

Introduction

This book is concerned with how structure, meaning and communicative function interact in human languages. Language is a system of communicative social action in which grammatical structures are employed to express meaning in context. While all languages can achieve the same basic communicative ends, different languages use different linguistic means to achieve them, and an important aspect of these differences concerns the divergent ways syntax, semantics and pragmatics interact across languages. For example, the noun phrase referring to the entity being talked about (the ‘topic’) may be signalled by its position in the clause, by its grammatical function, by its morphological case, or by a particle in different languages, and moreover in some languages this marking may have important grammatical consequences and in other languages it may not.

The framework in which this investigation is to be carried out is Role and Reference Grammar [RRG].¹ RRG grew out of an attempt to answer two basic questions, which were originally posed during the mid-1970s: (1) what would linguistic theory look like if it were based on the analysis of languages with diverse structures such as Lakhota, Tagalog and Dyirbal, rather than on the analysis of English?, and (2) how can the interaction of syntax, semantics and pragmatics in different grammatical systems best be captured and explained? Accordingly, RRG has developed typologically motivated descriptive tools and theoretical principles which are designed to expose this interaction and offer explanations for it. It posits three main representations: (1) a representation of the syntactic structure of sentences, which corresponds closely to the actual structural form of utterances, (2) a semantic representation representing important facets of the meaning of linguistic expressions, and (3) a representation of the information (focus) structure of the utterance, which is related to its communicative function. There is a set of rules, called the linking algorithm, which relates the syntactic and semantic representations to each other, and discourse-pragmatics plays a role in the linking. From an RRG perspective, one of the most important ways in which languages differ from each other is in terms of the manner in which discourse-pragmatics interacts with the linking between syntax and semantics. This is summarized in Figure 1.

¹ A bibliography of work in RRG is available on the RRG web site (linguistics.buffalo.edu/research/rrg.html), along with papers, master’s theses and dissertations in PDF format for downloading. It is the best source for the most up-to-date work in the theory.

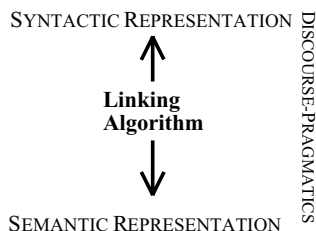


Figure 1 *Organization of Role and Reference Grammar*

RRG seeks to be more than a descriptive framework for the analysis of languages; it also strives to provide an explanatory framework for the analysis of language acquisition and language processing. See Van Valin and LaPolla (1997), ‘Epilog’, and the references cited therein, also Van Valin (1998, 2001a, 2002, 2005), Butler (2004), Weist (2002), and Weist *et al.* (2004).

Thus, this exploration of the syntax, semantics and pragmatics interface will also be an explication of RRG; this presentation presupposes no prior acquaintance with the theory.² The discussion will proceed as follows. The first five chapters deal with simple clauses: in chapter 1 the conception of clause structure proposed in RRG will be detailed; chapter 2 brings in the theory of lexical representation and semantic roles; in chapter 3 the notion of information structure is introduced and related to clause structure; chapter 4 introduces grammatical functions and their interaction with semantic roles and information structure; and chapter 5 presents the algorithm which links syntactic and semantic representations in simple clauses. The next two chapters deal with complex sentences: in chapter 6 the approach to the syntax and semantics of clause linkage is outlined, and in chapter 7 the linking algorithm is expanded to deal with argument linking in complex sentences.

² It does presuppose a basic introduction to syntax, as found in, e.g., Carnie (2002), Tallerman (1998), Van Valin (2001b). See Butler (2003) for a critical introduction to RRG along with a comparison of it with Functional Grammar and Systemic Functional Grammar.

1 Syntactic structure

1.0 General considerations

The first step in the exploration of the syntax, semantics and pragmatics interface in the grammatical systems of human languages is to characterize the nature of syntactic structure, which includes the structure of clauses, of adpositional phrases and noun phrases. There are two fundamental aspects of clause structure which every theory must deal with; they may be termed **RELATIONAL** and **NON-RELATIONAL**. Relational structure is concerned with relations between a predicate and its argument(s); they may be syntactic, semantic, pragmatic or some combination thereof; these relations will be the topics of chapters 2–4. Non-relational structure is concerned with the hierarchical organization of phrases, clauses and sentences; it is the topic of this chapter.

From an RRG point of view, there are two general considerations that a theory of clause structure must meet. They are given in (1.1).

