactic type. Here I examine adverbial subordinate clauses followed by *kara*/*node* 'because' and *ba*/*to*/*tara*/*nara* 'if'.[77] Although it is possible to form a WH question out of these subordinate clauses, they differ in terms of possible answers to the question.

(68)
A: Ken wa [dare ga iku nara] konsaato ni iku no?
who go if concert to go QT
'(lit.) Will Ken go to the concert if who goes (to the concert)?'

B: Hanako
'(It's) Hanako.'

B': ? Hanako ga iku nara
go if

'If Hanako goes, (Ken will go).'

B": [Hanako ga iku nara] iku tte
go if go QT
'(I heard) if Hanako goes, (Ken) will go.'

(69)
A: Ken wa [dare ga ike ba] konsaato ni iku no?
who go if concert to go QT
'(lit.) Will Ken go to the concert if who goes (to the concert)?'

B: Hanako
'(It's) Hanako.'

B': ? Hanako ga ike ba
go if
'If Hanako goes, (Ken will go).'

B": [Hanako ga ike ba] iku tte
go if go QT
'(I heard) if Hanako goes, (Ken) will go.'

(70)
A: Ken wa [dare ga i ttara] konsaato ni iku no?
who go if concert to go QT
'(lit.) Will Ken go to the concert if who goes (to the concert)?'

B: Hanako
'(It's) Hanako.'

B': ? Hanako ga i ttara
go if
'If Hanako goes, (Ken will go).'

B": [Hanako ga i ttara] iku tte
go if go QT
'(I heard) if Hanako goes, (Ken) will go.'

(71)
A: Ken wa [dare ga kuru to] isshoni terebi o miru no?
who come if together TV OBJ watch FP
'(lit.) Does Ken watch TV together if who comes?'

B: Hanako
'(It's) Hanako.'

B': ? Hanako ga kuru to
come if
'If Hanako comes, (Ken watches TV together).'

B": [Hanako ga kuru to] miru
come if watch
'If Hanako comes, (Ken) watches (TV together).'

In (68)-(71), it is possible to answer to the question either with only the value of the WH phrase or with the whole subordinate clause along with the matrix predicate. However, it is not acceptable to give an answer by repeating the subordinate clause without the matrix predicate. On the other hand, all three types of answers are possible with adverbial subordinate clause followed by *kara*/*node*.

(72)
A: Ken wa [dare ga iku kara/node] konsaato ni iku no?
who go because concert to go QT
'(lit.) Will Ken go to the concert because who goes (to the concert)?'

B: Hanako
'(It's) Hanako.'

B': Hanako ga iku kara/node
go because
'Because Hanako goes, (Ken will go).'

B": [Hanako ga iku kara/node] iku tte
go because go QT
'(I heard) because Hanako goes, (Ken) will go.'

The same contrast between 'if' clauses and 'because' clauses is shown by the acceptability in truncation of a sentence which negates the subordinate clause. For sentences with *kara*/*node*, it is perfectly acceptable to negate the subordinate clause without repeating the matrix predicate, as shown in (77). For sentences with *nara*/*ba*/*tara*/*to*, on the other hand, repeating the matrix predicate with the subordinate

clause is preferred over negating the subordinate clause without repeating the matrix predicate, as shown in (73)-(76).

(73)
A: Hanako wa [Ken ga kuru nara] konsaato ni ikanai tte?
come if concert to go:NEG QT
'Is Hanako not going to the concert if Ken comes?'

B: ? iya Taro ga kuru nara
no come if
'No, if Taro comes, (Hanako is not going to the concert).'

B': iya Taro ga kuru nara ikanai tte
no come if go:NEG QT
'No, if Taro comes, (Hanako) is not going (to the concert).'

(74)
A: Hanako wa [Ken ga kure ba] konsaato ni ikanai tte?
come if concert to go:NEG QT
'Is Hanako not going to the concert if Ken comes?'

B: ? iya Taro ga kure ba
no come if
'No, if Taro comes, (Hanako is not going to the concert).'

B': iya Taro ga kure ba ikanai tte
no come if go:NEG QT
'No, if Taro comes, (Hanako) is not going (to the concert).'

(75)
A: Hanako wa [Ken ga ki tara] konsaato ni ikanai tte?
come if concert to go:NEG QT
'Is Hanako not going to the concert if Ken comes?'

B: ? iya Taro ga ki tara
no come if
'No, if Taro comes, (Hanako is not going to the concert).'

B': iya Taro ga ki tara ikanai tte
no come if go:NEG QT
'No, if Taro comes, (Hanako) is not going (to the concert).'

(76)
A: Hanako wa [Ken ga kuru to] isshoni terebi o miru tte?

come if together TV OBJ watch QT
'Does Hanako watch TV together if Ken comes?'

B: ? iya Taro ga kuru to
no come if
'No, if Taro comes, (Hanako watches TV together).'

B': iya Taro ga kuru to miru
no come if watch
'No, if Taro comes, (Hanako) watches (TV together).'

(77)
A: Hanako wa [Ken ga kuru kara/node] konsaato ni ikanai tte?
come because concert to go:NEG QT
'Is Hanako not going to the concert because Ken will come?'

B: iya Taro ga kuru kara/node
no come because
'No, (Hanako is not going to the concert) because Taro will come.'

B': iya, Taro ga kuru kara/node ikanai tte
no come because go:NEG QT
'No, (Hanako) is not going (to the concert) because Taro will come.'

The observation above suggests that adverbial subordinate clause of *kara*/*node* 'because' exhibits greater matrixhood than that of *nara/ba/tara/to* 'if'; the former exhibits less informational dependency on the matrix clause than the latter. Summarizing the discussion above, clausal complements are less dependent on the matrix clause than noun-modifying clauses, including both noun complements and relative clauses, and 'if' clauses are less dependent on the matrix clause than 'because' clauses. Figure 24 illustrates this point.

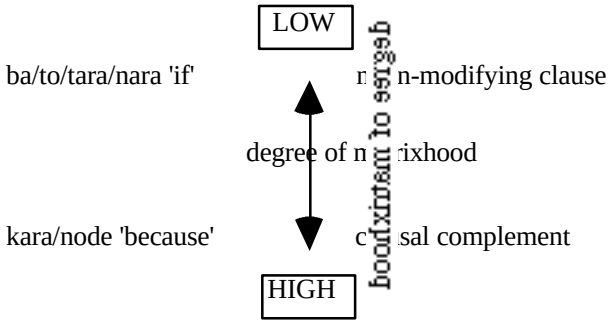


Figure 24: Matrixhood hierarchy of the four types of subordinate clause

In Section 3.4.5, I examined the use of *wa* and *ga* in subordinate clauses and discussed that the use of *wa*/*ga* correlates with some particular types of subordinate clause; there is more use of *ga* in the environment in which *wa* canonically appears in noun-modifying clauses than in clausal complements and in 'if' clauses than in 'because' clauses. I repeat in Figure 25 the hierarchy of the subordinate clause types in terms of *wa*/*ga* switch.

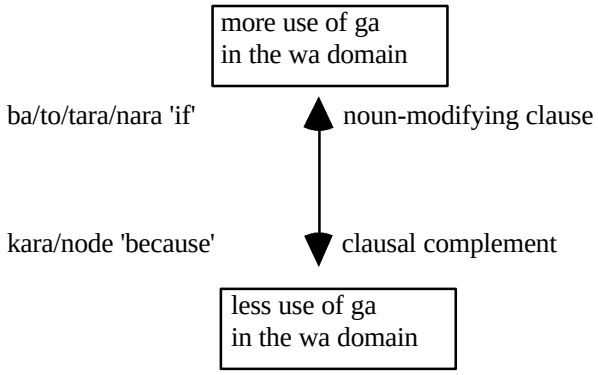


Figure 25: The use of *ga* in the *wa* domain and types of
embedded clause

As clear from Figures 24 and 25, there is a correlation between the degree of matrixhood and the degree in the substitution of *ga* for *wa*; the greater the matrixhood is, i.e. the more informational independence a subordinate clause exhibits, the more use of *wa* in its canonical environment there is within the subordinate clause. To illustrate this correlation, I combine the two hierarchies above in Figure 26.[78]

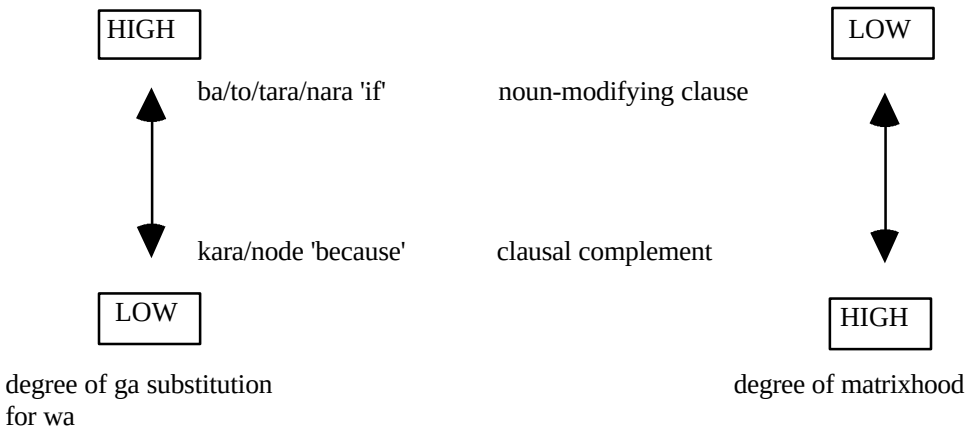


Figure 26: *Wa*/*ga*-switch and matrixhood hierarchy for the four subordinate clause types

Japanese is distinguished from languages such as English in terms of the range of focus domain, which is most clearly manifested by the possible range of WH question formation. In the case of Japanese, the potential focus domain covers the entire sentence, except for the elements in the LDP. However, the notion of focus domain alone does not provide a satisfactory accounts for the observation that *ga* is substituted for *wa* in some particular types of subordinate clause more frequently than in other types of subordinate clause. The close examination of different types of subordinate clauses in terms of some syntactic tests has revealed that there are different degrees of informational dependency on the matrix clause in subordinate clauses and *ga* is used in the place of *wa* in informationally dependent unit more frequently than in informationally independent unit.

In the following chapter, I will extend the RRG analysis to non-canonical word order and further investigate postposing constructions in Japanese, particularly on the basis of the theory of focus structure as well as the constituent structure.

CHAPTER 6

WORD ORDER FLEXIBILITY AND FOCUS STRUCTURE IN ROLE AND REFERENCE GRAMMAR

6.1. INTRODUCTION

In the previous chapter, I investigated *wa*/*ga* sentences within the framework of RRG, in terms of both its constituent projection and the focus structure projection. As seen in the discussion of the previous chapter, the use of *wa* and *ga* is fully explained only by the interplay between syntax and pragmatics; it cannot be accounted for satisfactorily by either structure alone or discourse-pragmatic factors alone. The purpose of this chapter is to extend the RRG analysis to non-canonical word order in Japanese in order to capture how much word order flexibility is allowed and how the constraints on non-canonical word order are accounted for by interplay between syntax and pragmatics. Furthermore, this chapter focuses particularly on the postposing construction in both simple and complex sentences to capture the principles which account for the constraints on postposing.

The discussion in this chapter will proceed as follows. Section 6.2 examines basic constraints on non-canonical word order in terms of both the constituent projection and the focus structure projection in order to see how sentences of non-canonical word order are represented in RRG. Section 6.3 extends the analysis particularly to the postposing construction to further discuss how the postposing construction is represented in both the constituent and focus structure projections and how the basic constraints on postposing are accounted for. Finally, Section 6.4 focuses on constraints on postposing of and out of a subordinate clause and proposes two different types of principles, one from the structural point of view and the other from the discourse-pragmatic point of view.

6.2. WORD ORDER FLEXIBILITY AND THE LAYERED STRUCTURE IN RRG

As discussed in Section 4.2, Japanese exhibits flexible word order to a considerable extent. This section examines the word order flexibility in terms of the constituent projection of RRG to see how the flexibility is represented in the theory.

As noted earlier, there are two basic structural constraints on non-canonical word order in Japanese: (i) dependent morphemes, including bound morphemes, particles, and the copula, cannot appear in a non-canonical position by themselves, and (ii) elements in a non-canonical position must be a *maximal projection* in the Government and Binding terminology. As to the first constraint, dependent morphemes are by definition those which cannot appear independently; for example, particles such as *wa*, *ga*, and *o* cannot be separated from the 'head' NP. In the formalist framework, the second constraint is accounted for in terms of the X-bar theory of clause structure, as discussed in Section 4.3.1. In postposing, for example, only a maximal projection can be postposed; NP can be postposed but not N alone, S' but not S alone, AP but not A alone, AdvP but not Adv alone, etc. I repeat the examples from Simon (1989) below.

(1) [$_{NP1}$ ___ [$_{N}$ shoosetsu]] o yonda no [$_{NP2}$ <u>furansugo no</u>]
novel OBJ read:PST FP French GEN
'(I) read a French novel.'

(2) * [$_{NP1}$ [$_{AP}$ mizikai] [$_{NP2}$ furansugo no] ___] yonda no [$_{N}$ <u>shoosetsu] o</u>
short French GEN read:PST FP novel OBJ
'(I) read a short French novel.'

(1) is grammatical because the whole genitive NP is postposed, while (2) is not acceptable because it is the head noun alone that is postposed. As to preverbal non-canonical word order, the constraint is more tight. For example, it is not possible to scramble NP2 out of NP1 in (3).

(3) * Ken ga <u>furansugo no</u> Hanako ni [$_{NP1}$ [$_{NP2}$ ___] [$_{N}$ shoosetsu]] o miseta
French GEN DAT novel OBJ show:PST

'Ken showed a French novel to Hanako.'

Since RRG does not posit X-bar type representations, the principle of maximal projection is impossible in this theory. In RRG, as in the case of clauses, NPs are represented in terms of both the layered structure and operator projection, which illustrates the striking parallels between the structures of NPs and clauses. The layered structure of the NP [LSNP] contains a nominal nucleus [NUC_N] as its most basic, which dominates a referring element [REF], i.e. a noun [N]. If the NUC_N dominates a relational noun, the nominal core [$CORE_N$] also dominates an argument in a PP headed by *of*, a non-predicative preposition.[79] Figure 1 is from Van Valin and LaPolla (in press: 26) to illustrate the LSNP in English.

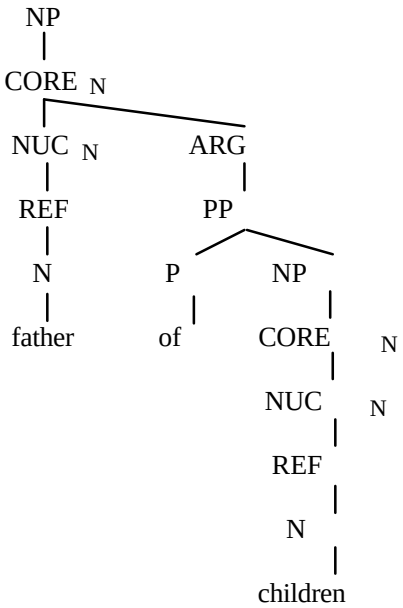


Figure 1: LSNP in English

The parallels between the structures of NPs and clauses are further illustrated by NPs headed by deverbal nominals, in which there is a core-periphery distinction. For example, the NP *arrest of Bill by FBI agents in New York* corresponds to the clause *Bill was arrested by FBI agents in New York*; the NP reflects the argument structure of the

source verb *arrest*, with *Bill* and *FBI agent* as core arguments and *in New York* as the periphery. Figure 2 is from Van Valin and LaPolla (in press: 27) to illustrate the LSNP.

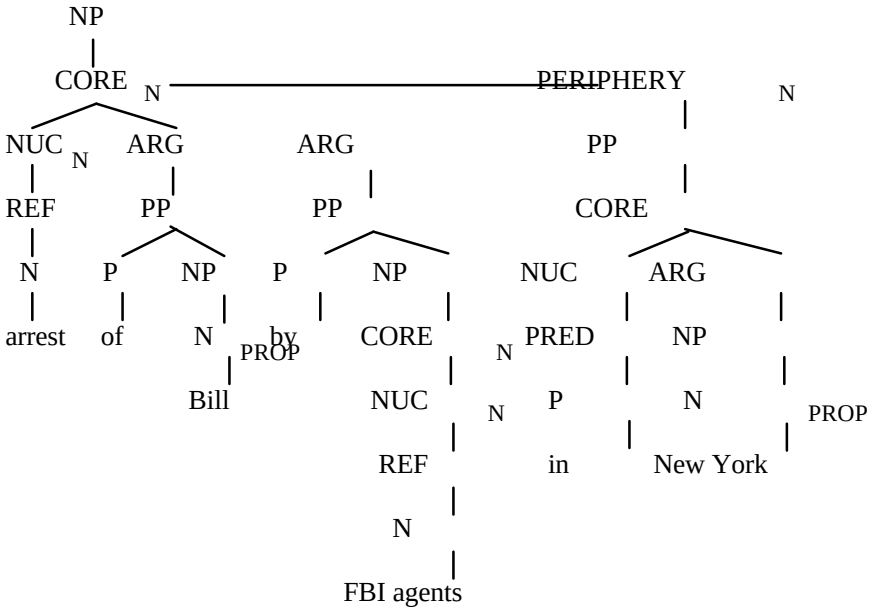


Figure 2: LSNP of English NP headed by deverbal nominal

The operator projection represents NP operators, including determiners, i.e. articles, demonstratives, and deictics, quantifiers, negation, and adjectival and nominal modifiers. Adjectival, nominal modifiers, and numeral classifiers are nuclear operators, the quantity operators modify the core of the NP, and the locality operators, i.e. deictics and definiteness, modify the NP as a whole. Van Valin and LaPolla (in press: 29) cite the examples in Figure 3 to illustrate the operator projection of NPs.

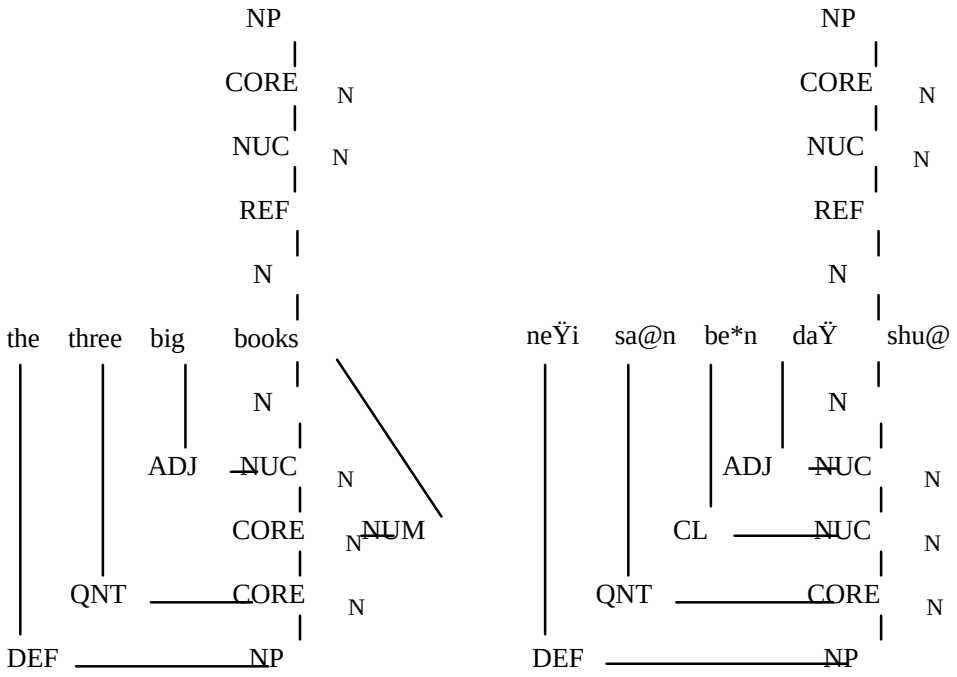


Figure 3: LSNP with operators in English and Mandarin Chinese
*nefi sa@n be*n daf shu@* [that three CL big books]

Given the LSNP with the operator projection of RRG discussed above, the second basic constraint on non-canonical word order is accounted for in the following way. In the case of postposing, a nominal nucleus alone cannot appear in the postverbal position, being separated from arguments and NP operators.

