1   Introduction

It is widely recognized that differences in givenness or newness of information to be conveyed influence the way the information is expressed and distributed in sentences. Likewise, the way the information is expressed by the speaker influences how it is interpreted by the hearer. The relationships between how the information is coded in sentences and the linguistic contexts in which the sentences are used have long been a central concern in studies of form and function in language. Yet aspects of givenness or newness of information that are relevant or irrelevant may be language-specific and may even be specific to a certain construction or usage within a language. This book presents analyses based on the notion of salience of information, one of the aspects of givenness that is related to one's attention or consciousness in local discourse (Chafe 1974; Prince 1981b), and shows how relative salience of information influences the morphosyntactic instantiation of sentences. Furthermore, this book also utilizes the bidimensional model of salience based on Clamons et al. (1993), Mulkern (2003; 2007), and Chiarcos (2009), which captures two independent features of salience: *backward-looking (given or inherent) salience* and *forward-looking (imposed) salience*. The former is related to the topicality of information, a commonly used sense of givenness which is determined by the preceding context. The latter is the speaker's foregrounding of information for the development of the subsequent discourse. The analysis in this book also includes the speaker's intended backgrounding of information, termed *forward-looking (imposed) non-salience*, which is an important aspect of forward-looking salience but has been neglected in previous literature. It will be shown that the analyses of Japanese discourse with respect to a range of morphosyntactic instantiation of information contribute to our understanding of salience of information in the bidimensional model.

The overarching goal of this book also includes discussions of the syntax–semantics–pragmatics interface by bringing together studies of discourse and grammar, and in particular by applying discourse-based observations to Role and Reference Grammar (RRG) (Van Valin and LaPolla 1997; Van Valin 2005). While discourse analyses and studies of syntax have typically been presented separately, this study attempts to break the tradition and combine the

1

two in unified and coherent frameworks. This study utilizes RRG as a primary guiding framework to describe regularities in discourse which are drawn from analyses based on a range of discourse-analytical frameworks. While RRG is a theory of syntax, discourse-pragmatics represents one of the critical components of the theory, which therefore provides an appropriate ground for the goal of creating a unified framework and for further development of the theory at the syntax–pragmatics interface.

The book discusses several Japanese morphosyntactic phenomena centering around topicalization and postposing of arguments and omission of arguments, predicates, and case marking, some of which are characteristic of spoken language specifically. Accordingly, the study addresses observations in a range of discourse types and genres, from written narratives to formal and informal conversational Japanese. Moreover, the goal is not to present the grammar of certain discourse types or genres, or grammatical descriptions of particular discourse data, but to discuss how discourse-based observations are represented in the theory of syntax and how discourse-pragmatics interacts with syntax. Thus regularities observed in the discourse data will be discussed and applied to the coherent framework of grammar. Furthermore, unlike many previous studies of syntax in which the target of analysis is the use of language among native speakers, this book explores the grammar of Japanese as a second or foreign language (henceforth L2 Japanese) in the same framework. This book applies these observations in L2 Japanese discourse to the theory and discusses variations in comparison with the L1 discourse.

One may ask why discourse-pragmatics should be a part of grammar. The answer is rooted in a view of language from the communication-and-cognition perspective (Van Valin and LaPolla 1997: 11), which acknowledges communication as an important function of language. While theories differ within this perspective with respect to how syntax, semantics, and pragmatics relate to each other, RRG claims that these are equally important components of grammar, and discourse-pragmatics is indeed one of the core concerns of the theory. As Van Valin (2005: 1) puts it, one of the central questions to pursue in RRG is, "how can the interaction of syntax, semantics and pragmatics in different grammatical systems best be captured and explained?" The importance of discourse-pragmatics in RRG is closely related to the typological orientation of the theory because there is a great deal of cross-linguistic variation in syntax–pragmatics interface. For example, it is language-specific whether discourse-pragmatics plays a role in the selection of a privileged syntactic argument (PSA) of a grammatical construction.[1] On one hand, there

---

[1] In RRG, privileged syntactic argument is defined as a restricted neutralization of semantic roles for syntactic constructions and it is construction-specific (Van Valin 2005: section 4.2). Traditionally, grammatical relations such as subject and object are used to describe the