- (1.1) General considerations for a theory of clause structure:
- a. A theory of clause structure should capture all of the universal features of clauses without imposing features on languages in which there is no evidence for them.
 - b. A theory should represent comparable structures in different languages in comparable ways.

These are very strong conditions, especially considering that RRG does not posit any abstract underlying syntactic representations; the syntactic representation of a sentence corresponds closely to its actually occurring form.

1.1 The layered structure of the clause

These considerations lead to a very different conception of clause structure in RRG from that assumed in other approaches. First, because it is concerned with the interplay of syntax, semantics and pragmatics in grammatical systems, the representation of clauses must allow for the representation of all of these factors, where necessary. This renders a representation in terms of purely syntactic features highly inappropriate. Second, the theory is greatly concerned

with typological issues. In particular, it seeks to uncover those aspects of clause structure which are found in all human languages; hence the conception of clause structure it posits must be equally applicable to free-word-order, flat-syntax languages such as Dyirbal and Malayalam, to head-marking languages like Lakota and Tzotzil (see section 1.4), and to fixed-order, configurational, dependent-marking languages like English and Icelandic. It must, further, be applicable to these languages without recourse to positing abstract underlying representations and derivations from abstract to overt representations.

The RRG notion of (non-relational) clause structure is called ‘the layered structure of the clause’ and it is based on two fundamental contrasts: between the predicate and non-predicating elements, on the one hand, and, among the non-predicating elements, between arguments and non-arguments, on the other, i.e. between those NPs and adpositional phrases which are arguments of the predicate and those which are not. These contrasts are found in all languages, regardless of whether they are configurational or non-configurational, head-marking or dependent-marking, free-word-order or fixed-word-order. On this view, the primary constituent units of the clause are the ‘nucleus’, which contains the predicate (usually a verb), the ‘core’, which contains the nucleus and the arguments of the predicate, and a ‘periphery’, which subsumes non-arguments of the predicate, e.g. setting locative and temporal phrases (see section 1.5). This may be informally represented as in the box diagrams in Figures 1.1 and 1.2. The semantic motivation for these units is summarized in Table 1.1. The semantic representations which underlie the notion of core argument will be presented in chapter 2.

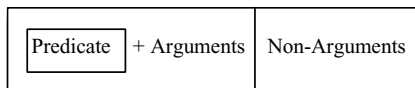


Figure 1.1 *Universal oppositions underlying clause structure*

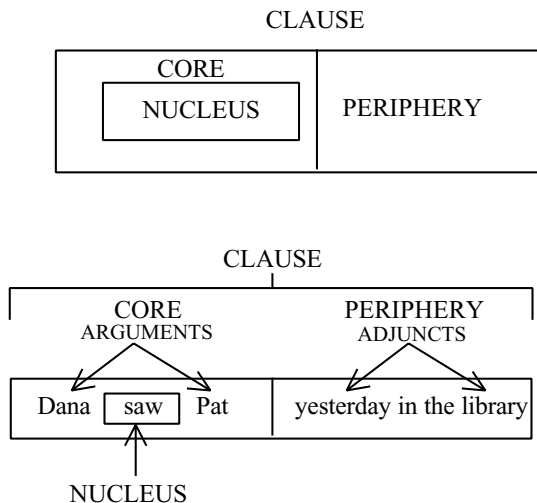


Figure 1.2 *Components of the layered structure of the clause*

Table 1.1 *Semantic units underlying the syntactic units of the layered structure of the clause*

Semantic element(s)	Syntactic unit
Predicate	Nucleus
Argument in semantic representation of predicate	Core argument
Non-arguments	Periphery
Predicate + Arguments	Core
Predicate + Arguments + Non-arguments	Clause (= Core + Periphery)

Since these hierarchical units are defined semantically and not syntactically, they are not dependent upon either immediate dominance or linear precedence relations. Accordingly, the elements in these units may in principle occur in any order, if a given language permits it. This is illustrated in (1.2) from Dyrbal, an Australian Aboriginal language with exceptionally free word order; all of the sentences in (1.2) mean ‘The man speared the wallaby in the mountains.’

- (1.2) Dyrbal (Dixon 1972)
- a. Bayi bargan_{CORE} baŋgul yaɾa-ŋgu_{CORE} ɟurɟa-ɲu_{NUC}
 DET.ABS wallaby-ABS DET.ERG man-ERG spear-TNS
 mountains-LOC
 gambi-ɾa_{PER}.
- b. Baŋgul yaɾaŋgu_{CORE} gambiɾa_{PER} bayi bargan_{CORE} ɟurɟaɲu_{NUC}.
 man mountains wallaby speared
- c. Dyurɟaɲu_{NUC} gambiɾa_{PER} bayi bargan_{CORE} baŋgul yaɾaŋgu_{CORE}.
 speared mountains wallaby man
- d. Bayi bargan_{CORE} gambiɾa_{PER} ɟurɟaɲu_{NUC} baŋgul yaɾaŋgu_{CORE}.
 wallaby mountains speared man
 ‘The man speared the wallaby in the mountains.’