(4)

a. [$_{DEF}$ ___ [$_{ADJ}$ kireena [$_{NUC}$ onnanohito]]] o mita yo <u>ano</u>
beautiful woman OBJ see:PST FP that
'(I) saw that beautiful woman.'

b. [$_{DEF}$ ano [$_{ADJ}$ ___ [$_{NUC}$ onnanohito]]] o mita yo <u>kireena</u>

c. [$_{DEF}$ ___ [$_{ADJ}$ ___ [$_{NUC}$ onnanohito]]] o mita yo <u>ano kireena</u>

d. [$_{DEF}$ ___ [$_{ADJ}$ ___ [$_{NUC}$ ___]]] mita yo <u>ano kireena onnanohito o</u>

e. * [$_{DEF}$ ano [$_{ADJ}$ kireena [$_{NUC}$ ___]]] mita yo <u>onnanohito o</u>

(4) shows that any NP operator or any combination of NP operators can appear in the postverbal position without the nominal nucleus, while the nominal nucleus alone cannot appear in the postverbal position, leaving the operators in the canonical position.

The same pattern is found with ARGs within a NP.

(5)

a. [$_{NP}$ [$_{ARG}$ ___] [$_{ARG}$ sono hannin no] [$_{NUC}$ taiho]] wa sugokatta yo keesatsu no
that criminal GEN arrest violent:PST FP police GEN
'The arrest of that criminal by the police was violent.'

b. [$_{NP}$ [$_{ARG}$ ___] [$_{ARG}$ ___] [$_{NUC}$ taiho]] wa sugokatta yo sono hannin no

c. [$_{NP}$ [$_{ARG}$ ___] [$_{ARG}$ ___] [$_{NUC}$ taiho]] wa sugokatta yo
keesatsu no sono hannin no

d. [$_{NP}$ [$_{ARG}$ ___] [$_{ARG}$ ___] [$_{NUC}$ ___]] sugokatta yo
keesatsu no sono hannin no taiho wa

e. * [$_{NP}$ [$_{ARG}$ keesatsu no] [$_{ARG}$ sono hannin no] [$_{NUC}$ ___]] sugokatta yo taiho wa

The constraint on scrambling is tighter than that on postposing; clausal arguments cannot be split up.

(6)

a. [$_{ARG}$ Hanako] ni [$_{ARG}$ ookii eego no zisho] o ageta
DAT big English GEN dic. OBJ give:PST
'(I) gave Hanako a big English dictionary.'

b. ookii eego no zisho o Hanako ni ageta

c. * ookii Hanako ni eego no zisho o ageta

d. * eego no Hanako ni ookii zisho o ageta

Scrambling the whole ARGs is acceptable as in (6b), while it is unacceptable to split the ARG by scrambling as in (6c) and (6d).

It is also the case with scrambling within a clausal argument that ARGs are minimal units for scrambling. Although (7a) is the most natural order for the three ARGs,

which parallels the word order of the corresponding clause, (7b)-(7f) are all acceptable, if not perfectly natural.

(7)

a. [$_{\text{PERIPHERY}}$ New York de no] [$_{\text{ARG}}$ keesatsu no] [$_{\text{ARG}}$ sono hannin no] [$_{\text{NUC}}$ taiho]
at GEN police GEN that criminal GEN arrest
'the arrest of the criminal by the police (which took place) in New York'

b. New York de no sono hannin no keesatsu no taiho

c. keesatsu no New York de no sono hannin no taiho

d. keesatsu no sono hannin no New York de no taiho

e. sono hannin no New York de no keesatsu no taiho

f. sono hannin no keesatsu no New York de no taiho

However, if the scrambling splits a nominal argument, it results in an ungrammatical phrase, as in (8c).

(8)

a. [$_{\text{ARG}}$ Ken ga UFO o mita toiu] [ARG Bill no] [$_{\text{NUC}}$ uwasa]
OBJ see:PST COMP GEN rumor
'the rumor by Bill that Ken saw a UFO'

b. Bill no Ken ga UFO o mita toiu uwasa

c. * Ken ga Bill no UFO o mita toiu uwasa

In general, it is also acceptable to scramble ARGs and NP operators; however, the phrase becomes unacceptable if locality operators, i.e. operators which express the location of the referent as to a reference point, appear closer to the head noun than more than one modifier of the inner layers. In (9e) and (9f), the demonstrative *ano*, which expresses *definiteness*, appears immediately next to the head noun, despite the adjective and the nominal argument. This is plausible given that the locality operators modify the whole NP, while adjectives and nominal arguments modify the inner layers, i.e. the nominal nucleus and the nominal core respectively.

(9)

a. [DEF ano] [ADJ muzukasii] [ARG furansugo no] [NUC hon][80]
that difficult French-lg. GEN book
'that difficult book about the French language'

b. ano furansugo no muzukasii hon

c. furansugo no ano muzukasii hon

d. muzukasii ano furansugo no hon

e. * muzukasii furansugo no ano hon

f. * furansugo no muzukasii ano hon

As shown in (9), it is acceptable to reverse an adjective and a nominal argument; however, if the argument contains a clause, the scrambling results in ungrammaticality. (10b) is unacceptable, where the noun complement and the adjective is reversed.

(10)

a. [ARG Ken ga UFO o mita toiu] [ADJ atarasii] [NUC uwasa]
OBJ see:PST QT new rumor
'the new rumor that Ken saw a UFO'

b. * atarasii Ken ga UFO o mita toiu uwasa

Furthermore, it is acceptable to reverse a nominal periphery and a nominal argument if they both consist of a clause, as in (11); however, it is not acceptable if the former consists of a clause, while the latter does not, as shown in (12b).

(11)

a. [PERIPHERY kinoo Fred ga kiita] [ARG Ken ga UFO o mita toiu] [NUC uwasa]
yesterday hear:PST OBJ see:PST COMP rumor
'the rumor that Ken saw a UFO which Fred heard yesterday'

b. Ken ga UFO o mita toiu kinoo Fred ga kiita uwasa

(12)

a. [PERIPHERY kinoo Fred ga kiita] [ARG Ken no] [NUC uwasa]
yesterday hear:PST GEN rumor
'Ken's rumor which Fred heard yesterday'

b. * Ken no kinoo Fred ga kiita uwasa

The same pattern is found in the reversal of a nominal periphery containing a clause and an adjective, which leads to ungrammaticality.

(13)
a. [$_{\text{PERIPHERY}}$ kinoo Fred ga mita] [$_{\text{ADJ}}$ atarasii] [$_{\text{NUC}}$ kuruma]
yesterday see:PST new car
'the new car which Fred saw yesterday'

b. * atarasii kinoo Fred ga mita kuruma

Given the discussion above, I summarize the constraints on postposing out of NPs and on word order variation within NPs in terms of the layered structure as follows.

(14) A sentence is unacceptable if the nominal nucleus of an NP alone occurs in the postverbal position, being separate from the modifiers of the head noun in the canonical position.

(15) Scrambling results in unacceptability if:
(a) it splits a nominal argument,
(b) the locality operator appears closer to the head noun than more than one modifiers of the inner layers,
(c) modifiers of two different layers are reversed when the modifier of the outer layer contain a clause and that of the inner layer does not.

I have discussed so far the constraints on non-canonical word order on the NP level. Below, I extend the analysis to the clause-level word order flexibility in terms of the layered structure of the clause in RRG. As seen in Section 4.2, there is flexibility in word order in simple sentences to a considerable extent. I repeat the sentence variants of scrambling for a simple ditransitive sentence in (16). All six variants of scrambling are possible, though the sentence becomes awkward, but not totally unacceptable, if there is more than one constituent fronted, as discussed in Section 4.2.

(16)
a. Ken ga Hanako ni sono hon o ageta
DAT that book OBJ give:PST
'Ken gave the book to Hanako.'

b. Hanako ni Ken ga sono hon o ageta

c. sono hon o Ken ga Hanako ni ageta

d. ? Hanako ni sono hon o Ken ga ageta

e. ? sono hon o Hanako ni Ken ga ageta

f. Ken ga sono hon o Hanako ni ageta

As seen in Section 5.4.2, however, if the sentence contains a *wa*-marked NP or PP, the scrambling makes the sentence awkward if an element crosses the *wa*-marked element, as shown in (17).

(17)

a. Ken wa Hanako ni sono hon o ageta
DAT that book OBJ give:PST
'Ken gave the book to Hanako.'

b. ? Hanako ni Ken wa sono hon o ageta

c. ? sono hon o Ken wa Hanako ni ageta

d. * Hanako ni sono hon o Ken wa ageta

e. * sono hon o Hanako ni Ken wa ageta

f. Ken wa sono hon o Hanako ni ageta

The sentences in (17b)-(17e) have the scrambling crossing the *wa*-marked NP and result in awkwardness as in (17b) and (17c) or complete unacceptability as in (17d) and (17e).

The same pattern is found in sentence variants with a *wa*-marked indirect object.

(18)

a. Hanako ni wa Ken ga sono hon o ageta
DAT that book OBJ give:PST
'Ken gave the book to Hanako.'

b. Hanako ni wa sono hon o ken ga ageta

c. * sono hon o Hanako ni wa Ken ga ageta

d. * Ken ga Hanako ni wa sono hon o ageta

e. * Ken ga sono hon o Hanako ni wa ageta

f. * sono hon o Ken ga Hanako ni wa ageta

Furthermore, we find the same pattern with a *wa*-marked object.

(19)

a. sono hon wa Ken ga Hanako ni ageta
that book DAT give:PST
'Ken gave the book to Hanako.'

b. sono hon wa Hanako ni Ken ga ageta

c. * Ken ga sono hon wa Hanako ni ageta

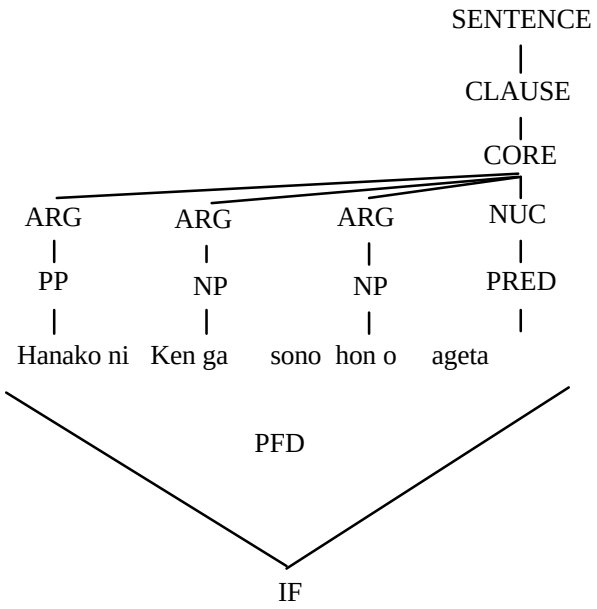
d. * Ken ga Hanako ni sono hon wa ageta

e. * Hanako ni Ken ga sono hon wa ageta

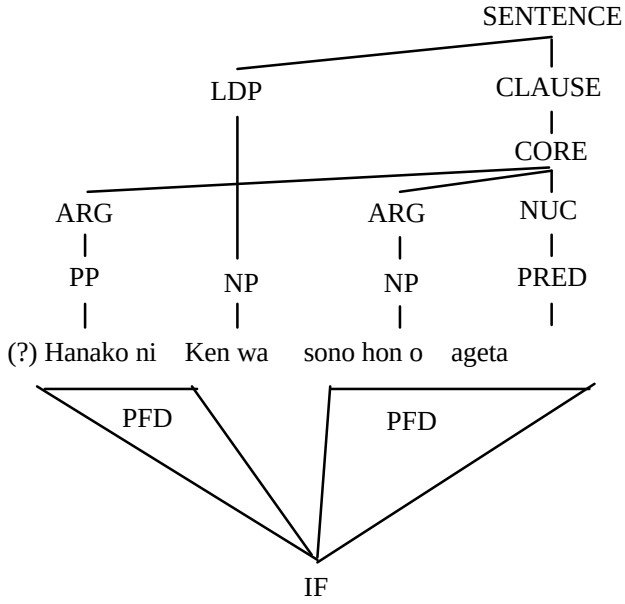
f. * Hanako ni sono hon wa Ken ga ageta

The observation above supports the existence of LDP for *wa*-marked elements; the sentence becomes awkward or unacceptable if the *wa*-marked element is preceded by other elements. As discussed in Section 5.4.2, *wa*-marked predicates and adverbs are not subject to the ordering constraint above; they are not in the LDP.

The principle that *wa*-marked elements in the LDP must be at the sentence initial position can be also accounted for in terms of the focus structure projection in Japanese. As discussed earlier, the PFD covers the whole sentence excluding the LDP. Sentences with the *wa*-marked elements in the LDP preceded by other elements exhibit a split in the PFD. Figures 4 and 5 illustrate this point by showing the constituent projection and the focus structure projection of (16b) and (17b). In Figure 4, *Ken ga* is an ARG; it is within the PFD, while *Ken wa* is in the LDP in Figure 5, which splits the PFD of the sentence.



'Ken gave the book to Hanako.' (=16b)
Figure 4



'Ken gave the book to Hanako.' (=17b)
Figure 5

At this point, I will turn to complex sentences for further investigation of non-canonical word order in terms of the constituent projection of RRG. As discussed in Section 5.2, interclausal relational structure is captured in terms of the following two

concepts in RRG, nexus relations and juncture levels. The former is the syntactic relations between the units; there are three possible juncture relations: *subordination*, *cosubordination*, and *coordination*. (See Section 5.2.) Juncture levels refer to the grammatical level of the linked units and there are three possible levels: clausal, core, and nuclear. I repeat the nine possible juncture-nexus types as in (20). As mentioned earlier, all human languages do not exhibit all nine types here; for example, Japanese lacks nuclear cosubordination.

(20) Nuclear Cosubordination
Nuclear Subordination
Nuclear Coordination
Core Cosubordination
Core Subordination
Core Coordination
Clausal Cosubordination
Clausal Subordination
Clausal Coordination

Following Hasegawa's (1992) classification of interclausal relational structures in Japanese, I examine below to what extent non-canonical word order is allowed in complex sentences.

To begin the discussion, I examine the following sentence variants of clausal subordination. In clausal subordination, one clausal unit is embedded in another. (22) is an example of clausal complement subordination with a *wa*-marked NP in the matrix clause.

(21)

a. Ken wa [Hanako ga Taro ni okane o kasita to] itta
DAT money OBJ lend:PST QT say:PST
'Ken said that Hanako lent money to Taro.'

b. [Hanako ga Taro ni okane o kasita to] Ken wa itta

c. * [Hanako ga Ken wa Taro ni okane o kasita to] itta

d. * [Hanako ga Taro ni Ken wa okane o kasita to] itta

e. * [Hanako ga Taro ni okane o Ken wa kasita to] itta

(21a) shows the canonical word order of the sentence with the *wa*-marked NP at the left-most position of the sentence and the subordinate clause with the order SUB+DAT+OBJ+V. The topic NP can also follow the subordinate clause as in (21b). However, (21c)-(21e) show that it cannot occur in any position within the subordinate clause.

It is not only a *wa*-marked NP but also a *ga*-marked NP of the matrix clause that cannot appear within the subordinate clause.

(21')

a. Ken ga [Hanako ga Taro ni okane o kasita to] itta
DAT money OBJ lend:PST QT say:PST
'Ken said that Hanako lent money to Taro.'

b. [Hanako ga Taro ni okane o kasita to] Ken ga itta

c. * [Hanako ga Ken ga Taro ni okane o kasita to] itta

d. * [Hanako ga Taro ni Ken ga okane o kasita to] itta

e. * [Hanako ga Taro ni okane o Ken ga kasita to] itta

As seen earlier, the elements within the subordinate clause can be reversed as well.

(21)

f. Ken wa [Hanako ga Taro ni okane o kasita to] itta
DAT money OBJ lend:PST QT say:PST
'Ken said that Hanako lent money to Taro.'

g. Ken wa [Hanako ga okane o Taro ni kasita to] itta

h. Ken wa [Taro ni Hanako ga okane o kasita to] itta

i. Ken wa [Taro ni okane o Hanako ga kasita to] itta

j. Ken wa [okane o Hanako ga Taro ni kasita to] itta

k. Ken wa [okane o Taro ni Hanako ga kasita to] itta

However, the scrambling within the subordinate clause cannot cross the subordinate predicate. As shown in (21l)-(21o), subordinate clauses in complex sentences exhibit limited word order flexibility relative to simple sentences because there is no postposing of any kind allowed within the subordinate clause.[81]

(21)

l. * Ken wa [kasita Hanako ga okane o Taro ni to] itta

m. * Ken wa [Hanako ga kasita okane o Taro ni to] itta

o. * Ken wa [Hanako ga okane o kasita Taro ni to] itta

It appears that the ungrammaticality of the sentences above may be explained by the general constraint that the complementizer *to* cannot be separated from the verb since it is not an independent morpheme. However, the following are also ungrammatical.

(21)

p. * Ken wa [kasita to Hanako ga okane o Taro ni] itta

q. * Ken wa [Hanako ga kasita to okane o Taro ni] itta

r. * Ken wa [Hanako ga okane o kasita to Taro ni] itta

Finally, the scrambling is possible within the whole matrix sentence.

(22)

a. Ken wa [Hanako ga Taro ni okane o kasita tte] itta n da yo[82]
DAT money OBJ lend:PST QT say:PST NOM COP FP
'Ken said that Hanako lent money to Taro.'

b. Ken wa itta n da yo [Hanako ga Taro ni okane o kasita tte]

c. itta n da yo Ken wa [Hanako ga Taro ni okane o kasita tte]

d. itta n da yo [Hanako ga Taro ni okane o kasita tte] Ken wa

e. [Hanako ga Taro ni okane o kasita tte] Ken wa itta n da yo

f. [Hanako ga Taro ni okane o kasita tte] itta n da yo Ken wa

Clausal adverbial subordination shows the same pattern as clausal complement subordination regarding the points observed above. A matrix *wa*/*ga*-marked NP can intervene between the subordinate clause and the matrix predicate, but it cannot occur within the subordinate clause.

(23)

a. Ken wa/ga [Hanako ga Taro ni okane o kasita node] odoroita
DAT money OBJ lend:PST because be-surprised:PST

'Ken was surprised because Hanako lent money to Taro.'

b. [Hanako ga Taro ni okane o kasita node] Ken wa/ga odoroita

c. * [Hanako ga Ken wa/ga Taro ni okane o kasita node] odoroita

d. * [Hanako ga Taro ni Ken wa/ga okane o kasita node] odoroita

e. * [Hanako ga Taro ni okane o Ken wa/ga kasita node] odoroita

The subordinate clause allows scrambling of its elements, except for the predicate, within the subordinate clause.