are languages, such as Amele, in which there is no pragmatic influence on the selection of PSA. On the other hand, in many languages, including English, discourse prominence of information affects the selection of PSA such that the PSA tends to represent the primary topic of discourse (Van Valin 2005: section 4.3). Japanese is similar in that the selection of PSA is robustly linked with continuity of information (Shimojo 2005) or maintenance of a viewpoint (Nakahama 2003), for example. Discourse influence is also relevant to word ordering. In many languages, word order is influenced by discourse-pragmatic factors such as the focus structure of a sentence. However, in some languages, including English, word ordering is structurally constrained and less constrained by focus structure. In other languages, such as Italian, word order is structurally flexible but constrained by focus structure.[2] Japanese allows flexible word ordering to some extent despite the basic SOV ordering (Shibatani 1990), and non-canonical ordering is sensitive to discourse-pragmatic grounds (Kuno 1995). The flexible word ordering in spoken Japanese includes postpredicate placement of sentence elements, which will be discussed in chapter 6. While post-predicate elements may be either focal or non-focal (Shimojo 1995), these elements exhibit a marked pragmatic function of backgrounding information. In addition, choice of argument forms exemplifies discourse-pragmatic influence. Indefinite NPs are the unmarked form as focus, while zero form is the unmarked form as topic (Van Valin and LaPolla 1997: 205). In Japanese, the zero coding of arguments, which will be discussed in chapters 3 and 4, is the unmarked form used to signal local continuity of information while overt phrases are commonly used as the focus as well as the non-focus of a sentence. Given the pervasive discourse-pragmatic influence on syntax, it is important to capture what role discourse-pragmatics plays in a given language and how it interacts with syntax and semantics. Also, it is important to keep in mind that different constructions of a given language may exhibit different aspects of syntax–pragmatics interface. Therefore we need to describe construction-specific pragmatic features where it is applicable.

The interaction of the three major representations of RRG – syntax, semantics, and discourse-pragmatics – is schematically shown in Figure 1.1.

The syntactic and semantic representations are linked via the bidirectional linking algorithm, which is a set of rules regarding how the two representations are mapped with each other, from semantics to syntax and from syntax to semantics (Van Valin 2005: chapter 5). The linking rules consist of the lexical phase, which represents universal properties, and the syntactic phase, which

---

phenomena, but RRG does not assume that grammatical relations are universal and terms like "subject" and "object" have no theoretical status in RRG.

[2] See Van Valin (1999) for detailed discussion of typology in terms of relative rigidity or flexibility of syntax and focus structure.

SYNTACTIC REPRESENTATION

DISCOURSE-PRAGMATICS

**Linking
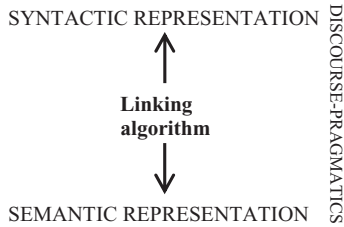algorithm**

SEMANTIC REPRESENTATION

Figure 1.1 Organization of Role and Reference Grammar (Van Valin 2005: 2)

exhibits cross-linguistic variation. One important feature of RRG is that, as shown in the figure, discourse-pragmatics spans syntax and semantics, parallel to the linking algorithm. This indicates the pervasive influence of discourse-pragmatics on the linking and also shows the typological concerns of RRG, because, as mentioned above, there is significant cross-linguistic variation in the interface with discourse-pragmatics and exactly what role it plays is specific to a given language, possibly specific to certain constructions and discourse types, and even specific to L1 and L2 grammars. Given the theoretical background, this book lays out what we observe in discourse with respect to the target phenomena for analysis and then applies the discourse-based observations to RRG. This approach is appropriate because the linguistic phenomena being investigated must first be described before they can be applied to the grammatical theory for explanation.

Because the present study will be based on discourse observations and generalizations, the approach employed in the study aligns with the so-called usage-based approach to grammar to some extent. Researchers who employ this approach share the view that the study of a language system cannot be separated from the study of language use. This is consistent with the basic claim of RRG that language is characterized by the syntax–semantics–pragmatics interface. The approach is characterized by the use of discourse data for the formulation of a linguistic theory, a trend that is becoming increasingly common in linguistic theories (Butler and Gonzálvez-García 2014: 6). In fact, this is consistent with Langacker's (1987: 494) characterization of the usage-based approach – "Substantial importance is given to the actual use of the linguistic system and a speaker's knowledge of this use" – and, as Langacker puts it, this approach assumes that "the grammar is held responsible for a speaker's knowledge of the full range of linguistic conventions" (Langacker 1987, 494). In this study, we assume that a theory of grammar must be able to capture how the language is used, and therefore the discourse reality must be represented properly in the theory of grammar. Thus it is essential to not limit our consideration to the grammaticality of sentences alone and consider discourse regularities as a reflection of the grammar.

On the other hand, it is important to acknowledge that RRG is not a usage-based model in the sense of grammar being the *product* of regularities in discourse or frequency of usage, as claimed in some radical approaches.[3] In RRG, syntax, semantics, and discourse-pragmatics are equally dominant components of grammar and their bidirectional interactions are the core of the theory. For this reason, the goal of this book includes illustration of the bidirectional interface of these components with respect to the Japanese morphosyntactic phenomena discussed.

The discussion will proceed as follows. Chapter 2 discusses the background and key concepts, including salience of information, as well as RRG representations of discourse-pragmatics. The subsequent discussions use the same basic approach consisting of a discourse analysis and an application to RRG. Chapter 3 presents the discourse analysis of zero-coded and topicalized arguments, two representative forms which are associated with salience of information. Chapter 4 applies the results from the discourse analysis to RRG. In chapter 5, the zero coding of verbs is discussed for a certain numeral classifier construction, and the analysis will be extended to the zero coding of predicates in a broader context. Chapter 6 presents an analysis of post-nominal zero marking and post-predicate placement of arguments, which are both associated with non-salience of information in subsequent discourse. The analysis is based on both L1 and L2 data in chapters 3 and 4 (topic and zero coding of arguments) and chapter 5 (zero coding of predicates). Meanwhile, only L1 data are used in chapter 6 (the zero marking and post-predicate placement of arguments).