Bayi bargan ‘wallaby (ABS)’ and *baŋgul yaɾaŋgu* ‘man (ERG)’ are arguments in the core of this clause (i.e. ‘core arguments’), regardless of the word order in the clause, because they are arguments of the predicate *ɟurɟa-* ‘spear’, which is the nucleus of the clause. *Gambiɾa* ‘mountains (LOC)’ is not an argument of *ɟurɟa-*, and therefore it is not part of the core; it is, rather, an element in the periphery of the clause. The hierarchical structure of the clause is semantically motivated and not strictly syntactically based.

There are additional elements which may occur in a simple sentence, i.e. a single-clause sentence. The first is the ‘precore slot’ [PrCS], the position in which question words appear in languages in which they do not occur *in situ*, e.g. English, Italian, Zapotec; it is also the location in which the fronted element in a sentence like *BEAN SOUP I can’t stand* appears. This position is clause-internal but core-external. There is also a ‘postcore slot’ [PoCS] in some verb-final languages, e.g.

Japanese (Shimojo 1995), Dhivehi (Indo-Aryan; Cain and Gair 2000). In Dhivehi, both WH-words and non-WH NPs can occur in the postcore slot; they are in small caps.

- (1.3) a. Alī bunī kīKE ta?
 Ali say.PAST.FOC what Q
 ‘What did Ali say?’
- a’. Alī kīke bunī ta?
 Ali what say.PAST.FOC Q
 ‘What did Ali say?’
- b. Māle uḷunīma aharen bonī AIS KURĪMU.
 Male be.PAST.PROG.when 1sg drink.PRES.FOC ice cream
 ‘When in Male, it is ice cream that I eat.’
- b’. Māle uḷunīma ais kurīmu bonī AHAREN.
 Male be.PAST.PROG.when ice cream drink.PRES.FOC 1sg
 ‘When in Male, it is I who eat ice cream.’

In addition to a clause, a simple sentence may also include a phrase in a detached position, most commonly in the ‘left-detached position’ [LDP]. This is the location of sentence-initial elements, most commonly adverbials, which are set off from the clause by a pause, e.g. *YESTERDAY, I bought myself a new car* or *AS FOR JOHN, I haven’t seen him in a couple of weeks*. There is also a ‘right-detached position’ [RDP], as in sentences like *I know them, THOSE BOYS*. When the element in a detached position functions as a semantic argument of the verb, there is normally a resumptive pronoun in the core referring to it.

A language which clearly indicates the precore slot – left-detached position contrast is Tzotzil (Aissen 1987, 1992), a VOS Mayan language. In this language the contrast is marked both intonationally and morphologically. This is exemplified in (1.4).

- (1.4) a. ?i-Ø-s-pet lokel ?antz ti tul-e.
 ASP-3ABS-3ERG-carry away woman DEF rabbit-DEF
 ‘The rabbit carried away the woman.’
- b. Buchu Ø-s-tam?
 who 3ABS-3ERG-take
 ‘Who took it?’
- c. Kusi ch-Ø-a-kan?
 what ASP-3ABS-2ERG-want
 ‘What do you want?’
- d. Vo?ot la ch-a-bat ?un, vo?on la ch-i-kom ?un.
 2sg CL ASP-2ABS-go CL, 1sg CL ASP-1ABS-stay CL
 ‘It’s you who’s going, I’m staying.’
- e. ?a ti tzeb-e, ?i-Ø-s-sa? s-malal.
 TOP DEF girl-DEF ASP-3ABS-3ERG-search 3ERG-husband
 ‘The girl, she searched for her husband.’

Question words occur in immediately preverbal position, as in (b) and (c); there is no pause following them. An NP may also occur in this position, as in (d). This immediately preverbal position is the precore slot. A left-detached position phrase is both set off by a pause and marked by the ‘topic particle’ *?a*, as in (e). The contrast in the discourse functions of phrases in the two positions is illustrated by the following text excerpt (Aissen 1987:158).