(23)

f. Ken wa/ga [Hanako ga okane o Taro ni kasita node] odoroita

g. Ken wa/ga [Taro ni Hanako ga okane o kasita node] odoroita

h. Ken wa/ga [Taro ni okane o Hanako ga kasita node] odoroita

i. Ken wa/ga [okane o Hanako ga Taro ni kasita node] odoroita

j. Ken wa/ga [okane o Taro ni Hanako ga kasita node] odoroita

k. * Ken wa/ga [kasita Hanako ga okane o Taro ni node] odoroita

l. * Ken wa/ga [Hanako ga kasita okane o Taro ni node] odoroita

m. * Ken wa/ga [Hanako ga okane o kasita Taro ni node] odoroita

Finally, the scrambling is possible in the whole matrix sentence.

(24)

a. odoroita n da yo
be-surprised:PST NOM COP FP

Ken wa/ga [Hanako ga okane o Taro ni kasita node]
money OBJ DAT lend:PST because
'Ken was surprised because Hanako lent money to Taro.'

b. Ken wa/ga odoroita n da yo [Taro ni Hanako ga okane o kasita node]

c. [Taro ni okane o Hanako ga kasita node] odoroita n da yo Ken wa/ga

Now I examine the following sentence variations of clausal cosubordination. In clausal cosubordination, neither unit is embedded in another; however, one unit is dependent on the other for the clause-level operators. In (25), the imperative IF operator is shared by the two clausal units.

(25)

a. kyoo wa [toshokan de hon o yonde] [hayaku uchi ni kaerinasai yo]
today library in book OBJ read early home to return:IMP FP
'Today, read books in the library and go home early.'

b. [toshokan de hon o yonde] kyoo wa [hayaku uchi ni kaerinasai yo]

c. ? [toshokan de kyoo wa hon o yonde] [hayaku uchi ni kaerinasai yo]

d. ?? [toshokan de hon o yonde] [hayaku kyoo wa uchi ni kaerinasai yo]

e. kyoo wa [hayaku uchi ni kaerinasai yo] [toshokan de hon o yonde]

f. * kyoo wa [toshokan de hon o uchi ni yonde] [hayaku kaerinasai yo]

(25a) shows the canonical order of the sentence. (25b) shows that the *wa*-marked NP can appear between the two juncts. However, the sentence becomes awkward if the matrix NP appears within either the first or the second clause, as in (25c) and (25d). Furthermore, the clauses cannot be scrambled together with an element of one clause in the other clause, as shown in (25f). Clausal cosubordination shows the same pattern as in clausal subordination in that a matrix NP can intervene the two clausal juncts, which indicates that the dependent clause in clausal subordination and cosubordination forms its own

information unit; it exhibits matrixhood. (See Section 5.4.4 for discussion of matrixhood of dependent clauses.)

In clausal coordination, neither unit is embedded in another; so, the two units are independent of each other in terms of clausal operators.

(26)

a. kinoo wa [daigaku de kaigi ga arimasita] ga [ikimasita ka]
yesterday university at meeting exist:PST CONJ go:PST Q
'Yesterday, there was a meeting at the university, and did you go (there)?'

b. ?? [daigaku de kaigi ga arimasita] ga kinoo wa [ikimasita ka]

c. * kinoo wa [daigaku de arimasita] ga kaigi ga [ikimasita ka]

d. kinoo wa [ikimasita ka] [daigaku de kaigi ga arimasita] ga

(26a) shows the canonical order of the sentence. In (26b), the *wa*-marked NP intervenes between the two juncts and it makes the original reading of the sentence difficult. The more likely reading of (26b) is that the *wa*-marked NP covers the second clause only, i.e. 'There have been meetings held at the university, and did you go to yesterday's?'. This suggests an important implication as to the difference between clausal coordination and clausal cosubordination/subordination. The two juncts in the former are more informationally independent of each other than those in the latter since the *wa*-marked NP intervening the two juncts normally covers only the following clause in clausal coordination, while it covers both of the two juncts in clausal cosubordination and subordination. This conforms to the nature of coordination, cosubordination, and subordination in RRG since in cosubordination and subordination one clause is dependent on the other, whether it is embeddedness or operator dependency, while in coordination the two juncts are independent of each other in both embeddedness and operators.

As in the clause-level juncture, Japanese exhibits the three nexus types in the core-level juncture. First, I examine core coordination sentences. In core coordination, there is no embeddedness relation between the two core juncts and they are independent of each other in terms of the core operators; however, they share an argument of the clause. A

verb with *zu ni* 'without -ing' is an example of core coordination in Japanese (Hasegawa 1992: 56). In (27), the *ga*-marked NP is the subject of both cores, while the negative operator modifies only the first core, i.e. there is no operator dependency on the core level.

(27)

a. Ken ga [benkyoo o sezu ni] [tesuto o uketa yo]
study OBJ do:NEG CMPL test OBJ take:PST FP
'Ken took the test without studying.'

b. [benkyoo o sezu ni] Ken ga [tesuto o uketa yo]

c. ?? [benkyoo o Ken ga sezu ni] [tesuto o uketa yo]

d. ?? [benkyoo o sezu ni] [tesuto o Ken ga uketa yo]

e. Ken ga [tesuto o uketa yo] [benkyoo o sezu ni]

(27a) shows the canonical word order of the sentence. The matrix NP can intervene between the two juncts, as in (27b); however, if it appears within either of the juncts, the sentence becomes awkward as in (27c) and (27d). Finally, the two juncts can be reversed as in (27e).

In core cosubordination, the linked cores share core operators. As in core subordination, the linkage is on the core level because the linked cores share an argument of the clause. A verb with *ni* 'in order to' is an example of core cosubordination in Japanese (Hasegawa 1992: 58). In (28), the negative operator is shared by the two linked cores.

(28)

a. Ken wa [kaimono o si ni] [Toronto ni ikanakatta yo]
shopping OBJ do CMPL to go:NEG:PST FP
'Ken didn't go to Toronto to do the shopping.'

b. [kaimono o si ni] Ken wa [Toronto ni ikanakatta yo]

c. ?? [kaimono o Ken wa si ni] [Toronto ni ikanakatta yo]

d. ?? [kaimono o si ni] [Toronto ni Ken wa ikanakatta yo]

e. Ken wa [Toronto ni ikanakatta yo] [kaimono o si ni]

As in core coordination, the matrix NP can intervene between the two juncts as in (28b). (28c) and (28d) show that the matrix NP cannot appear within either core. (28e) shows that the two juncts can be reversed.

In core subordination, there is an embeddedness relation between the two cores in that an argument of the matrix core contains the embedded core. The verbals such as *i-* 'be permitted', *daizyoobu* 'be all right', and *sum-* 'be settled' take a core argument with the *te* form of a verb; therefore, they exhibit core subordination (Hasegawa 1992).[83]

(29)

a. anata wa [pikunikku ni Ken o tsuretekite] ii yo
you picnic to OBJ bring be-permitted FP
'You may bring Ken to the picnic.'

b. * [pikunikku ni Ken o tsuretekite] anata wa ii yo

c. ?? [pikunikku ni Ken o anata wa tsuretekite] ii yo

d. anata wa ii yo [pikunikku ni Ken o tsuretekite]

Core subordination is analogous to core coordination and cosubordination in that the matrix NP cannot appear in the junct, as shown in (29c), while the two juncts can be reversed, as shown in (29d). However, unlike core coordination and cosubordination, (29b) shows that the sentence becomes ungrammatical with the matrix NP intervening the two juncts, and this suggests that the juncture in core subordination is more tight than the juncture in core coordination and cosubordination; in other words, the embedded core in core subordination exhibits less matrixhood than the juncts in core coordination and cosubordination.

The nuclear juncture in Japanese is primarily manifested by linked nuclei by the *te* form of verbs. As unique features of nuclear juncture, Hasegawa (1992: 69) notes three points as follows. Unlike core and clausal juncture, "[n]uclear juncture is subject to (i)

restrictions on possible intervening elements, (ii) obligatory joint participation in the domains of core-level and clause-level operators, and (iii) and intonational restriction: obligatory absence of *major phrase* boundaries." There are two nexus types for the nuclear-level juncture in Japanese, nuclear coordination and nuclear subordination, and Hasegawa (1992: 60) gives the following criterion to distinguish the two: "A subordinate nucleus does not participate in determination of the core arguments, only modifying the matrix nucleus, whereas coordinated nuclei jointly specify the arguments."

(30) is an example of nuclear coordination, given the observation that the valence of the *te* predicate, i.e. 'to repair', is changed in that the direct object *sono tokee* 'that clock' receives the nominative marking by *ga*, instead of the accusative marking by *o*. Furthermore, the negative operator modifies only the first nucleus; the two nuclear juncts do not share the operator.

(30)

a. sono tokee ga [mada shuurisinaide] [aru yo]
that clock yet repair:NEG exist FP
'That clock hasn't been repaired yet.'

b. * [mada shuurisinaide] sono tokee ga [aru yo]

c. * sono tokee ga [aru yo] [mada shuurisinaide]

As suggested by Hasegawa's (1992) characterization of nuclear-level juncture, the matrix NP cannot intervene between the two juncts, as shown in (30b). Also, the coordinated two juncts cannot be reversed, as shown in (30c).

This rigid linkage between two juncts is also the case with nuclear subordination. (31) is a case of nuclear subordination; unlike (30), the verb *simau*- 'to put' rarely takes the nuclear level negative operator *naide*, and also the verb does not specify the core arguments (Hasegawa 1992: 72); the core arguments are determined solely by the *te*-predicate *itte* 'to go.

(31)

a. Ken wa gaikoku ni itte [simatta yo]
foreign-country to go put:PST FP

'Ken has gone to a foreign country, to my regret.'

b. * gaikoku ni itte Ken wa [simatta yo]

c. * Ken wa [simatta yo] gaikoku ni itte

As in nuclear coordination, the matrix NP cannot intervene the two nuclear juncts, as shown in (31b), and the two juncts cannot be reversed, as shown in (31c).

Having examined complex sentences of eight juncture-nexus types in Japanese, we can summarize the systematicity in terms of word order flexibility. First, matrix elements can intervene the linked juncts on the clause and core levels, while they cannot for the nuclear level. Second, the two linked units can be reversed on the clause and core level, while they cannot be reversed on the nuclear level. (32) summarizes the word order flexibility in terms of the two points above.

(32)

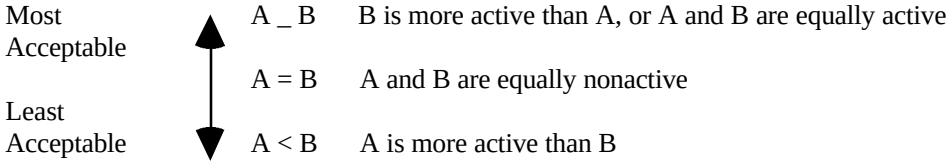
Juncture-Nexus Types	Intervenability of Matrix Elements	Reversibility of Juncts
Nuclear Subordination	no	no
Nuclear Coordination	no	no
Core Cosubordination	yes	yes
Core Subordination	no	yes
Core Coordination	yes	yes
Clausal Cosubordination	yes	yes
Clausal Subordination	yes	yes
Clausal Coordination[84]	yes?	yes

The pattern in (32) conforms to the principle as to the tightness of the syntactic bond involved in the linkage (Foley and Van Valin 1984, Van Valin 1993a); the more tight the syntactic bond between the juncts, the less word order flexibility there is. There is both intervenability and reversibility in the clausal juncture and it is also the case in the core juncture, except for core subordination where there is not intervenability. In the nuclear juncture, on the other hand, there is no intervenability nor reversibility. Yet, one peculiar pattern in (32) is that a matrix element cannot intervene the two juncts in core subordination. This indicates that the embedded core of core subordination exhibits less matrixhood than the core of core coordination and cosubordination.

6.3. POSTPOSING AND FOCUS DOMAINS

In the previous section, I examined word order flexibility in general in terms of the constituent projection and the focus structure projection of RRG. In this section, I particularly focus on postposing construction in Japanese in terms of RRG analyses.

In Section 4.5.4, I proposed the acceptability hierarchy for postposing which predicts that the postposing construction is most acceptable when the postverbal elements are more active than the preceding elements or the postverbal elements and the preceding elements are equally active. The hierarchy is repeated below in Figure 6.



A: proposition expressed by the preverbal elements, including the predicate
B: elements in the postverbal position

Figure 6: Acceptability hierarchy for postposing

For the three activation patterns in the hierarchy above, there are five possible patterns in terms of the focus structure projection. For the most acceptable type, there are three possible patterns. First, the postverbal elements and the preceding elements are both within the actual focus domain [AFD]; however, the preceding elements are less active, i.e. more focused, than the postverbal elements. Given the different degrees of focus as in this pattern, I posit two subdomains of AFD: the primary focus domain [PrFD] and the secondary focus domain [SFD]. Elements in the PrFD exhibit higher degree of focus than

those in the SFD, and in the above type of postposing the postverbal elements are within the SFD and the preceding elements within the PrFD, which is shown in Figure 7.

Another focus pattern in the most acceptable postposing type is exemplified by the cases in which the postverbal elements are marked with *wa*. *Wa*-marked elements, unless they are within the focus domain such as *wa*-marked adverbs and predicates, are in the left-detached position [LDP] in the preverbal position. Those *wa*-marked elements which appear in the postverbal position are in the *right-detached position* [RDP], which is analogous to the LDP in that it is outside of the IF operator; therefore, it is outside the focus domain. Figure 8 shows a sentence with a postverbal element in the RDP.

One other focus pattern in the most acceptable postposing type is the cases in which the postverbal and preceding elements are equally active. This focus pattern is exemplified by *ga*-marked elements in the postverbal position which is outside the AFD, which is shown by Figure 9.

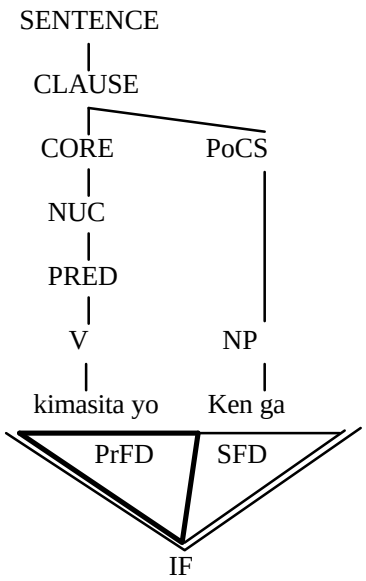


Figure 7: 'Ken came.'

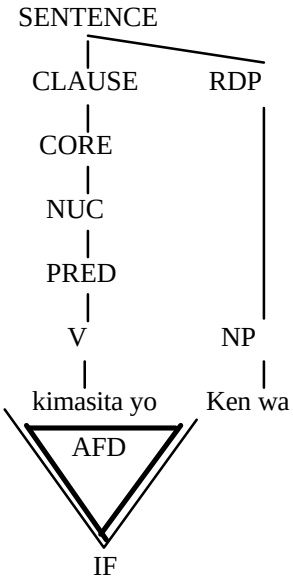


Figure 8: 'Ken came.'

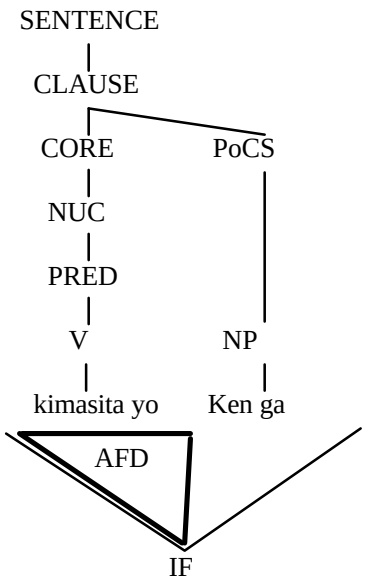


Figure 9: 'Ken came.'

In the second most acceptable postposing type, the postverbal elements and the preceding elements are equally nonactive, hence, the whole sentence is within the AFD. However, the postverbal elements and the preceding elements are differentiated within the AFD: the postverbal elements in the SFD, while the preceding elements in the PrFD since the information conveyed by the preceding elements is more important, i.e. more focused, than the information conveyed by the postverbal elements, as discussed in Section 4.4. The focus pattern for this postposing type is identical with the one which is shown in Figure 7.

Finally, in the least acceptable postposing type, the postverbal elements are less active than the preceding elements. There are two logical possibilities for this type: (i) the SFD is followed by the PrFD in the postverbal position, which is shown by Figure 10, and (ii) the AFD appears in the postverbal position, as in Figure 11.

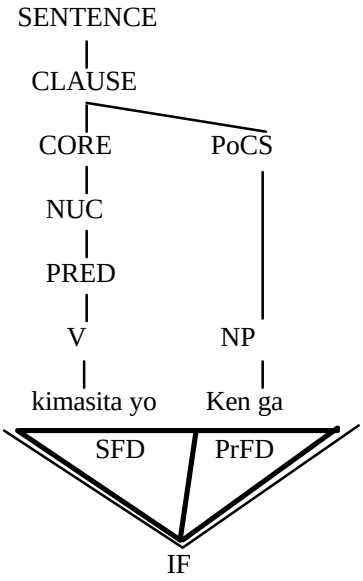


Figure 10: 'Ken came.'

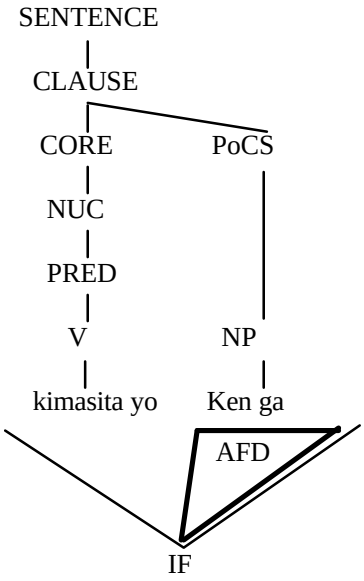


Figure 11: 'Ken came.'

In terms of the constituent projection, I posit the *postcore slot* [PoCS] for the postverbal elements within the PFD. While the PoCS corresponds to the precore slot [PrCS] in terms of the layered structure, i.e. they are both within the CLAUSE but outside the CORE, the PoCS contrasts with the PrCS in terms of the focus structure in that the PrCS represents an element of the primary focus, e.g. a *ga*-marked focused argument in the argument focus construction, while the PoCS represents elements within the PFD, which are either focus or nonfocus. Structurally, the PoCS distinguishes the postverbal elements in Japanese from elements which are canonically placed postverbally, for example, an object in SVO languages, as in English. I will give further argument for the PoCS in later sections.

In Section 6.2, I examined non-canonical word order of sentences with a *wa*-marked NP and pointed out that the elements of such sentences cannot be scrambled with a *wa*-marked NP. I repeat variants of a ditransitive sentence with *wa*-marked NP in (33). In (33b)-(33e), the sentences are not acceptable; the *wa*-marked NP in the LDP intervenes the focus domain.

(33)

a. Ken wa Hanako ni sono hon o ageta
DAT that book OBJ give:PST
'Ken gave the book to Hanako.'

b. ? Hanako ni Ken wa sono hon o ageta

c. ? sono hon o Ken wa Hanako ni ageta

d. * Hanako ni sono hon o Ken wa ageta

e. * sono hon o Hanako ni Ken wa ageta

f. Ken wa sono hon o Hanako ni ageta

The principle that the *wa*-marked NP should be in the *detached* position from the rest of the sentence also applies to the postposing construction with a *wa*-marked NP in the postverbal position. (34) shows the variants of a ditransitive sentence with a *ga*-marked NP and another argument in the postverbal position, while (34') shows the variants with an argument and a *wa*-marked NP in the RDP. In (34), reversing the order of the two postverbal arguments does not change the acceptability of the sentence. In (34'), on the other hand, a sentence becomes awkward if the *wa*-marked NP intervenes between the predicate and the argument NP in the postverbal position.