The inclusion of the analysis of L2 Japanese in this book may appear unusual for the scope of study described above since it falls outside the common realm of second-language (acquisition) research. However, the relevance of inclusion of L2 analysis is twofold: describing and explaining linguistic phenomena, and applying linguistic theory to a broader scope of phenomena. With respect to the former, undoubtedly describing and explaining linguistic phenomena are primary goals of linguistic analysis. Therefore L2 analysis is an attempt to describe and explain L2 linguistic systems, which makes it a natural extension of the study of native speakers' knowledge of their language. This extension leads us to the following general, but fundamental, questions: what is common to L1 and L2? How is L2 usage different from native-speaker usage for the given linguistic phenomena? How are such differences explained? For these goals, analysis of L2 discourse is extremely useful for finding regularities of L2 usage, which may or may not be consistent with the native speaker's system of the same language. In L2 discourse studies,

[3] See Butler and Gonzálvez-García (2014: section 8.2.2) for discussions of RRG in this regard.

L2 usage is often compared with native speakers' usage in the same discourse settings and thereby L2-specific usage is effectively identified and described. This approach is found in a number of studies, including those cited in this book, such as Nakahama (2003; 2011), Polio (1995), and Yanagimachi (2000). It is also advocated by recent development of L1–L2 Japanese-language corpora intended for general and varying research purposes, such as BTSJ Japanese Natural Conversation Corpus (Usami 2020). Describing L2 usage patterns per se has been an important agenda in second-language research regardless of whether or not the findings are used as the basis for drawing theoretical or pedagogical conclusions in the same analysis. While L2 analysis has not drawn as much attention outside second-language research, explicating L2 usage would contribute to our understanding of the nature of L2 linguistic systems and how the "same" language varies across speaker groups. While L2 grammar exhibits divergence from native norms, elucidation of L2 linguistic systems would also confirm that L2 grammar is not random and unpatterned.[4]

The inclusion of L2 analysis is also important for applying linguistic theory to a broader scope of phenomena. This is because the analysis of L2 grammar may serve as the basis for testing the validity of theories and claims based on a native-speaker linguistic system. This approach is justified by the general consideration that a linguistic theory should be able to capture comparable phenomena in different languages in comparable ways, including second languages. For example, this is the primary impetus to universal-grammar (e.g. Chomsky 1981) research in second-language acquisition. As White (2015: 36) puts it, "Given that linguistic theory offers a model of the linguistic competence of native speakers, it may be able to provide a characterization of nonnative competence as well. This is the assumption of researchers working on second language acquisition (SLA) from the perspective of generative linguistics."[5] For the researchers described above, it is important to investigate whether universal grammar, which is claimed to constrain native-speaker grammar, is accessible to L2 learners while they construct the linguistic system for the target language. Obviously, the particular goal stated above is irrelevant to the theoretical orientation employed in this book. However, from the RRG perspective, it is a valid question whether the theory offers a model of L2 speakers' linguistic system, or more broadly communicative competence (see

---

[4] This is not to disregard the fact that learners make nonsystematic errors, including pre-systematic errors (Corder 1973; Brown 2007), which are made because of the learners' unawareness of the target language rules, and mistakes due to a performance failure (Corder 1981). However, with respect to the morphosyntactic phenomena investigated in this book, deviations from native norms collectively observed in the L2 discourse are too systematic to be dismissed as random error.

[5] See also White (2003), inter alia, for discussion of universal grammar in relation to second-language acquisition.

section 2.4.1). While this is an important question to be addressed with regard to RRG, its application to L2 analysis has been largely unaddressed to date, except for limited studies conducted on clause linkage in L2 English and French (see section 2.4). Furthermore, as a parallel-architecture (non-syntacto-centric) theory in which discourse-pragmatics plays a major role (section 2.2), it is also relevant to ask whether RRG offers a model of L2 speakers' discourse-pragmatic knowledge of the target language beyond their linguistic competence.

The L2 analysis presented in this book deals with L2 speakers' topic coding and zero-coding of argument forms (chapters 3 and 4), as well as the zero-coding of predicates (chapter 5), including incomprehensible "incomplete" predicate-less sentences.[6] Accordingly, the L2 analysis provides comparative discussion of L1–L2 usage and, at the same time, the analysis involves theoretical application to the second-language data based on RRG. Thus this book is a demonstration of the relevance of L2 analysis to issues outside traditional second-language research.

---

[6] One may consider these incomprehensible cases as performance failure, i.e. mistakes; however, some of these "incomplete" utterances show little sign of interrupted fluency and none of them are self-corrected by adding missing information afterwards, unlike typical mistakes which are self-corrected (Corder 1981). Further discussion of errors versus mistakes is beyond the scope of the study and the distinction does not affect the claim that the collective L2 usage contrasts with the native norm and how the predicate-less utterances in L2 are described in RRG.