- (1.5) [Something had landed at the foot of the tree, they went to look. There was a straw mat. ‘Hell, what could it be? Come on, let’s untie the straw mat!’ the two men said to each other. They untied it. You know what?–]
- a. Tzeb san-andrex la te Ø-s-ta-ik ?un.
 girl San Andreas CL there 3ABS-3ERG-find-pl CL
 ‘A San Andreas girl they found there.’
- b. ?a ti tzeb san-antrex ?un-e, ?i-Ø-y-ik-ik la
 TOP DEF girl San Andreas CL-DEF ASP-3ABS-3ERG-take-pl CL
 ech’el ?un.
 away CL
 ‘The San Andreas girl, they took her with them.’

In (1.5a) *tzeb san-andrex* ‘a San Andreas girl’ is a piece of unexpected new information and is the major focus of the assertion, and consequently it must occur clause-internally in order to be within the domain of the illocutionary force operator. It is preverbal but is not set off by a pause and is not marked by the topic marker *?a*, and accordingly it is in the precore slot. In (b), on the other hand, the same NP is marked by *?a* and set off by a pause, and, in addition, it has already been introduced as a significant participant in the context. Therefore here it is in the left-detached position.

An English sentence containing a precore slot is presented in Figure 1.3. In this representation, an arrow indicates that the periphery is an optional modifier of the core. The periphery will be discussed further in section 1.5. In the sentence in Figure 1.3, *Robin* and *Pat* are arguments within the core, but *Robin* is a direct core argument, while *Pat* is an oblique core argument, because it is marked by a preposition. In languages with case systems, nominative, accusative and dative (or ergative, absolutive and dative) are the direct cases, while the other cases count as oblique.

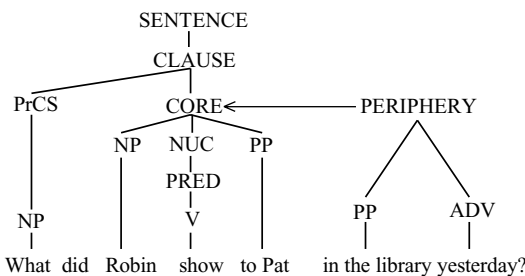


Figure 1.3 *The layered structure of the clause in English*

The RRG conception of the layered structure of the clause is thus a semantically based theory of non-relational syntactic structure; that is, the fundamental units in the hierarchical organization of sentences and clauses are semantically motivated by the contrast between predicate and argument, on the one hand, and that between XPs, i.e. NPs and PPs, which are related to the predicate and those which are not, on the other (see Table 1.1). These units are, however, syntactic units. There are dissociations between the semantic motivations and the syntactic instantiation of these concepts. For example, while the notion of nucleus is based on the semantic notion of predicate, in some languages a nucleus may contain more than the predicate, e.g. an incorporated argument. In Lakhota, the noun *čhá* ‘tree, wood’ can be compounded with the verb *kaksá* ‘chop’ to create *čhákáksa* ‘wood-chop’, a noun + verb combination that can function as the nucleus of a Lakhota clause. The notion of core argument is motivated by the notion of argument in the semantic representation of the verb, but there are clear instances in English and other languages in which this correlation does not hold. For example, the dummy *it* in *it is snowing* occupies a core argument position but is not a semantic argument of *snow*. Furthermore, in a passive construction like *The bagel was eaten by Chris*, *Chris* is a semantic argument of *eat* but occurs as an adjunct in the periphery. Hence while the semantic distinctions in Table 1.1 motivate the syntactic distinctions, there is no absolute correlation between them.

There is an interesting difference between the universal and non-universal aspects of clause structure. The universal aspects (the nucleus, core, periphery and clause) are all semantically motivated, as shown in Table 1.1. The non-universal aspects (the detached phrases, the extra-core slots) are not semantically motivated; rather, they seem to be pragmatically motivated (or at least are associated with constructions that have strong pragmatic conditions on their occurrence). A major difference between RRG and some other theories is that the category of VP, which plays a central role in Chomskyan approaches, has no analogue in the layered structure of the clause.¹

1.2 Operators

In Figure 1.3, the auxiliary verb *did* is not attached to anything, and this is because it is not part of the nucleus, core or periphery. It is, rather, the morphological realization of a tense ‘operator’ which modifies the clause. Grammatical categories like aspect, tense and modality are treated as operators modifying different layers of the clause. Each of the clause levels may be modified by one or more operators. The operators are summarized in Table 1.2. The nuclear operators have scope over the nucleus; they modify the action, event or state itself

¹ The source of VPs in languages which have them will be discussed in section 3.5.