(34)

a. sono hon o ageta n da tte <u>sensee ga Ken ni</u>
that book OBJ give:PST NOM COP QT teacher DAT
'(I heard) the teacher gave the book to Ken.'

b. sono hon o ageta n da tte <u>Ken ni sensee ga</u>

c. Ken ni ageta n da tte <u>sensee ga sono hon o</u>

d. Ken ni ageta n da tte <u>sono hon o sensee ga</u>

(34')

a. ? sono hon o ageta n da tte <u>sensee wa Ken ni</u>
that book OBJ give:PST NOM COP QT teacher DAT
'(I heard) the teacher gave the book to Ken.'

b. sono hon o ageta n da tte <u>Ken ni / sensee wa</u>

c. ? Ken ni ageta n da tte sensee wa sono hon o

d. Ken ni ageta n da tte sono hon o / sensee wa

As expected, if there is a considerable pause between the postverbal elements and the preceding elements in (34'a) and (34'c), the sentence becomes acceptable. In (34a) and (34c), on the other hand, the sentence is acceptable even without such a pause. In (34'b) and (34'd), the sentences require a pause between the argument and the *wa*-marked NP, which suggests that *wa*-marked elements of non-focus must be in a detached position from the rest of the sentence, whether they are in the LDP or the RDP. However, if the RDP is the only element in the postverbal position, the sentence does not require a pause between the *wa*-marked element and the predicate, as shown in (34'e). This appears to be contradictory to the claim that postverbal *wa*-marked elements are detached from the preceding string; however, the *wa*-marked elements in the postverbal position are already detached from the rest of the sentence even without an intervening pause, given the fact that the *wa*-marked NP appears in the postverbal position, which violates the canonical word order of SOV.

(34')

e. Ken ni sono hon o ageta n da tte sensee wa

In summary, as in the case of LDP, the RDP must be detached from the rest of the sentence in that they do not interrupt the focus domain of the sentence. This point is illustrated by Figures 12 and 13, which shows (34'a) and (34'b) respectively.

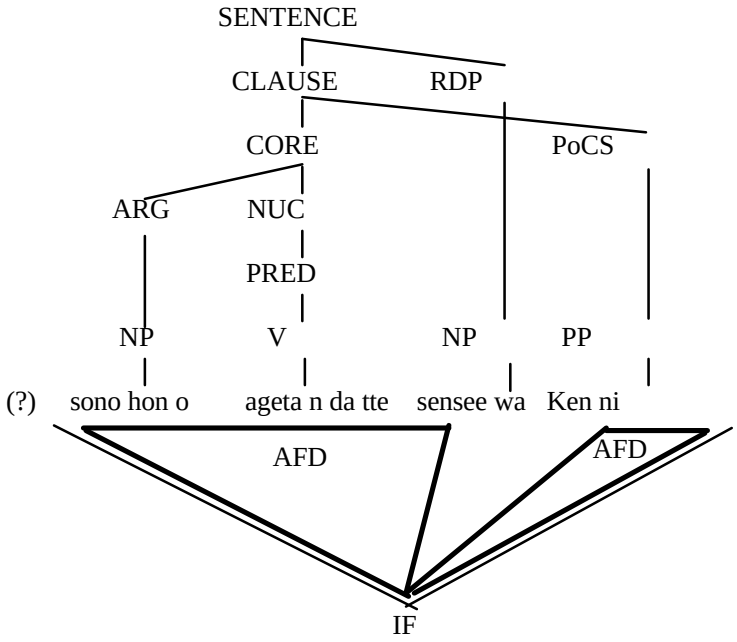


Figure 12: '(I heard) the teacher gave the book to Ken.' (=34'a)

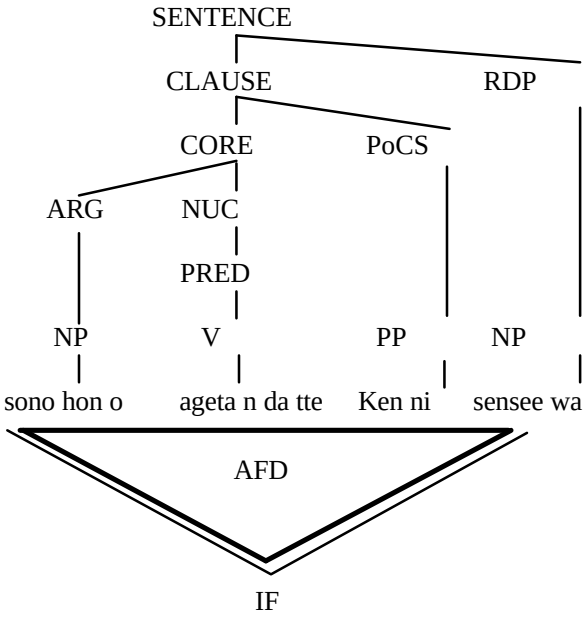


Figure 13: '(I heard) the teacher gave the book to Ken.' (=34'b)

The actual focus domain of the sentence in Figure 13 is interrupted by the *wa*-marked element in the RDP; so the sentence is awkward. On the other hand, it is not the case in the sentence in Figure 14; the element in the RDP is detached from the rest of the sentence

in that the RDP is outside the focus domain and there is a pause required in this sentence between the clause and the RDP.

6.4. POSTPOSING IN COMPLEX SENTENCES

In Section 4.2, I discussed the fact that postposing is widely possible in complex sentences, and in Section 6.2, I discussed possible non-canonical word order in terms of different clause linkage types and summarized that juncts are reversible on the clause and core level linkage. Despite of the considerable flexibility in postposing, there are constraints on postposing as well, as seen in Section 4.3.1. In this section, I discuss constraints on postposing in complex sentences and propose principles in terms of matrix-subordinate clause relationship from two different angles, the syntactic and discourse-pragmatic points of view.

6.4.1. Acceptability of Postposing and Matrixhood of Subordinate Clause

As discussed in Section 4.3.1, it has been noted (cf. Simon 1989, Hudson 1993) that a sentence is usually awkward or unacceptable if an element is postposed out of a subordinate clause. For example, a sentence is unacceptable if an element is postposed out of a relative clause and a noun complement clause, as shown in (35) and (36) respectively.

(35) * [kinoo ___ katta] kuruma o mita yo Ken ga
yesterday buy:PST car OBJ see:PST FP
'(I) saw the car which Ken bought yesterday.'

(36) * [Ken ga kinoo ___ katta tteyuu] uwasa o kiita yo kuruma o
yesterday buy:PST QT rumor OBJ hear:PST FP car OBJ
'(I) heard the rumor that Ken bought a car yesterday.'

Unlike noun-modifying clauses above, however, postposing out of clausal complements is often acceptable, as in (37).

(37) John wa [Ken ga Hanako ni ___ ageta tte] itteta yo hon o
DAT give:PST QT say:PST FP book OBJ
'John was saying that Ken gave a book to Hanako.'

The same kind of contrast between noun-modifying clauses and clausal complements is found in postposability of the subordinate clause as a whole; placing the whole subordinate clause results in awkwardness in the case of noun-modifying clauses, while it is not the case with clausal complements.

(38)
a. ? ___ gakusee wa saiyoosare soo desu ka [Suzuki kyoozyu ga suisensiteiru]
student employ:PAS likely COP Q prof. recommending
'Is the student whom Prof. Suzuki is recommending likely to be employed?'

b. ___ saiyoosare soo desu ka [Suzuki kyoozyu ga suisensiteiru] gakusee wa
'Is the student whom Prof. Suzuki is recommending likely to be employed?'

(39)
a. ? Ken wa ___ uwasa o sinziteru no? [Taro ga okane o nusunda toiu]
rumor OBJ believe FP money OBJ steal:PST QT
'Does Ken believe the rumor that Taro stole the money?'

b. Ken wa ___ sinziteru no? [Taro ga okane o nusunda toiu] uwasa o
'Does Ken believe the rumor that Taro stole the money?'

(40) Ken wa sinziteru no? [Taro ga okane o nusunda tte]
believe FP money OBJ steal:PST QT
'Does Ken believe that Taro stole the money?'

In (38) and (39), placing only the subordinate clause in the postverbal position results in awkwardness, and this is especially the case if there is no intervening pause between the matrix predicate and the postverbal elements. Note that placing a noun-modifying clause with the head noun in the postverbal position is perfectly acceptable, even without the intervening pause before them, as shown in (38b) and (39b). As in (40), on the other hand, postposing the clausal complement is invariably acceptable, regardless of presence or absence of the intervening pause.

In Section 5.4.4, I proposed a matrixhood hierarchy of the four types of subordinate clauses: noun-modifying clauses, clausal complements, 'if' clauses, and 'because' clauses. I repeat the hierarchy below in Figure 14.

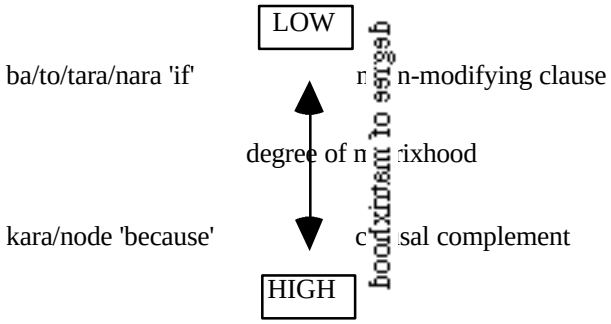


Figure 14: Matrixhood hierarchy of the four types of subordinate clause

Given the observation that postposing of or out of clausal complements is more acceptable than postposing of or out of noun-modifying clauses, I propose the correlation between the matrixhood of subordinate clause and the acceptability of postposing; namely, the greater the matrixhood a subordinate clause exhibits, the more acceptable it is to place a part of or the whole subordinate clause in the postverbal position.

The same correlation is found by the contrast between adverbial subordinate clauses of 'if' and 'because' in postposing construction with no intervening pause; placing whole 'because' clauses is more acceptable than placing whole 'if' clauses.

(41) ? Ken wa ___ konsaato ni ikanai tte [Hanako ga iku nara]
concert to go:NEG QT go if
'(I heard) Ken won't go to the concert if Hanako goes.'

(42) ? Ken wa ___ konsaato ni ikanai tte [Hanako ga ike ba]
concert to go:NEG QT go if
'(I heard) Ken won't go to the concert if Hanako goes.'

(43) ? Ken wa ___ konsaato ni ikanai tte [Hanako ga i ttara]
concert to go:NEG QT go if
'(I heard) Ken won't go to the concert if Hanako goes.'

(44) ? Ken wa ___ benkyoo dekinai tte [Hanako ga kuru to]
study do:POT:NEG QT come if
'(I heard) Ken cannot study if Hanako comes.'

(45) Ken wa ___ konsaato ni ikanai tte [Hanako ga iku kara/node]
concert to go:NEG QT go because

'(I heard) Ken won't go to the concert because Hanako goes.'

Note that the postposing constructions in (41)-(44) are acceptable if there is an intervening pause between the postverbal elements and the preceding elements, in which case the postverbal elements are construed as separate from the preceding matrix clause. However, without such a unit boundary, (41)-(44) are less acceptable than (45), if not completely unacceptable.

The observation above indicates that the degree of acceptability in postposing of or out of a subordinate clause correlates with the degree of matrixhood, which varies according to the types of subordinate clauses. Figure 15 illustrates the correlation between the matrixhood of subordinate clause and the acceptability of postposing, discussed above.

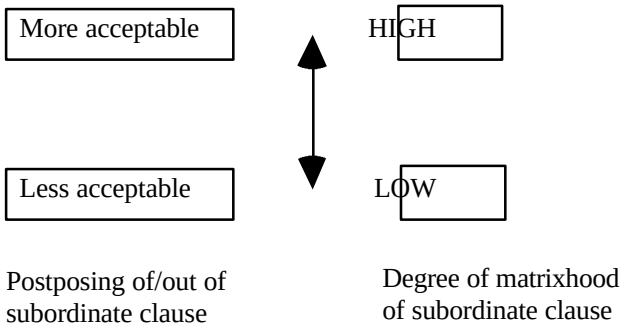


Figure 15: Acceptability of postposing and degree of matrixhood of subordinate clause

The hierarchy in Figure 15 accounts for the varying degrees of acceptability in postposing; however, it accounts for the data only partially. It is not only the matrixhood of the subordinate clause but also the types of matrix clause that influence the acceptability of postposing out of a subordinate clause. This point is illustrated by the following.

(46)

a. * kore wa [___ kinoo kabutteta] boosi da yo Ken ga
this yesterday wear:PST hat COP FP
'This is the hat which Ken was wearing yesterday.'

b. kore wa [___ kinoo kabutteta] no da yo Ken ga
this yesterday wear:PST NOM COP FP
'This is the one which Ken was wearing yesterday.'

(47)

a. * [Ken ga ___ ageta] hon o yonda? Hanako ni
give:PST book OBJ read:PST DAT
'Did (you) read the book which Ken gave to Hanako?'

b. [Ken ga ___ ageta] no yonda? Hanako ni
give:PST NOM read:PST DAT
'Did (you) read the one which Ken gave to Hanako?'

(46a) and (47a) are cases of postposing out of a relative clause, which results with ungrammaticality. In (46b) and (47b), however, the same postposing is acceptable if the head noun is replaced with the nominalizer *no*.[85]

The same contrast is found in sentences with noun complement clauses.

(48)

a. ? [Ken ga ___ hon o ageta tteyuu] uwasa o sitteru? Hanako ni
book OBJ give:PST QT rumor OBJ knowing DAT
'Do (you) know the rumor that Ken gave a book to Hanako?'

b. [Ken ga ___ hon o ageta tteyuu] no sitteru? Hanako ni
book OBJ give:PST QT NOM knowing DAT
'Do (you) know the one that Ken gave a book to Hanako?'

(48a) shows that placing an element of the noun complement in the postverbal position is unacceptable, while (48b) shows that the same postposing is acceptable if the head noun is replaced with the nominalizer. The observation above indicates that the postposing out of a subordinate clause is acceptable when the subordinate clause functions like a matrix argument. Unlike (46a), (47a), and (48a), the subordinate clauses in (46b), (47b), and (48b) are analogous to clausal complements in that the head noun is lexically absent. The contrast in terms of the *matrix-argumenthood* is illustrated by the fact that in (46a), (47a), and (48a) the matrix clause can remain grammatical without the subordinate clause, while in (46b), (47b), and (48b) the matrix clause requires the subordinate clause to be a grammatical sentence. (46')-(48') show that the sentences with the full head noun can stand as a grammatical sentence without the subordinate clause, while it is not the case with the sentences with the nominalizer.

(46')
a. kore wa boosi da yo
this hat COP FP
'This is a hat.'

b. * kore wa no da yo
this NOM COP FP
'This is the one.'

(47')
a. hon o yonda?
book OBJ read:PST
'Did (you) read the book?'

b. * no yonda?
NOM read:PST
'Did (you) read the one?'

(48')
a. uwasa o sitteru?
rumor OBJ knowing
'Do (you) know the rumor?'

b. * no sitteru?
NOM knowing
'Do (you) know the one?'

The observation above further supports the claim that there are different degrees of matrixhood of subordinate clause and the degree of matrixhood correlates with acceptability of postposing out of the subordinate clause, as illustrated in Figure 15. In the case of (46)-(48), the more the subordinate clause function like a valence-satisfying unit, i.e. an argument, the more acceptable the postposing out of the subordinate clause is.

The correlation between the degree of matrixhood of subordinate clause and the acceptability of postposing is further illustrated by the contrast between ordinary relative clauses and the construction which is so-called internally headed relative clauses (cf. Kuruda 1975-76, 1976, 1976-77, Ishii 1989, Hirose and Ohori 1992, Ohara 1992, 1994). Unlike ordinary externally headed relative clauses, internally headed relative clauses are characterized by its head NP occurring within the subordinate clause, which is an

argument of both the subordinate predicate and the matrix predicate; i.e. there is no gap in the subordinate clause.[86] (49a) and (50b) show examples of ordinary externally headed relative clause and internally headed relative clause respectively.

(49)

a. keesatsu wa [ginkoo ni haitta] gootoo o tsukamaeta n da yo
police bank to enter:PST robber OBJ capture NOM COP FP
'The police captured the robber who entered the bank.'

b. keesatsu wa [gootoo ga ginkoo ni haitta] no o tsukamaeta n da yo
police robber bank to enter:PST NOM OBJ capture:PST NOM COP FP
'The police captured the robber who entered the bank.'

Interestingly, the contrast between externally headed relative clause and internally headed relative clause is also shown by the acceptability in postposing an element out of the subordinate clause; it is generally unacceptable to postpose an subordinate element of ordinary externally headed relative clauses, while it is often acceptable to do so out of internally headed relative clauses.[87] The examples in (50) illustrate this point.

(50)

a. * keesatsu wa [___ haitta] gootoo o tsukamaeta n da yo <u>ginkoo ni</u>
police enter:PST robber OBJ capture:PST NOM COP FP bank to
'The police captured the robber who entered the bank.'

b. keesatsu wa [gootoo ga ___ haitta] no o tsukamaeta n da yo <u>ginkoo ni</u>
police robber enter:PST NOM OBJ capture:PST NON COP FP bank to
'The police captured the robber who entered the bank.'

Internally headed relative clauses are themselves arguments of the matrix predicate, while externally headed relative clauses are not, as shown by the contrast in (50'); the matrix sentence is grammatical without the relative clause in (50'a), while it is not the case in (50'b). This contrast accounts for the fact that internally headed relative clauses exhibit greater acceptability in postposing out of the relative clause than ordinary externally headed relative clauses, further supporting the claim made above that postposing out of a subordinate clause is acceptable when the subordinate clause functions as a matrix argument.

(50')

a. keesatsu wa gootoo o tsukamaeta n da yo
police robber OBJ capture:PST NOM COP FP
'The police captured the robber.'

b. * keesatsu wa no o tsukamaeta n da yo
police NOM OBJ capture:PST NON COP FP
'The police captured (the robber).'

6.4.2. Acceptability of Postposing and Focus on Subordinate Clause

The previous section examined the acceptability of postposing of or out of a subordinate clause in terms of the matrixhood of subordinate clause. In this section, I examine the acceptability of postposing out of a subordinate clause from a different angle, in terms of focus structure of complex sentences. The previous section discussed that postposing out of noun-modifying clauses is usually less acceptable than postposing out of clausal complements, and I claimed that this general tendency is due to the difference in the degree of matrixhood of the subordinate clause, i.e. clausal complements exhibit greater matrixhood than noun-modifying clauses.

Despite the general patterns discussed above, there are cases in which postposing out of a noun-modifying clause is acceptable, and furthermore, postposing out of a clausal complements is unacceptable. In (51B), postposing out of the relative clause is unacceptable, which follows the general principle discussed above, i.e. postposing out of a noun-modifying clause is normally unacceptable. In (52B), on the other hand, postposing out of the relative clause is acceptable, despite the fact that (52B) is structurally parallel to (51B).

(51)

A: kore ga [Ken ga kinoo moratta] biiru na no?
this yesterday receive:PST beer NOM FP
'Is this the beer which (someone) gave to Ken yesterday?'

B: * iya kore wa [___ kinoo moratta] wain da yo <u>Ken ga</u>
no this yesterday receive:PST wine COP FP

'No, this is wine which (someone) gave to Ken yesterday.'

(52)
A: kore ga [Ken ga kinoo moratta] biiru na no?
this yesterday receive:PST beer NOM FP
'Is this the beer which (someone) gave to Ken yesterday?'

B: iya kore wa [___ kinoo katta] biiru da yo Ken ga
no this yesterday buy:PST beer COP FP
'No, this is the beer which Ken bought yesterday.'

In terms of the constituent projection, (51B) is identical with (52B); however, their difference in acceptability of postposing is accounted for by the difference in the focus structure of the sentences. In (51B), the subordinate clause represents the active proposition *Ken ga kinoo moratta*, which is mentioned in (51A), while the matrix clause represent the nonactive proposition *kore wa/ga wain* due to the element *wain*, which is not mentioned in (51A). The focus structure for (51B) exhibits the AFD which covers the matrix clause but not the subordinate clause. Figure 16 shows the focus structure projection of the sentence in (51B), along with its constituent projection.

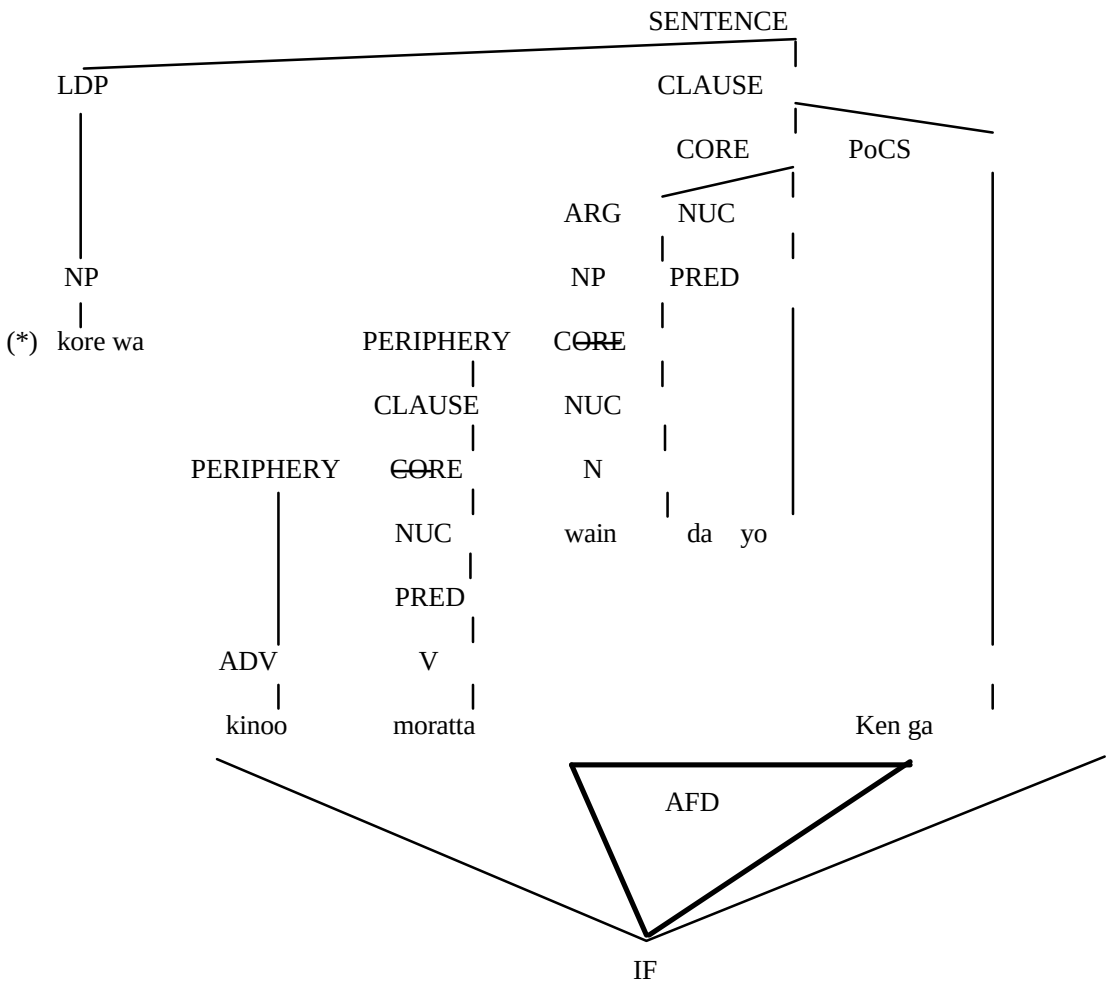


Figure 16: 'This is wine which (someone) gave to Ken yesterday.' (=51B)

In (52B), on the other hand, the subordinate clause represents a nonactive proposition because of the nonactive predicate *katta*, while the matrix clause represents the active proposition *kore wa/ga biiru*, which is mentioned in (52A). In terms of the focus structure, the AFD covers a part of the subordinate clause, while it is not the case with the matrix clause. The focus structure projection and the constituent projection for (52B) are shown in Figure 17.

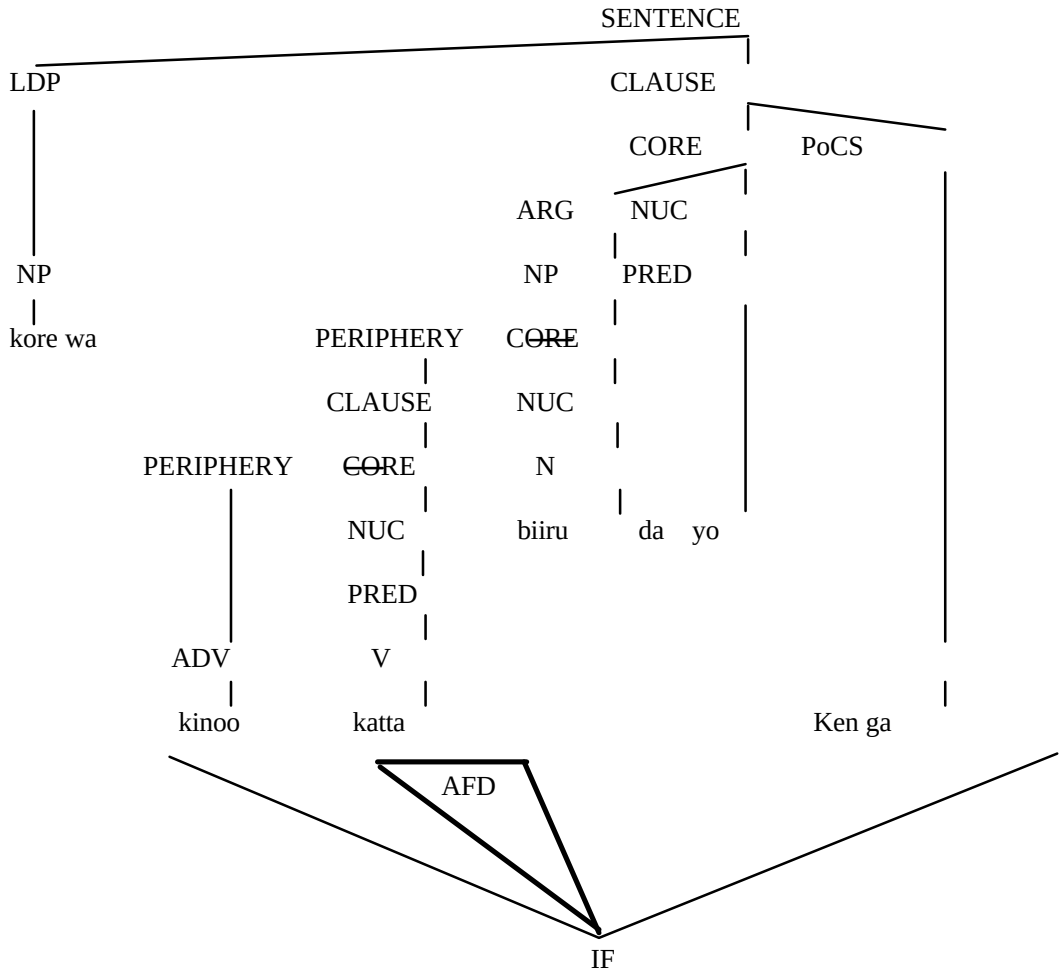


Figure 17: 'This is the beer which Ken bought yesterday.' (=52B)

As suggested by Figures 16 and 17, the difference in acceptability of postposing between (51B) and (52B) is accounted for in term of the focus structure; postposing is acceptable if it is out of a subordinate clause which carries the actual focus of the sentence, while it is unacceptable if it is out of a subordinate clause of non-focus.

The same pattern is found in postposing out of a noun complement.

(53)
A: [Ken ga takarakuzi o ateta tteyuu] uwasa kiita?
lottery OBJ hit:PST QT rumor hear:PST
'Did (you) hear the rumor that Ken hit the lottery?'

B: * iya sore wa [___ takarakuzi o ateta tteyuu] demakase da yo Ken ga
no that lottery OBJ hit:PST QT lie COP FP
'No, it is the lie that Ken hit the lottery.'

(54)
A: [Ken ga takarakuzi o ateta tteyuu] uwasa kiita?
lottery OBJ hit:PST QT rumor hear:PST
'Did (you) hear the rumor that Ken hit the lottery?'

B: iya sore wa [___ takarakuzi o moosukoside ateta tteyuu] uwasa da yo Ken ga
no that lottery OBJ almost hit:PST QT rumor COP FP
'No, it is the rumor that Ken almost hit the lottery.'

In (53B), the subordinate clause represents the active proposition *Ken ga takarakuzi o ateta*, which is mentioned in (53A), while the matrix clause carries the focus because of the nonactive element *demakase*. The AFD of this sentence covers the matrix element, not the elements of the subordinate clause. This pattern is reversed in (54B); the subordinate clause conveys the nonactive proposition *Ken ga takarakuzi o moosukoside ateta*, while the matrix clause consists of the active element *uwasa*, which is mentioned in (54A). In terms of the focus structure projection, the AFD falls within the subordinate clause, while the matrix clause does not carry the actual focus. Figures 18 and 19 shows the focus structure projection for (53B) and (54B) respectively.

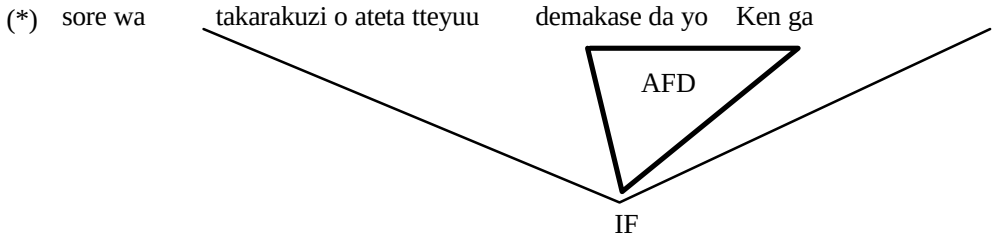


Figure 18: 'It is the lie that Ken hit the lottery.' (=53B)

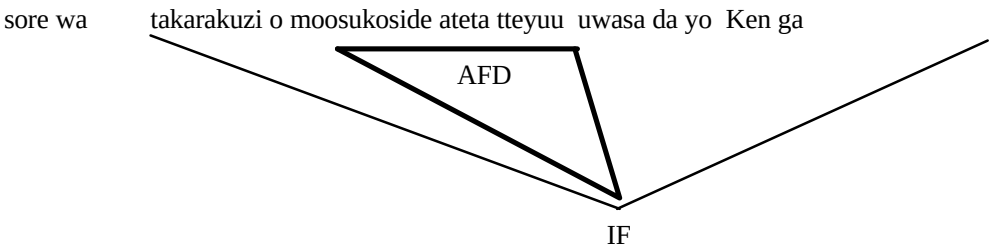


Figure 19: 'It is the rumor that Ken almost hit the lottery.' (=54B)

We have seen so far the cases in which postposing out of a noun-modifying clause, which is normally unacceptable, becomes acceptable because the postverbal elements are linked with the subordinate clause which carries the focus of the sentence. On the other hand, there are cases in which postposing out of a clausal complement, which is normally acceptable, becomes unacceptable. In (55B), the postposing out of the clausal complement is acceptable, while it is not acceptable in (56B), despite the structural parallel between the two.

(55)
A: Ken wa [Hanako ga kinoo kuru tte] omotteta no?
yesterday come QT thinking:PST FP
'Did Ken think that Hanako would come yesterday?'

B: iya Ken wa [___ ototoi kuru tte] omotteta n da yo <u>Hanako ga</u>
no day-before-yes. come QT thinking:PST NOM COP FP
'No, Ken thought that Hanako would come the day before yesterday.'

(56)
A: Ken wa [Hanako ga kinoo kuru tte] omotteta no?
yesterday come QT thinking:PST FP
'Did Ken think that Hanako would come yesterday?'

B: * iya Ken wa [___ kinoo kuru tte] kanchigaisiteta n da yo <u>Hanako ga</u>
no yesterday come QT misunderstanding:PST NOM COP FP
'No, Ken wrongly understood that Hanako would come yesterday.'

In (55B), the subordinate clause contains the nonactive element *ototoi*; therefore, the AFD covers a part of the subordinate clause. On the other hand, the matrix clause

conveys only the active proposition *Ken wa omotteta*, which is mentioned in (55A); therefore, the AFD does not include the matrix clause. In (56B), on the other hand, the subordinate clause represents the active proposition *Hanako ga kinoo kuru*, which is mentioned in (56A), while the matrix clause represents the nonactive entity *kanchigaisiteta*; therefore, the AFD covers a part of the matrix clause. Given the contrast in focus structure between (55B) and (56B), the same principle applies here; postposing is acceptable if it is out of the subordinate clause which carries the focus of the sentence. Figures 20 and 21 illustrate this point with the focus structure projection for the two sentences above.

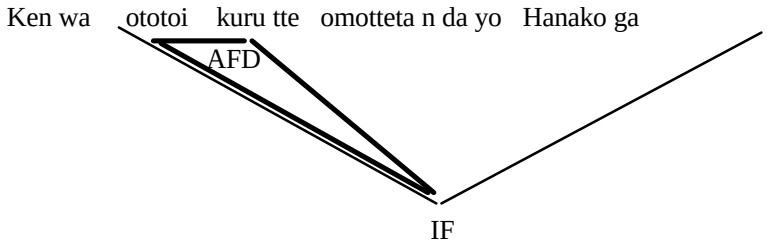


Figure 20: 'Ken thought that Hanako would come the day before yesterday.' (=55B)

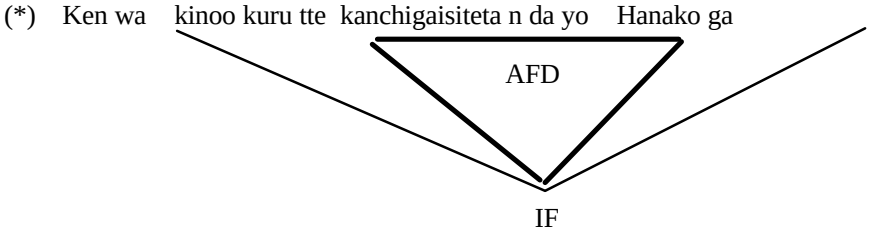


Figure 21: 'Ken wrongly understood that Hanako would come yesterday.' (=56B)

The discussion above suggests that the varying acceptability of postposing is accounted for by an independent principle, separate from the principle discussed in the previous section in terms of the matrixhood of subordinate clause. (57) summarizes the principle which I have discussed throughout this section.

(57) Postposing out of a subordinate clause is unacceptable if the subordinate clause does not carry the focus of the sentence.

This principle stated in (57) is analogous to the general restriction on question formation in complex sentences proposed by Van Valin (1993a, b), which was discussed in the previous chapter. I repeat the principle below as (58).

(58) General restriction on question formation: The element questioned (the WH-word in a simple, direct WH-question or the focal NP in a simple, direct yes-no question) must function in a clause which is within the PFD of the sentence.

Although in principle the restriction in (58) does not apply to Japanese since question formation is widely possible in subordinate clauses, the general restriction on postposing out of subordinate clauses in Japanese, stated in (57), appeals to the same general principle as the principle in (58) does; an element extracted out of a complex sentence, whether it is question formation or postposing, must be linked with the unit which carries the primary focus of the sentence.

In terms of the constituent projection, the fact that postposing out of subordinate clauses is possible provides further evidence for the PoCS. As discussed in Section 6.2, there is a general principle that Japanese does not allow non-canonical word order by placing an element across clause boundaries in complex sentences; however, this principle may be violated only if an element crossing a clause boundary is placed in the postverbal position. Therefore, the PoCS, instead of tangling elements of the layered structure, provides a better representation for postverbal elements, showing that there is a 'landing site' only at the postcore position for elements postposed. In this sense, the PoCS is analogous to other matrix-level elements of the layered structure of the clause, the LDP, the PrCS, and the RDP; these represent particular sentence elements which appear in a fixed syntactic position; the LDP and the RDP, for example, represents non-focus elements, which appear in a detached position from the clause.

6.5. CONCLUSION

This chapter has revealed an interesting interaction between syntax and pragmatics as to the acceptability of non-canonical word order in Japanese; satisfactory accounts cannot be provided by either one alone. The constraints on non-canonical word order are structural to a considerable extent, while the fundamental principle behind postposing is accounted for by contextual factors; i.e. the postverbal elements in general represent active information in the immediately preceding context.

I have also discussed that there are structural constraints to a considerable extent on non-canonical word order in complex sentences as well, depending on the clause linkage types. Furthermore, I have examined varying degrees of acceptability in postposing of or out of a subordinate clause and found that the accounts require the interplay between syntax and pragmatics here as well. I have proposed that there are at least two types of principles which correlate with the degree of acceptability in postposing of or out of a subordinate clause: (i) dependency relationship between matrix clause and subordinate clause, and (ii) focus relationship between matrix clause and subordinate clause. The first is viewed from two different angles: the matrixhood of the subordinate clause, and the dependence of the matrix clause on the subordinate clause. In terms of the former, the greater the matrixhood a subordinate clause exhibits, the more acceptable it is to postpose the subordinate clause or out of the subordinate clause. In terms of the latter, the more dependent a matrix clause is on the subordinate clause, the more acceptable it is to postpose out of the subordinate clause.

The focus relationship between matrix and subordinate clauses forms another principle, which is independent of the two principles discussed above. The two principles above are context-independent in that the degree of matrixhood of subordinate clause and dependency of matrix clause on subordinate clause varies depending on sentence structures, rather than the discourse context. The principle based on the focus relationship, on the other hand, predicts that the acceptability of postposing is also affected by

preceding discourse context since the focus structure of a sentence is determined by the preceding context. The postposing should be out of a subordinate clause which represents the primary focus of the sentence, i.e. the proposition conveyed by the subordinate clause should be less active than that conveyed by the matrix clause.

CHAPTER 7

CONCLUDING REMARKS

In this dissertation I have investigated the interface between morphosyntax and focus structure in Japanese. I have demonstrated throughout that seemingly two separate phenomena, *wa* and *ga* in morphology and non-canonical word order in syntax, are in fact related to each other in terms of focus structure of the sentence. This final chapter will illustrate this point by summarizing the findings in each chapter.

In Chapter 2, I investigated the use of *wa* and *ga* in mini-discourses and proposed the functional contrast between the two in terms of the focus structure, as stated in (1).

(1)

WA: A *wa*-marked element is non-focus and it combines with the open proposition which is the focus of the utterance.

GA: a. A *ga*-marked element and the open proposition which it combines with are both the focus of the utterance.

b. A *ga*-marked element is focus and it combines with the open proposition which is non-focus of the utterance.

c. A *ga*-marked element and the open proposition which it combines with are both non-focus of the utterance.

The functional contrast in (1) in principle illustrates the function of *wa* as a non-focus marker and it separates the *wa*-marked element from the rest of the sentence which is the focus of the sentence, while *ga* does not exhibit such a function; it simply plays a role as a case marker and does not manifest a specific focus type. In this sense, *ga* is *neutral* in terms of focus, unlike *wa* which is solely for the [non-focus-*wa*+focus] pattern. The characterization in terms of relative degree of activation clearly illustrates the contrast between *wa* and *ga* stated above, in terms of all possible focus patterns in *wa*/*ga* sentences, which is shown in Figure 1 and 2.