Table 1.2 *Operators in the layered structure of the clause*

Nuclear operators:
Aspect
Negation
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)
Event quantification
Modality (root modals, e.g. ability, permission, obligation)
Internal (narrow scope) negation
Clausal operators:
Status (epistemic modals, external negation)
Tense
Evidentials
Illocutionary force

without reference to the participants. Core operators modify the relation between a core argument, normally the actor, and the action; this is especially true of core directionals and modality. Clausal operators, as the name implies, modify the clause as a whole. They fall into two groups, one containing tense and status, and the other evidentials and illocutionary force. Tense and status situate the proposition expressed by the clause within temporal and realis–irrealis continua; evidentials indicate the epistemological basis of the state of affairs (the proposition plus tense and status operators) expressed, i.e. how the speaker came to have the information being uttered, while illocutionary force specifies the type of speech act. Hence evidentials and illocutionary force are modifiers of the sentence or utterance as a whole, rather than one of its constituent clauses; they are thus ‘sentential’ in nature. They occur only in main (root) clauses, i.e. clauses which are immediately dominated by the sentence node. Negation is the only operator that occurs at all three levels: nuclear negation has only the nucleus in its scope, core negation has one or more core arguments (and possibly also the nucleus) in its scope, and clausal negation has the entire clause in its scope. The classification of a particular operator as nuclear, core or clausal is a direct function of its meaning. No language need have all of these operators as grammatical categories; for example, English, unlike Kewa and Quechua, does not have evidentials as a grammatical category. The only operators which every language has are illocutionary force and negation.

Examples of operators from a variety of languages are given in (1.6)–(1.12). The Kewa examples in (1.6) involve aspect, nuclear directionals (nuclear because they modify the motion of the action, not a participant), and tense. The Turkish examples in (1.7) illustrate aspect and tense in (a) and modality, status and tense

in (b). Both Kewa and Turkish are verb-final, left-branching languages, and consequently the operator-bearing morphemes follow the verb; this is also the case in Barasano in (1.10) and Amele in (1.12). English and Tiwi, which are verb-medial, right-branching languages, exhibit the opposite arrangement: the operator-bearing morphemes precede the verb, as in (1.8) and (1.11). Note that the order of tense, aspect and the verb in the Tiwi example in (1.11) is the mirror image of the order in the Kewa, Turkish and Barasano examples.

- (1.6) Kewa (Papua-New Guinea; Franklin 1971)
- a. Íra-pa-niaa-ru.
 cook-PERF-down-1sgPAST (V-ASP-DIR-TNS)
 ‘I burned it downward (as a hill).’
- b. Íra-pa-saa-ru.
 cook-PERF-up-1sgPAST (V-ASP-DIR-TNS)
 ‘I burned it upward (as a hill).’
- (1.7) Turkish (Watters 1993)
- a. Gel-iyor-du-m.
 come-PROG-PAST-1sg (V-ASP-TNS)
 ‘I was coming.’
- b. Gel-emi-yebil-ir-im.
 come-ABLE.NEG-PSBL-AOR-1sg (V-MOD-STA-TNS)
 ‘I may be unable to come.’
- (1.8) English
- a. He may be leaving soon. (IF/TNS-STA-ASP-V)²
- b. She was able to see them. (IF/TNS-MOD-V)
- c. Will they have to be leaving? (IF/TNS-MOD-ASP-V)
- (1.9) Korean (Yang 1994)
- An-tul-li-wu-(si)-lswuiss-cianh-ass-keyss-up-nita.
 NEG-hear-CAUS-PASS-(SH)-ABLE-NEG-PAST-PRESUM-POL-DEC
 (NEG-V-MOD-NEG-TNS-EVID-IF)
 ‘(I) guess that (he) [HON] might not be heard.’
- (1.10) Barasano (Tucano, South America; Jones and Jones 1991)
- Gahe-rîbî bota-ri kēã-kudi-ka-bã ìdã.
 other-day post-pl chop-ITER-FPAST-VIS 3pl (V-ASP-TNS-EVID)
 ‘The next day they went from place to place chopping down posts.’
- (1.11) Tiwi (Australia; Osborne 1974)
- ŋə-ru-uŋt̪iŋ-apa.
 1sg-PAST-PROG-eat (TNS-ASP-V)
 ‘I was eating.’

² The reason that illocutionary force and tense are linked in the English examples is that, in English, illocutionary force is indicated by the position of the tense marker in the main clause: interrogative by core-initial tense, declarative by core-internal tense, and imperative by no tense.