As shown in Figure 1, the only possible activation pattern for *wa* is the combination of active *wa*-marked element and nonactive predicate. *Ga*, on the other hand, can manifests any of the other possible activation patterns, as shown by the possible range between the two extremes (b) and (c) in the figure. For example, the *ga*-marked referent and the proposition represented by the predicate are equally nonactive, i.e. the neutral description in Kuno (1972, 1973) and the sentence focus in Lambrecht (1994), the predicate proposition is more active then the *ga*-marked referent, i.e. the exhaustive listing in Kuno (1972, 1973) and the argument focus in Lambrecht (1994), or the *ga*-marked referent and the predicate proposition are equally active, which I refer to as neutral focus in the present study. (a), (b), and (c) in Figure 2 illustrates the three focus types of *ga* above, which shows the flexibility of *ga* in terms of the possible range of focus structure of the sentence.

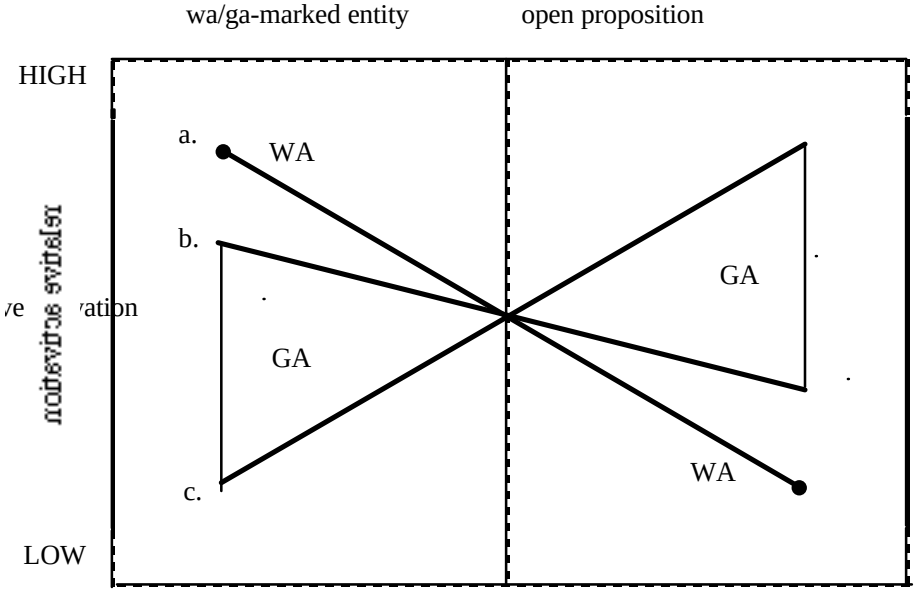


Figure 1: The contrast between *wa* and *ga* in terms of relative degrees of activation

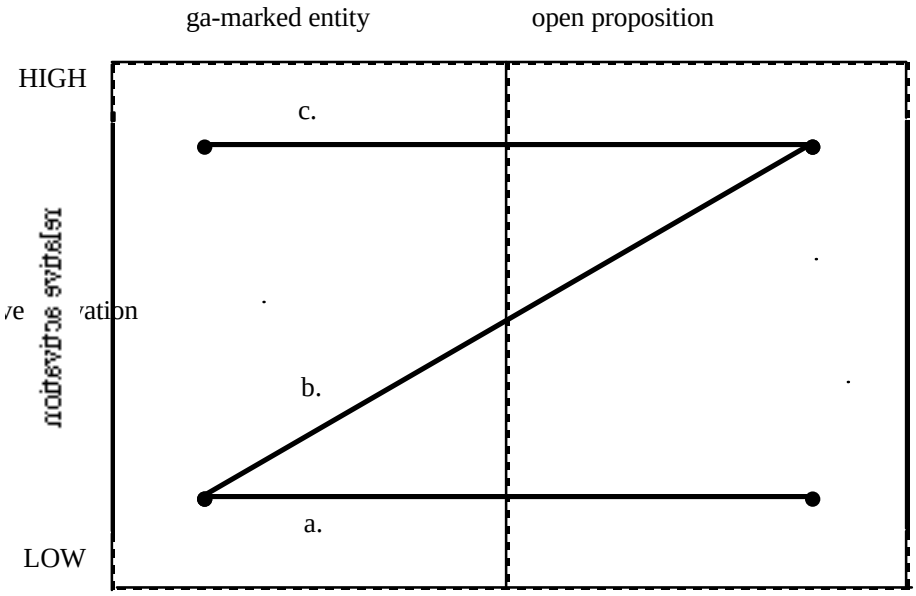


Figure 2: The three basic activation patterns of *ga*

In Chapter 3, I investigated conversational Japanese on the basis of the referential distance measurement [RD] and further supported the functional contrast between *wa* and *ga*, which I proposed in Chapter 2. It is worth noting the empirical finding that the third pattern of *ga*, (c) in Figure 2, is rather frequent in actual conversation, though this particular focus pattern of *ga* has not been discussed much in previous studies. The results from the RD measurement also showed that matrix clauses are not strikingly different from subordinate clauses in terms of the basic contrast between *wa* and *ga*; subordinate clauses show the same focus patterns of *wa* and *ga* as matrix clauses. At the same time, however, I found a tendency that *ga* tends to be substituted for *wa* in certain types of subordinate clauses than in other types, which is illustrated by the hierarchy in Figure 3.

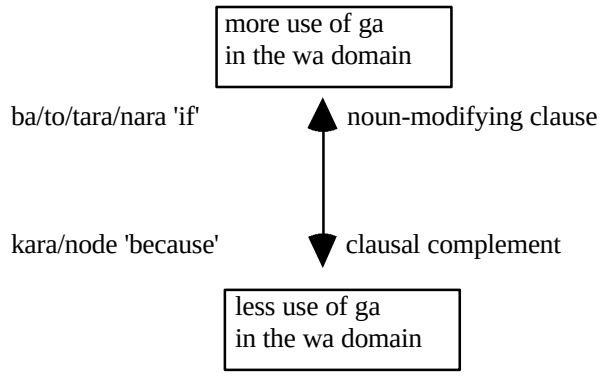


Figure 3: The use of *ga* in the *wa* domain and types of embedded clause

In Chapter 4, I extended the analysis in terms of focus structure to the syntactic-level phenomenon, postposing, specifically those cases without an intervening pause between the postverbal elements and the preceding elements. The observation of postposing on the mini-discourse level suggested that the acceptability for postposing is predicted by one general principle in terms of the notion of importance, which is repeated in (2).

(2) The postposing construction is acceptable if and only if the element in postverbal position is less important than the preceding elements.

The degree of importance may be indicated morphologically, e.g. sentence-final particles, or prosodically, i.e. prosodic emphasis on particular elements, or both. In terms of the notion of activation, the principle in (2) can be stated in terms of the hierarchy in Figure 4 with the four activation patterns of postposing.

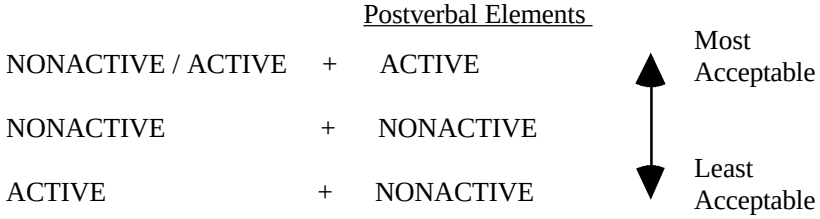


Figure 4: Acceptability hierarchy for postposing in terms of focus types

The hierarchy assumes the correlation between the relative degree of importance and the relative degree of activation; what is nonactive is always more important than what is

active, relative to the goal of the utterance in the surrounding discourse. Given this background, the optimal postposing type is the one with active elements in the postverbal position, while the least acceptable type is the one with nonactive elements in the postverbal position which follow active elements in the preceding string, which exhibits the pattern that the postverbal elements are more important than the preceding elements. The hierarchy in Figure 4 is empirically supported by the results from the RD measurement of the same conversational Japanese database as used in Chapter 3. The hierarchy reflects the token distribution in the database; the majority of the tokens fall on the most acceptable type, while there is almost no token on the least acceptable type.

The acceptability hierarchy for postposing above and the focus patterns of *wa*/*ga* sentences, which I proposed in Chapter 3, together account for the use of *wa* and *ga* in the postverbal position. Given the focus patterns of *wa*/*ga* sentences, there are four different types of *wa*/*ga* marked elements which may appear postverbally: (1) *wa*-marked element, i.e. the referent represented by the *wa*-marked postverbal element is more active than the proposition expressed by the preceding string, (2) *ga*-marked element whose referent is equally nonactive to the proposition expressed by the preceding element, (3) *ga*-marked element whose referent is less active than the proposition expressed by the preceding string, and (4) *ga*-marked element whose referent is equally active to the proposition expressed by the preceding string. However, the acceptability hierarchy in Figure 4 rules out the possibility stated in (3) above, since it is the least acceptable type of postposing construction. This interface between the focus patterns of *wa*/*ga* sentences and the acceptability hierarchy for postposing is reflected in the results from the RD measurement in the conversation database; there is not a single token of (3) above in the database.

In Chapters 5 and 6, I examined the same morphosyntactic phenomena in the RRG framework, more specifically in terms of the constituent projection and the focus structure projection of RRG. In Chapter 5, I discussed *wa*/*ga*-marked elements in terms of the scope of the IF operator and proposed that *wa*-marked elements are not always outside

the focus domain, i.e. non-focus; there is a demarcation between *wa*-marked NP/PP/clause and *wa*-marked adverb/predicate/WH phrase; the former is outside the focus domain, while the latter is within the focus domain.

In terms of the focus structure projection, it is useful to posit different degrees of focus within the actual focus domain [AFD]. I have referred to the most highly focused domain within the AFD as the primary focus domain [PrFD] and the less focused domain within the AFD as the secondary focus domain [SFD]. In this way, for example, we can properly represent the postposing construction where the postverbal elements are within the AFD but less focused than the preceding elements.

In Chapter 5, I also pointed out that different types of subordinate clauses exhibit different degrees of matrixhood and proposed a correlation between the degree of matrixhood and the use of *wa* and *ga* in subordinate clauses; namely, the more a subordinate clause functions like a matrix clause, the more acceptable it is for *wa* to be used in the subordinate clause. I examined four types of subordinate clauses and claimed the correlation between the degree of matrixhood and the degree of *ga* substitution for *wa,* which is shown in Figure 5.

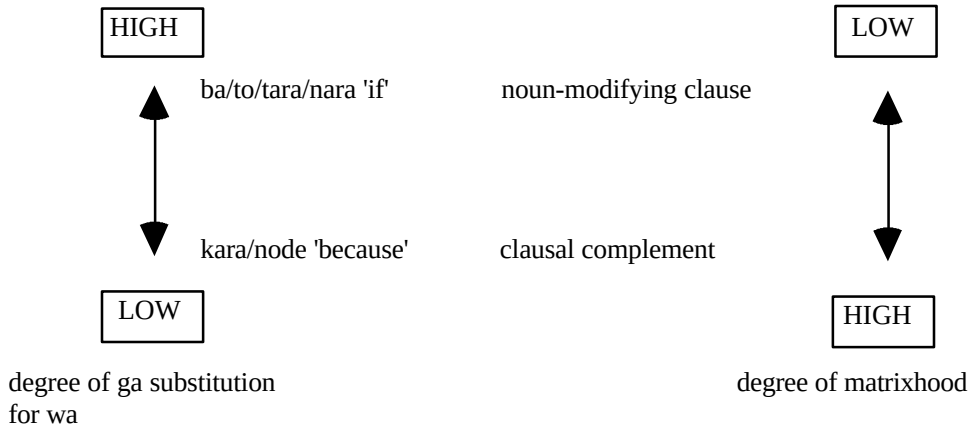


Figure 5: *Wa*/*ga*-switch and matrixhood hierarchy for the four subordinate clause types

In Chapter 6, I extended the RRG analysis to the non-canonical word order in general and particularly to the postposing construction. For non-canonical word order on

the phrase level, I laid out the basic constraints on non-canonical word order on the phrase level in terms of the layered structure of the clause in RRG, as in (3) and (4).

(3) A sentence is unacceptable if the nominal nucleus of an NP alone occurs in the postverbal position, being separate from the modifiers of the head noun in the canonical position.

(4) Scrambling results in unacceptability if:
(a) it splits a nominal argument,
(b) the locality operator appears closer to the head noun than more than one modifiers of the inner layers,
(c) modifiers of two different layers are reversed when the modifier of the outer layer contain a clause and that of the inner layer does not.

Furthermore, I discussed a basic clause-level constraint on non-canonical word order that *wa*-marked elements of non-focus must be in a detached position, either preceding the clause, i.e. in the left-detached position [LDP], or following the clause, i.e. in the right-detached position [RDP].

Japanese contrasts with languages such as English in that in Japanese the extraction out of a subordinate clause is acceptable to a considerable extent; however, it is not the case that postposing out of a subordinate clause is constraint-free. I discussed constraints on postposing out of subordinate clauses from two different angles, the structural and discourse-pragmatic points of view. Structurally, the constraints on extraction out of subordinate clauses can be accounted for by the notion of matrixhood of subordinate clause. I proposed a correlation between the degree of matrixhood of a subordinate clause and the acceptability of postposing of or out of the subordinate clause; the more a subordinate clause functions like a matrix clause, the more acceptable it is to postpose the subordinate clause or an element out of the subordinate clause, which is illustrated in Figure 6.

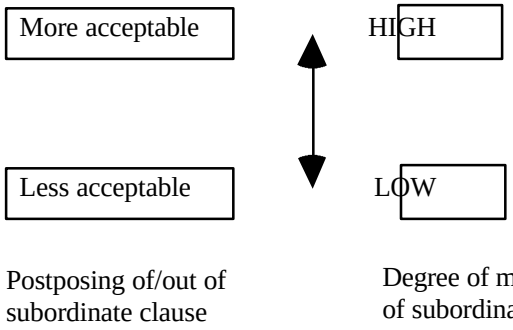


Figure 6: Acceptability of postposing and degree of matrixhood of subordinate clause

I also proposed a principle in terms of the focus structure, which is independent of the hierarchy in Figure 6. This principle accounts for varying degrees of extractability out of subordinate clauses in terms of focus structure. I repeat the principle in (5).

(5) Postposing out of a subordinate clause is unacceptable if the subordinate clause does not carry the focus of the sentence.

In essence, the principle predicts that for an extraction out of a subordinate clause to be acceptable, the extracted element must have a link with a subordinate clause which is the (primary) focus of the sentence. In this sense, despite the difference in degrees of extractability out of subordinate clauses, Japanese is not radically different from languages such as English, since the same sort of principle is applicable to those languages, as seen in Van Valin's (1993a, b) claim repeated in (6).

(6) General restriction on question formation: The element questioned (the WH-word in a simple, direct WH-question or the focal NP in a simple, direct yes-no question) must function in a clause which is within the PFD of the sentence.

The analysis of the postposing construction in Japanese has provided further support for the two additional elements to the traditional layered structure of the clause in RRG, the right-detached position [RDP] and the postcore slot [PoCS]. These elements are particularly useful to represent different types of postverbal elements in the postposing

construction in Japanese. The RDP represents postverbal nonfocus elements which are detached from the clause, i.e. typically *wa*-marked elements in postverbal position. The PoCS represents focus or nonfocus elements which are non-canonically placed in the postverbal position, e.g. *ga*-marked postverbal elements. Figure 7 illustrates the relationship between the constituent projection and the focus structure projection in terms of the four CORE-external elements of the layered structure of the clause.

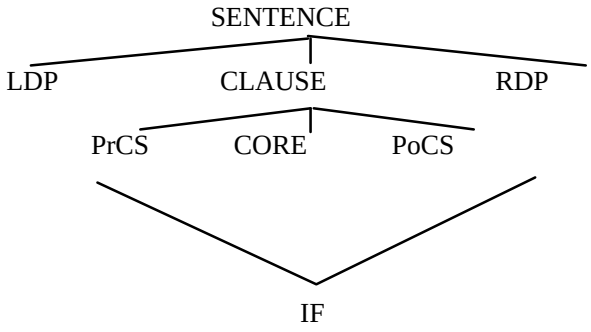


Figure 7: Constituent and focus structure projection
for the CORE-external elements of the LSC

Although the LDP corresponds to the RDP in that they are both outside the focus domain and CLAUSE-external but SENTENCE-internal; however, they contrast with each other in that the RDP represents elements which are typically characterized as afterthoughts in discourse, while it is not the case with the LDP. Similarly, the PrCS corresponds to the PoCS in that they are both within the focus domain and CLAUSE-internal; however, they contrast with each other in that the PrCS holds elements of the primary focus of the sentence, such as WH-phrases in English, while the PoCS typically holds elements of the secondary focus, e.g. *ga*-marked postverbal elements.

Finally, throughout the analyses of some morphosyntactic phenomena in Japanese, this dissertation has revealed a necessary interplay between the structural principles and the functional principles. Although the use of *wa*/*ga* and the postposing construction is tightly connected with the discourse-pragmatic factors, the functional principles alone cannot account for those morphosyntactic phenomena satisfactorily. In this regard, RRG

has provided an ideal theoretical basis for the present study, with its structure-function integration.

REFERENCES

Ando, Sadao. 1986. *Eego no ronri nihongo no ronri*. Tokyo: Taishukan.

Chafe, Wallace. 1974 Language and consciousness. *Language* 50, 111-133.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (ed.), *Subject and Topic*, 27-55. New York: Academic Press.

Chafe, Wallace. 1980. The deployment of consciousness in the production of a narrative. In W. Chafe (ed.), *The Pear Stories*, 9-50. Norwood, New Jersey: Ablex.

Chafe, Wallace. 1982. Integration and involvement in speaking, writing, and oral literature. In D. Tennen (ed.), *Spoken and Written Language*, 35-53. Norwood, New Jersey: Ablex.

Chafe, Wallace. 1984. How people use adverbial clauses. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. 437-449. Berkeley: University of California.

Chafe, Wallace. 1987. Cognitive constraints on information flow. In R. Tomlin (ed.), *Coherence and Grounding in Discourse*, 21-51. Amsterdam: John Benjamins.

Chafe, Wallace. 1994. *Discourse, Consciousness, and Time*. Chicago: University of Chicago Press.

Chomsky, Noam. 1971. Deep structure, surface structure, and semantics interpretation. In D. Steinberg and L Jacobovits (eds.), *Semantics*, 183-216. Cambridge: Cambridge University Press.

Chomsky, Noam. 1973. Conditions on transformations. Reprinted in *Essays on Form and Interpretation*. New York: North-Holland, 1977.

Clancy, Patricia. 1980. Referential choice in English and Japanese narrative discourse. In W. Chafe (ed.), *The Pear Stories*, 127-202. Norwood, New Jersey: Ablex.

Clancy, Patricia. 1982. Written and spoken style in Japanese narratives. In D. Tannen (ed.), *Spoken and Written Language*. Norwood: Ablex.

Clancy, Patricia and Pamela Downing. 1987. The use of *Wa* as a cohesion marker in Japanese oral narratives. In J. Hinds, S. Maynard, and S. Iwasaki (eds.), *Perspectives on Topicalization: The Case of Japanese WA*, 3-56. Amsterdam: John Benjamins.

Clark, Herbert and Susan Haviland. 1977. Comprehension and the given/new contract. In R. Freedle (ed.), *Discourse Production and Comprehension*, 1-40. Norwood, New Jersey: Ablex.

Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.

Dryer, Matthew. 1994. Focus, pragmatic presupposition and activated propositions. To appear in *Journal of Pragmatics*.

Du Bois, John. 1980. Beyond definiteness: The trace of identity in discourse. In W. Chafe (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspect of Narrative Production*. Norwood, 203-274. New Jersey: Ablex.

Erteschik-Shir, Nomi. 1977. *On the Nature of Island Constraints*. Ph.D. dissertation, Massachusetts Institute of Technology.

Erteschik-Shir, Nomi and Shalom Lappin. 1979. Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* 6, 41-85.

Finn, Alice N. 1984. Intonational accompaniments of Japanese morphemes *wa* and *ga*. *Language and Speech* 27, 47-57.

Foley, William and Robert Van Valin. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Fox, Barbara and Sandra Thompson. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66, 2, 297-316.

Fujii, Yoko. 1991. Reversed word order in Japanese: A discourse pragmatic analysis. *Gengo Kenkyu* 99, 58-81.

Giv n, Talmy. (ed.) 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Studies*. Amsterdam: John Benjamin.

Giv n, Talmy. 1983. Topic continuity in discourse: An introduction. In T. Giv n (ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study*, 4-41. Amsterdam: John Benjamins.

Giv n, Talmy. 1984. *Syntax Volume I*. Amsterdam: John Benjamin.

Giv n, Talmy. 1988. The pragmatics of word-order: Predictability, importance and attention. In M. Hammond (ed.), *Studies in Syntactic Typology*, 243-284. Amsterdam: John Benjamins.

Giv n, Talmy. 1989 The grammar of referential coherence as mental processing instructions. *Technical Report No. 89-7*. University of Oregon: Institute of Cognitive and Decision Science.

Greenberg, Joseph. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In J Greenberg (ed.), *Universals of Language*, 73-113. Cambridge, Mass: MIT Press.

Grice, H. P. 1975. Logic and conversation. In P. Cole and J. L. Morgan (eds.), *Syntax and Semantics 3*, 41-58. New York: Academic Press.

Gundel, Jeanette. 1985. Shared knowledge and topicality. *Journal of Pragmatics* 9, 83-107.

Haiman, John. 1983. Iconic and economic motivation. *Language* 59, 781-819.

Halliday, M. A. K. 1967. Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3, 199-244.

Halliday, M. A. K. and Ruqaiya Hasan. 1976. *Cohesion in English*. Longman.

Harada, Shinichi. 1977. Nihongo ni henkei wa hitsuyoo da. *Gengo* 6.

Hasegawa, Yoko. 1989. Questioning vs. identifying: A functionalist analysis of the [a candidate that which professor recommended was hired?] construction in Japanese. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society*. 138-149. Berkeley: University of California.

Hasegawa, Yoko. 1992. *Syntax, Semantics, and Pragmatics of TE-Linkage in Japanese*. Ph.D. dissertation, University of California, Berkeley.

Hinds, John. 1983. Topic continuity in Japanese. In T, Giv n (ed.), *Topic Continuity in Discourse: A Quantitative Cross-language Study*, 43-93. Amsterdam: John Benjamins.

Hinds, John. 1984. Topic maintenance in Japanese narratives and Japanese conversational interaction. *Discourse Processes* 7, 465-482.

Hinds, John and Wako Hinds. 1979. Participant identification in Japanese narrative discourse. In G. Bedell, E. Kobayashi, and M. Muraki (eds.), *Explorations in Linguistics: Papers in Honor of Kazuko Inoue*, 201-212. Tokyo: Kenkyusha.

Hirose, Kyoko and Toshio Ohori. 1992. Japanese internally headed relative clauses revisited. Paper presented at the 1992 Meeting of the Linguistic Society of America.

Hopper, Paul. 1987. Emergent Grammar. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. 139-157. Berkeley: University of California.

Hudson, Mutsuko E. 1993. Processing Constraints on Japanese Postposing. Paper presented at the 4th International Pragmatics Conference, Kobe, Japan.

Inoue, Kazuko. 1978. *Nihongo no bunpo kisoku*. Tokyo: Taishukan.

Inoue, Kazuko. 1980. Hurui joohoo, atarashii joohoo. *Gengo* Oct, 22-34.

Inoue, Kazuko. 1983. *Nihongo no kihon koozoo.* Tokyo: Sanseido.

Inoue, Kazuko. 1984. Bun-bunpoo to danwabunpoo no setten. *Gengo Kenkyu* 84, 17-44.

Ishii, Yasuo. 1989. Head-internal relative clauses in Japanese. *Proceedings of the Fifth Eastern States Conference on Linguistics*. 234-245.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.

Kasuga, Masaharu. 1918. *Jinjoo shoogaku kokugodokuhon no kenkyuu*. Tokyo: Shubunkan.

Keenan, Edward L. and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63-99.

Kimball, John P. 1975. Predicative analysis and over-the-top parsing. In J. Kimball (ed.), *Syntax and Semantics 4*, 155-179. New York: Academic Press.

Kuno, Susumu. 1972. Functional sentence perspective: A case study from Japanese and English. *Linguistic Inquiry* 3, 269-320.

Kuno, Susumu. 1973. *The Structure of the Japanese Language*. Cambridge, Mass: MIT Press.

Kuno, Susumu. 1976. Subject, theme, and the speaker's empathy - a reexamination of relativization phenomena. In C. Li (ed.), *Subject and Topic*, 417-444. New York: Academic Press.

Kuno, Susumu. 1978a. *Danwa no bunpoo*. Tokyo: Taishukan.

Kuno, Susumu. 1978b. Japanese: A characteristic OV language. In W. Lehmann (ed.), *Syntactic Typology: Studies in the Phenomenology of Language*, 57-138. Austin: University of Texas Press.

Kuno, Susumu. 1980. A further note on Tonoike's intra-subjectivization hypothesis. In Y. Otsu and A. Parmer (eds.), *MIT Working Papers in Linguistics 2: Theoretical Issues in Japanese Linguistics*.

Kuroda, Shigeyuki. 1965. *Generative Grammatical Studies in the Japanese Language*. Ph.D. dissertation, Massachusetts Institute of Technology.

Kuroda, Shigeyuki. 1972. The categorical and the thetic judgment. Evidence from Japanese syntax. *Foundations of Language* 9, 153-85.

Kuroda, Shigeyuki. 1975-76. Pivot-independent relativization in Japanese. *Papers in Japanese Linguistics* 4: 85-97.

Kuroda, Shigeyuki. 1976. Headless relatives in modern Japanese and the relevancy condition. *Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society*. 269-279. Berkeley: University of California.

Kuroda, Shigeyuki. 1976-77. Pivot-independent relativization in Japanese. *Papers in Japanese Linguistics* 5: 157-179.

Kuroda, Shigeyuki. 1980. Bunkoozoo no hikaku. In T. Kunihiro (ed.), *Nichieigo Hikaku Kooza 2: Bunpoo*. Tokyo: Taishukan.

Lambrecht, Knud. 1986. *Topic, Focus, and the Grammar of Spoken French*. Ph.D. dissertation, University of California, Berkeley.

Lambrecht, Knud. 1987. Sentence focus, information structure, and the thetic-categorical distinction. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*. 355-82. Berkeley: University of California.

Lambrecht, Knud. 1988. When subjects behave like objects: A markedness analysis of sentence focus constructions across languages. MS. University of Texas, Austin.

Lambrecht, Knud. 1992. Sentence-focus structures as grammatical constructions. Paper presented at the 1992 Meeting of the Linguistic Society of America.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: A Theory of Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Makino, Seiichi. 1982. Japanese grammar and functional grammar. *Lingua* 57, 125-173.

Matsushita, Daizaburo. 1930. *Hyoojun nihon koogohoo*. Tokyo: Kaitakusha.

Maynard, Senko K. 1981. The given/new distinction and the analysis of the Japanese particles WA and GA. *Papers in Linguistics* 14, 1, 109-130.

Maynard, Senko K. 1987. Thematization as a staging device in the Japanese narrative. In J. Hinds, S. Maynard, and S. Iwasaki (eds.), *Perspectives on Topicalization: The Case of Japanese WA*, 57-82. Amsterdam: John Benjamins.

Maynard, Senko K. 1989. *Japanese Conversation: Self-contextualization Through Structure and Interactional Management*. Norwood, New Jersey: Ablex.

Maynard, Senko K. 1990. *An Introduction to Japanese Grammar and Communication Strategies*. Tokyo: The Japan Times.

Maynard, Senko K. 1993. *Kaiwa bunseki*. Tokyo: Kuroshio Shuppan.

McGloin, Naomi H. 1987. The role of *Wa* in negation. In J. Hinds, S. K. Maynard, and S. Iwasaki (eds.), *Perspectives on Topicalization: The Case of Japanese WA*, 165-183. Amsterdam: John Benjamins.

Mikami, Akira. 1963. *Nihongo no ronri - wa to ga*. Tokyo: Kuroshio Shuppan.

Miyagawa, Shigeru. 1987. *Wa* and the WH phrase. In J. Hinds, S. Maynard, and S. Iwasaki (eds.), *Perspectives on Topicalization: The Case of Japanese WA*, 185-217. Amsterdam: John Benjamins.

Mizutani, Nobuko. 1983. Aizuchi to ootoo. In O. Mizutani (ed.), *Hanashikotoba no hyoogen*, 37-44. Tokyo: Chikuma Shoboo.

Muraki, Masatake. 1970. *Presupposition, Pseudo-clefting, and Thematization*. Ph.D. dissertation, University of Texas, Austin.

Myhill, John. 1992. *Typological Discourse Analysis*. Oxford: Blackwell.

Nichols, Johanna. 1986. Head marking and dependent marking grammar. *Language* 62, 56-119.

Nishigauchi, Taisuke. 1984. Japanese LF: Subjacency vs. ECP. In *Seoul Papers On Formal Grammar Theory*.

Ohara, Kyoko Hirose. 1992. On Japanese internally headed relative clauses. *Proceedings of the Eighteenth Annual Meeting of the Berkeley Linguistics Society*, 100-108. Berkeley: University of California.

Ohara, Kyoko Hirose. 1994. Japanese pseudo-relatives: are they internally headed relativization? Paper presented at the 1994 Meeting of the Linguistic Society of America.

Ono, Hideichi. 1973. *Japanese Grammar*. Tokyo: The Hokuseidoo Press.

Payne, Doris L. 1985. *Aspects of the Grammar of Yagua*. Ph.D. dissertation, University of California, Berkeley.

Payne, Doris L. 1990. *The Pragmatics of Word Order: Typological Dimensions of Verb Initial Languages*. New York: Mouton de Gruyter.

Peng, F. C. C. 1977. Josei gengo no chiikisa, nendaisa to kojinsa: Toshika ni yoru ikkoosatsu. In F. C. C. Peng (ed.), *Language and Context*, 73-112. Tokyo: Bunka Hyoron.

Perlmutter, David (ed.) 1983. *Studies in Relational Grammar 1*. Chicago: University of Chicago Press.

Prince, Ellen. 1979. On the given/new information. In *Papers from the Fifteenth Regional Meeting, Chicago Linguistic Society*, 268-278. Chicago: Chicago Linguistic Society.

Prince, Ellen. 1981a. Topicalization, Focus-Movement, and Yiddish-Movement: A pragmatic differentiation. *Proceedings of the Seventh Annual Meeting of the Berkeley Linguistics Society*. 249-264. Berkeley: University of California.

Prince, Ellen. 1981b. Toward a taxonomy of given-new information. In P. Cole (ed.), *Radical Pragmatics*, 223-255. New York: Academic Press.

Prince, Ellen. 1986. Fancy syntax and 'shared knowledge'. *Journal of Pragmatics* 9, 65-81.

Rochemont, Michael. 1986. *Focus in Generative Grammar*. Amsterdam: John Benjamins.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Ph.D. dissertation, Massachusetts Institute of Technology.

Saeki, Tetsuo. 1975. *Gendai nihongo no gojun*. Tokyo: Kasama Shoboo.

Saito, Mamoru. 1985. *Some Asymmetries in Japanese and Their Theoretical Implications*. Ph.D. dissertation, Massachusetts Institute of Technology.

Shibamoto, Janet. 1985. *Japanese Women's Language*. New York: Academic Press.

Shibatani, Masayoshi. 1990. *The Languages of Japan*. Cambridge: Cambridge University Press.

Simon, Mutsuko E. 1989. *An Analysis of the Postposing Construction in Japanese*. Ph.D. dissertation, University of Michigan.

Sunakawa, Yuriko. 1994. Word order in Japanese: Pragmatics of cleft constructions. Paper presented at the 1994 Meeting of the Linguistic Society of America.

Tokieda, Motoki. 1950. *Nihon bunpoo koogohen*. Tokyo: Iwanami.

Uyeno, Tazuko. 1971. *A study of Japanese Modality - A Performative Analysis of Sentence Particles*. Ph.D. dissertation. University of Michigan.

Vallduv , Enric. 1992. *The Informational Component*. New York: Garland.

Van Valin, Robert. 1993a. A synopsis of role and reference grammar. In R. Van Valin (ed.), *Advances in Role and Reference Grammar*, 1-164. Amsterdam: John Benjamins.

Van Valin, Robert. 1993b. The interaction of pragmatics and syntax: A case study in restrictions on question formation, topicalization and relativization. MS. State University of New York at Buffalo.

Van Valin, Robert. 1993c. The lexical representation of verbs and the syntax-semantic interface. Paper presented at the Workshop on thematic Roles at the University of Kansas.

Van Valin, Robert and Randy LaPolla. In press. *Morphosyntax*. Cambridge: Cambridge University Press.

Van Valin, Robert and David Wilkins. 1993. Predicting syntactic structure from semantic representations: *Remember* in English and its equivalents in Mparntwe Arrernte. In

R. Van Valin (ed.), *Advances in Role and Reference Grammar*, 499-534. Amsterdam: John Benjamins.

Watanabe, Noriko. 1990. WA and GA: From the perspective of the deictic center in discourse. In H. Hoji (ed.), *Japanese/Korean Linguistics*, 129-140. Stanford, CA: CSLI.

Watanabe, Yasuko. 1989. *The Function of WA and GA in Japanese Discourse*. Ph.D. dissertation, University of Oregon.

Yang, Byong-Seon. 1994. *Morphosyntactic Phenomena of Korean in Role and Reference Grammar: Psych-verb Constructions, Inflectional Verb Morphemes, Complex Sentences, and Relative Clauses*. Ph.D. dissertation, State University of New York at Buffalo.

Yoshimoto, Kei. 1982. Wa and ga. *Gengo Kenkyu* 81, 1-17.

Zubin, David. 1979. Discourse function of morphology: The focus system in German. *Syntax and Semantics, Volume 12: Discourse and Syntax*, 469-504. New York: Academic Press.

---

[1]Vallduv (1992: 2) uses the terms *informational meaning* and *informational packaging* to refer to this *meaning*, which is not defined by truth-conditional semantics.

[2]As Dryer (1994) notes, the same type of claim has been made in the literature, as in Chafe (1976), Cruttenden (1986), Myhill (1992), Rochemont (1986), and Vallduv (1992). However, they do not make the claim as explicit as Dryer by stating that nonfocus exclusively involves activation and that it does not involve pragmatic presupposition.

[3]Other types of particles include those such as final particles (*shuu-zyoshi*), interjectional particles (*kantoo-zyoshi*), conjunctive particles (*setsuzoku-zyoshi*). See Shibatani (1990) inter alia for the traditional classification of particles in Japanese.

[4]However, Kuno (1973: 94) notes that *ga* is not as acceptable as *o* (the "canonical" object particle) when it marks object of desiderative and potential derivatives of the Sino-Japanese verbs, such as *koonyuusitai* 'want to buy' and *koonyuudekiru* 'be able to buy'.

[5]In the analysis of conversational Japanese in Chapter 3, *subject* here includes object marked by *ga* only for the purpose of the present study. As will

be seen, *ga*-marked subject and object are not distinct from each other in terms of focus patterns of sentence.

[6]For example, Kuno (1972) refers to *predictability* of an entity, while Chafe (1976) appeals to *saliency* for the notions of *newness* or *givenness*.

[7]Kuno (1972: 46) states the use of *wa* here is contrastive. The definition of *contrastive* in Kuno is not necessarily clear, however.

[8]Finn (1984) suggests a tendency on the basis of her speech experiment that the contrastive *wa* gets a larger frequency drop in phrases leading to the contrastive *wa* than the thematic *wa* and also there is more pause after the contrastive *wa* than the thematic one. However, it is not clear whether this result is statistically significant. Also, since the cue sentences were given in isolation to examine the acoustic properties, it is not clear whether the results here can be applied to the occurrence of *wa* in natural discourse.

[9]This condition for the distribution for the two functions of *ga* has been noted by Kuroda (1965).

[10]This function of *ga* has long been noted since Kasuga (1918) and Matsushita (1928).

[11]A sentence can be converted into a question simply with a rising intonation at the end in informal speech. A sentence-final particle such as *no* often accompanies a question without the question marker *ka* in informal speech.

[12]The most natural way to answer to a question such as (41A) is just giving the information asked for, simply saying *Mary* in the case of (41B). This is also the case in English; however, Japanese exhibits a more robust tendency that a speaker does not explicitly utter propositions which can be inferred from the discourse context, as seen in its heavy use of zero anaphora.

[13]An open proposition refers to a proposition which contains a variable argument, i.e. an argument whose referent is unspecified. An active open proposition implies that the referent of the variable argument is unspecified, i.e. nonactive, though the role of the argument, i.e. the agent role of X in *X brought John here today* is active at the time of mention of (41A).

[14]Sunakawa (1994) examined the information structure of cleft constructions on the basis of written Japanese and found dominant patterns in her data that the *wa*-marked cleft constituent is [+accessible] and the following constituent is [-accessible], while the *ga*-marked cleft constituent is [-accessible]

and the following constituent is either [-accessible] or [+accessible]. Sunakawa's definition of *accessibility* is not totally clear; however, her finding supports the discussion here since activated information is accessible and unaccessible information is not activated though not vice versa.

[15]In (61B), there are four possible answers in English as well depending on where the focal accent falls. Namely;

No, Sßlly is British and Ken is Am rican.
No, Sßlly is British and K n is American.
No, Sally is Br tish and K n is American.
No, Sally is Br tish and Ken is Am rican.

[16]Shibatani (1990) makes a similar argument as seen earlier.

[17]Needless to say, the sense of contrast may also be evoked by other factors as well, such as prosodic emphasis, i.e. emphasis on the NP-*wa* phrase in (80.4), contrastive predicate, e.g. the contrastive pairs such as 'cold' vs 'warm'. Here, I am focusing on a single factor, i.e. degrees of activation in competing entities, ignoring those other factors.

[18]The focus types in Lambrecht (1986, 1992, 1994) will be discussed in Section 4.4.

[19]For example, an element can be activated by a concept related to the concept conveyed by the element even though there is no previous mention of the element in the preceding discourse; therefore, the RD measurement does not capture this factor. See the discussion for *accessible to activation* (Dryer 1994) in Chapter 2. Later in this chapter, the notion of *accessible to activation* will be used to account for some cases of *wa* in the database which the RD measurement method fails to capture.

[20]I use the term *predicate* to refer to a whole sentence except the *wa*/*ga*-marked elements; therefore, *predicate* represents an open proposition with a variable for the *wa*/*ga*-marked elements. Since *wa* may mark a non-subject, *predicate* here may refer to a part of a sentence including the subject in such cases. The RD measurement methodology for predicate will be discussed more later.

[21]Maynard (1993: 62), for example, defines *conversation* in her conversation analysis as *daily conversation* which is characterized by the following.

(i) Two or more participants control and carry out the turn taking at the time of conversation rather than following a predetermined conversation turns.
(ii) There is at least one turn taking with meaningful utterances for a given topic.
(iii) The participants engage in simply talking to each other rather than achieving a certain predetermined purpose of the conversation.

[22]RD 20+ refers to RD 20 or over. NPM stands for *no previous mention*, which indicates that there is no antecedent referent in the preceding discourse; namely, an element of RD NPM is the first mention in the discourse.

[23]The use of *ga* in the repeated utterance is similar to the use of the indefinite article in repetition in English. For example:

A: I saw a man there.
B: You saw a man?

In B's utterance, the use of the definite article for *man* is unnatural even though it is mentioned by A in the immediately preceding context. Namely,

A: I saw a man there.
B: # You saw the man?

It is probably the case that the use of *ga* for repeating an utterance with *ga* is simply a copy of the same sentence structure including the particle *ga*, regardless of the activation status of the *ga*-marked elements in the repetition. I am indebted to Matthew Dryer for the English example above.

[24]Maynard (1989, 1993) calls the utterance units identified by pause Pause-bounded Phrasal Unit [PPU].

[25]The use of intonation contour for identification of utterance boundaries is also found in Chafe (1980, 1984), in which intonationally-defined units are called *intonation units*, and Clancy (1982), in which they are called *intonation groups*.

[26]Typically, during the pause which makes syntactically unnatural divisions there is no back-channel response (e.g. "un-huh") by the hearer, which may indicate that the hearer does not regard this kind of *unnatural* pause as an utterance boundary. See Maynard (1993: 97) for a similar observation.

[27]*Te*-form has been reported to be the most frequently occurring connective in Japanese (cf. Saeki 1975, Inoue 1983). Among numerous studies of *te* connective, see Hasegawa (1992) for a recent comprehensive analysis of *te* connective from both a syntactic and a semantic point of view.

[28]Her analyses of the *te* connectives are based on Role and Reference Grammar [RRG] (cf. Van Valin 1993, see the next chapter for the basic concepts of the theory). RRG assumes that a clause has a layered structure with *nucleus*, the innermost layer which contains the predicate (usually a verb), *core* which contains the nucleus and the arguments of the predicate, and *clause* which contains the core, and the *periphery*, which contains non-arguments of the predicate. Accordingly, Hasegawa classifies the *te*-connectives into the three categories: *nuclear* level, *core* level, and *clause* level.

[29]Hasegawa (1992: 191-192) adds one exception to this generalization. Because of Japanese possessing morphological means for delimiting potential subject referents, if a subject refers to a human, a noncoreferential reading is necessary even when either subject or both subjects is not present. For example,

Joan ga setsumee site nattoku nasaimasita
explanation do compliance do:HON:PST
'Joan explained, and (he) understood (it).'

In the sentence above, the honorific form of the verb in the second clause makes the coreferential reading impossible because of the non-honorific form of the verb in the first clause.

[30]Maynard (1989, 1993) calls the gap between speaking turns, in which the filler utterances are found, a *turn-transition state*. The conversation is in this state if the current speaker has yielded the turn and the turn state is open at the moment.

[31]The same kind of peculiar results are found in Watanabe (1989). In her data, 39.6% of all *ga* (59 out of all 149 *ga* tokens) falls on the RD of 1 through 10.

[32]The *wa* tokens here include both *wa*-marked subjects and non-subjects. There is a discrepancy in the total numbers of *wa* and *ga* between the tables in the previous sections and the tables in this section because some *wa*/*ga* utterances in the database are not accompanied by predicate proposition due to the speaker's aborting or repairing the utterance after mentioning the *wa*/*ga* phrase.

[33]Interestingly, the results here may suggest that *wa*-marked subjects and non-subjects exhibit different token distribution patterns for noun-modifying clauses, including both relative clauses and noun complement constructions. There are six *wa*-marked non-subject tokens, while there is no *wa*-marked subject in noun-modifying clauses, which may suggest that *wa* does not mark subjects but may mark non-subjects in noun modifying clauses.

[34]To simplify the token counts, the table divides the RD into the four groups for both *ga*-marked elements and predicates, i.e. the actual token count of *ga* was done in terms of the two domains defined by the RD groups as shown in Table 19.

[35]Frequency of mention is in fact used in quantitative analyses of discourse. Zubin (1979), for example, uses frequency of mention as a measure of the prominence of an entity in the speaker's discourse plan and has found a correlation between frequency of mention and case marking patterns in German. Zubin's finding that a discourse-prominent entity is associated with a certain morphological form, i.e. the nominative case in German, is very insightful for the observation of *wa* marking in the database; the degree of discourse prominence of a referent seems to correlates with the degree of activation, hence, the likelihood of *wa* marking.

[36]I define the term *cohesion* loosely here, following Halliday and Hasan (1976: 4); i.e. "[c]ohesion occurs where the interpretation of some element in the discourse is dependent on that of another. The one presupposes the other, in the sense that it cannot be effectively decoded except by recourse to it."

[37]The temporal expressions discussed here include only those in the subject position of a clause since the non-subject temporal expressions, hence, those which cannot take *ga*, were excluded from the data for the RD measurement, as discussed earlier.

[38]McGloin (1987) examines the role of *wa* in negation and suggests a hierarchy, which is similar to the Keenan-Comrie accessibility hierarchy (cf. Keenan and Comrie 1977), for the accessibility of *wa*-NPs to contrastive interpretation, along which whether *wa* is thematic (outside the scope of negation) or contrastive (focus of negation) is determined. In essence, McGloin claims that the *wa*-marked subject is the most unmarked entity for the thematic reading of *wa* and *wa* which marks a predicate is always contrastive. Although the present study does not use the notion of contrastiveness, it is insightful that McGloin claims that *wa*-marked predicates are distinct from *wa*-marked NPs. Inoue (1984: 37) makes a similar claim that the closer a wa-marked phrase appear to the predicate, the more contrastive the phrase is.

[39]The notion of *grounding* by Fox and Thompson (1990) may be useful to describe the contrast here. *Wa* marking an active proposition makes a nonactive proposition relevant at the point of the discourse at which it is introduced; what is unique to *wa* is its function as a bridge between an active proposition and a nonactive proposition.

[40]The core of Clancy and Downing's (1987) claim here is the function of *wa* which exhibits a semantic contrast between local linguistic elements. The characteristic of *contrastive wa* is not demonstrated by this study, however, as discussed in Section 2.2.2.

[41]The term *verb* here refers to *predicate* in general, including predicate adjectives and the copula, as well as regular verbs.

[42]Also, it is usually the case that an element cannot be postposed out of an embedded clause. This point will be discussed in the next section.

[43]The canonical position of the fronted element is indicated by a gap only for expository purposes. This study does not claim a syntactic operation of moving an element in order to generate the sentence.

[44]Throughout this thesis, elements in postverbal position are underlined for expository purposes.

[45]This does not mean that postposing is constraint-free, however. I will discuss constraints on postposing in the next section and in Chapter 6.

[46]Interestingly, Peng (1977) and Shibamoto (1985) observed that in their databases female speakers used postposing more frequently than male speakers.

[47]In the token counts here, *wa*-marked elements were counted according to the grammatical relations in the sentence, e.g. *wa*-marked subjects were counted as subjects, *wa*-marked objects as objects, etc. *Topic* NP, mostly *wa* marked, refers to elements which have no corresponding gap in the sentence.

[48]Simon says that, as for PPs, a postposition cannot be separated from what it is attached to because it is a bound morpheme. She gives a similar argument for VPs in subordinate clauses; neither V alone nor the whole VP cannot be postposed without the following complementizer because complementizer is a clitic; therefore, it cannot be separated from its host.

[49]In this regard, the constraint is tighter in the case of scrambling. For example, it is not possible to scramble NP2 out of NP1 in the following, which suggests a difference in degree of flexibility between postposing and scrambling.

* Ken ga <u>furansugo no</u> Hanako ni [$_{NP1}$ [$_{NP2}$ ___] [$_{N}$ shoosetsu]] o miseta
French GEN DAT novel OBJ show:PST
'Ken showed a French novel to Hanako.'

[50]Interestingly, Saito (1985: 242) analyzes postposing as different from scrambling by claiming that postposing "should not be analyzed in terms of movement and not even as instances of right-dislocation." See Simon (1989) for her argument against this claim.

[51]The alternative analysis to the postposing movement rule within the formalist theories is that postposed sentences are generated without any movement rule; sentences are generated with the postposed elements in the first place (e.g. in the D-structure) or the postposed elements are added, which is similar to *right-dislocation* but without any resumptive pronoun (e.g. Inoue 1978, Saito 1985). Simon (1989) rejects these non-movement analyses on the basis of her conclusion that the movement analysis gives more descriptive accuracy, explanatory adequacy, and simplicity. Kuroda (1980) suggests a movement analysis of postposing; however, his argument is not as explicit as Simon (1989).

[52]Similarly, Erteschik-Shir (1977) and Erteschik-Shir and Lappin (1979) observe in Danish that question formation is possible out of some relative clauses whose main clause is relatively semantically empty and head noun is non-specific or generic.

[53]Maynard (1989: 37) adds one other explanation for postposing to the list here; namely, the speaker adds elements in the postverbal position that he/she found difficult to specify in the preverbal string due to memory or production difficulties. I will not discuss this particular type any further since it requires a different type of notions.

[54]In the *pause type* of postposing, where there is a noticeable pause preceding the postverbal element, Simon (1989) labels the postverbal elements as *addition*. She suggests two functions of adding the postverbal elements in this type: (i) postverbal elements are used to soften the speech, and (ii) the postverbal elements are used to fill gaps between conversational turns.

[55]Some studies (e.g. Kuno 1978a, Ando 1986) suggest a principle that new information precedes old information not only in postposing constructions but

also in ordering preverbal elements. To my knowledge, however, no previous study directly supports this claim empirically.

[56]*No* is also used with informal questions accompanied by rising intonation, as in (48A). Martin (1975: 916-918) observes that *yo* is "used in asserting a claim, advocating a course of action, or emphasizing a warning," and *ne* and *na* are used to "soften a statement and invite confirmation." See also Uyeno (1971) and Maynard (1990) inter alia for discussion of final particles.

[57]The copula *da* has the corresponding polite form *desu*. It is normally viewed (e.g. Martin 1975, Maynard 1990) that the sentence-final particle *no* and the nominalizer *no*/*n* in the explanatory ending are the same morpheme.

[58]Giv n (1983: 20) claims a similar principle from a typological perspective: "Attend first to the most urgent task". For example, making the comment is more urgent, hence uttered first, than establishing the topic when the topic is more obvious than the comment, while establishing the topic is more urgent when the topic is less obvious than the comment.

[59]Maynard (1990: 258) calls the adverbials as in (ii) and (iii) *openers* and *fillers* respectively. Openers express that the speaker is about to say something, marking a new topic, adding to the current topic, etc. Fillers fill in otherwise awkward gaps in conversation. The appearance of openers in the postverbal position seems contradictory since the elements introducing come after the elements introduced; however, the opener utterances are usually followed by utterances by the same speaker in the data base, i.e. the openers are introduction of the speaker's turn as a whole. It also seems contradictory that fillers occur in the postverbal position immediately following the preceding preverbal utterance without an intervening pause; however, the filler utterance as a whole, including the preverbal string, typically occur filling the pause in the data base, i.e. the filler utterances are typically preceded and followed by a considerable pause in the data base.

[60]Technically, the open proposition *X is board for the landing boats* itself is not actually mentioned in the preceding context; however, both *board* and *landing boats* are both mentioned and it is implied by the context that the landing boats are made of board. So, the element *board for the landing boats* is considered to be more than accessible to activation.

[61]Besides the LDP and the PrCS, later I will posit one additional element for the LSC, the postcore slot [PoCS], which is within the clause but outside the core, like the PrCS. See Chapter 6 for discussion of the PoCS in connection with postposing construction in Japanese.

[62]See Van Valin (1993c) for discussion of the hierarchy and the latest formulation of syntax-semantic interface.

[63]**Do**' in activity verb LSs was not in Dowty's original proposal.

[64]See Van Valin (1993a) and Van Valin and Wilkins (1993) for detailed discussion of the hierarchy and the semantic relations expressed by the nine juncture-nexus types.

[65]The *immediately preceding context of the utterance* here refers to the interval before the utterance but does not include a brief interval just before the utterance. See Section 2.4 for discussion of the three intervals of utterance.

[66]*Ga* in uppercase indicates the prosodic prominence. According to Hasegawa (1992: 39), narrow-focus *ga* "bears a high tone, and/or the NP is followed by a pause." It is not necessarily the case and the prosodic prominence is not always clear in my conversation database, however.

[67]Hasegawa (1992: 39) states "*wa* can bear a high tone with a contrastive NP, but not with a topic NP." As in the distinction between the narrow-focus *ga* and the sentence-focus *ga*, however, this prosodic characteristic of contrastive *wa* is not necessarily clear in my conversation database.

[68]In Hasegawa (1992), LDP is attached to CLAUSE in her representation; however, it is attached to SENTENCE in the latest formulation of RRG (cf. Van Valin 1993a).

[69]The fact that extraction out of an active referent or proposition is blocked is further illustrated by the impossibility of extraction out of definite complex NPs. For example:

* Ken wa sono [dare ga katta] kuruma o untensita no?
that who buy:PST car drive:PST FP
'Who did Ken drive the car which ___ bought?'

With the overt marking of definite, the extraction out of a complex NP results in ungrammaticality. The same sort of impossibility is found in Lakhota (Van Valin 1993b).

wic&ha!s&a! ki [$_{NP}$ [$_{S}$ s&u!4ka wa4 ta!ku yaxta!ke] ki le] wa4ya!4ka he?
man the dog a *what/something bite the this saw Q
*'What did the man see the dog which bit ___ ?'
'Did the man see the dog which bit something?'

With the object NP in the definite restrictive relative clause, the sentence can only be interpreted as a yes-no question, not as a WH-question.

[70]If the preposed phrase *kaigi ni* is marked by *wa*, the sentence is more acceptable, i.e. *kaigi ni wa Taro wa ikimasita.* This tendency applies to (36)-(38) as well.

[71]The canonical order of this sentence is (a); the predicative adposition (non-core argument) *kono hikooki de* precedes the ADV.

[72]This sentence exhibits nuclear juncture, i.e. nuclear subordination. Following the criteria suggested in Hasegawa (1992), it is the nuclear-level juncture because there is no major phrase boundary between *nomi* and *sinakatta*; it is only such particles as *wa* and *mo* 'also' that can appear between the two, and it is subordination because (i) the matrix predicate *nomi* cannot take the negative operator, and (ii) the subordinate predicate *sinakatta* does not change the argument structure of *nomi* 'drink'. On the other hand, there are cases in which *wa* appears within nuclear coordination. For example:

> zutto kuruma o uranaide wa ita keredo kyonen utte simatta
> long-time car OBJ sell:NEG be:PST but last-year sell end-up:PST
> '(I) kept the car for a long time, but (I) ended up selling (it) last year.'

The negative operator *naide* modifies only the first predicate 'sell'; so, *wa* appears between the coordinated conjuncts. I do not discuss the linkage types of *wa*-marked predicates any further since it is beyond the scope of the present study. See Hasegawa (1992) for detailed discussion of the nuclear juncture in Japanese. The crucial point is that, regardless of the juncture type, *wa*-marked predicates are within the IF operator.

[73]I treat *paatii ni* as an oblique core argument here, assuming that 'come' is an *accomplishment* verb.

[74]"The qualifications "simple" and "direct" are meant to indicate that this does not apply to echo questions, rhetorical questions, indirect questions, and other marked question types which have different pragmatic (felicity) conditions on them" (Van Valin 1993a: 158).

[75]Although the complement clause is an argument in the matrix verb's logical structure, it is not syntactically a core argument, as shown in Figure 21. See Foley and Van Valin (1984: 251-255) for discussion as to this point.

[76](56) is from Nishigauchi (1984) and (57) from Hasegawa (1989). Nishigauchi does not consider (56B') as counter-evidence to Subjacency; he simply analyzes that (B') is truncation of (B), which is allowed because 'people' and 'political party' are sufficiently disjoint in reference; therefore, it is easy to recover the identity of the truncated elements. As Hasegawa (1989: 143) argues, however, Nishigauchi's argument above does not account for the fact that (57B') is acceptable. In (57), as Hasegawa notes, that Jeff Chang was a student at the University of California, Berkeley, and the president of the student body at the same time.

[77]*Ba / to / tara / nara* all express conditions; however, they express slightly different meanings. Martin (1975: 552-554), for example, lists the following categories of condition and corresponding morphemes: (1) 'if=provided (that)', expressing a *perquisite* condition: *ba / to*, (2) 'if=if perchance', expressing a *contingent* condition: *tara*, (3) 'whenever, every time that', expressing a *consequential* condition: *ba / to*, (4) 'when (in the past)', expressing a *perfect temporal* condition: *to / tara*, (5) 'when (in the future)' expressing an *imperfect temporal* condition: *tara. Nara* is treated as the same as *ba*; however, Martin (1975: 983) notes that *nara* often expresses condition by "anticipating a specific instance by a general case."

[78]It seems that the two hierarchies here also correlate with the degree of activation of subordinate clauses; the higher the degree of *ga* substitution for *wa* and the lower the degree of matrixhood, the more likely the subordinate clauses to be active.

[79]*Of* is non-predicative if it does not license the argument and it can appear with argument NPs of different semantic functions. For example, the argument is agent in the *attack of the killer bees*, theme in *the gift of a new car*, patient in *the destruction of the city*, etc. (Van Valin and LaPolla, in press: 26).

[80]This phrase is ambiguous as to the relationship between *furansugo no* and the NUC; the phrase can be interpreted as 'book about the French language' as glossed here or as 'book written in French'. In the latter, *furansugo no* is a modifier, rather than an argument.

[81]The same observation has been made by Kuroda (1980).

[82]The postposing here sounds more natural with the emphatic ending *n da yo* than without it, since we are examining the sentence in isolation, without a particular previous discourse context which makes the postposed elements active. (See the acceptability hierarchy for postposing discussed in Chapter 4.)

[83]Hasegawa (1992: 159) notes that the verbal *i-* 'be permitted' is distinct from *i-* 'be good'. The latter can be modified by degree adverbials such as *totemo* 'very', while the former cannot. Also, the negative form of the former is *ike-na-*, while that of the latter is *yoku-na-*.

[84]As discussed above, a *wa*-marked matrix NP can intervene between two CLAUSEs in clausal coordination; however, it changes the reading of the sentence. See (26) for examples.

[85]Martin (1975: 860) classifies *no* here as *post-appositional no*; he subsumes its use under nominalizations.

[86]Furthermore, according to Hirose and Ohori (1992), internally headed relative clauses are also characterized by the following respects: (i) the target of relativization is often difficult to identify, (ii) an internally headed relative clause normally appears as the direct object or the subject of a matrix predicate, not as the indirect object or an adjunct of a matrix clause, and (iii) the relative clause often exhibits an event nominalization reading (e.g. 'Taro knew the fact / event in which Hanako came running.') rather than a relativization reading (e.g. 'Taro knew Hanako who came running.') depending on the matrix predicate.

[87]Yang (1994) observes a contrast between externally headed relative clause and internally headed relative clause in Korean in terms of focus structure. Although WH-question formation is possible out of both types of relative clauses in Korean, as in Japanese, externally headed relative clauses are presupposed, while internally headed relative clauses are not, hence asserted. Given this contrast, Yang proposes that the former is typically outside of the actual focus domain, while the latter is within the actual focus domain.