Oxford Handbooks Online

Role and Reference Grammar as a Framework for Linguistic Analysis

Robert D. Van Valin Jr.

The Oxford Handbook of Linguistic Analysis (2 ed.)
Edited by Bernd Heine and Heiko Narrog

Print Publication Date: Feb 2015 Subject: Linguistics, Morphology and Syntax
Online Publication Date: Jul 2015 DOI: 10.1093/oxfordhb/9780199677078.013.0028

Abstract and Keywords

Role and Reference Grammar is a typologically grounded theory of grammar which strives to be both useful for linguists interested primarily in language description and for linguists concerned with the central theoretical issues in the field. This chapter presents the basics of the theory in the context of the questions that motivated it and of its implications for cognitive science.

Keywords: linking, lexical decomposition, grammatical relations, information structure, semantics, complex sentences, typology, endocentrism

30.1 Introduction: Motivation, Goals, and Evidence

THERE are many motivations for proposing and developing a theoretical framework in linguistics. For example, a leading idea in the development of Generalized Phrase Structure Grammar (Gazdar et al. 1985) was to determine whether natural language syntax could be adequately described in terms of a context-free phrase structure grammar. Lexical-Functional Grammar [LFG] (Bresnan 1982*a*, 2001) had a number of motivations, including applying the formalism of unification grammar to natural language phenomena and showing that lexical rules were superior to transformational rules. Construction Grammar (Fillmore 1988; Goldberg 2006) emerged as a reaction to Chomsky's claim that constructions are mere epiphenomena derived from the interaction of more basic general rules and principles. Role and Reference Grammar [RRG] was inspired by both typological and theoretical concerns. The motivating questions for RRG were, "what would a linguistic theory look like if it were based on the analysis of

languages with diverse structures, such as Lakhota, Tagalog, Dyirbal, and Barai (Papua New Guinea), rather than on the analysis of English?", and "how can the interaction of syntax, semantics, and pragmatics in different grammatical systems best be captured and explained?" These two questions contain both theoretical and descriptive content. On the one hand, they both emphasize the importance of taking account of typologically diverse languages in the formulation of a linguistic theory, and, on the other, they indicate that the resulting theory will be one in which semantics (p. 708) and pragmatics play significant roles. One of the reasons that some of the constructs posited by RRG are rather different from those in other theories is precisely because of this starting point. For example, theories starting from English and other familiar Indo-European languages often take the notion of subject for granted, whereas for one that starts from syntactically ergative and Philippine languages, this is not the case, and the notion of subject as a theoretical construct is called seriously into question. Indeed, the initial work in RRG concerned the question of the universality of subject and the cross-linguistic validity of grammatical relations in general (Foley and Van Valin 1977, 1984; Van Valin 1977*a*, 1981).

Since the late 1980s there have been two additional questions that RRG seeks to answer: "can language acquisition be accounted for without recourse to an autonomous Language Acquisition Device?", and "can a model of grammar that answers the typological and theoretical questions posed above provide any insights into the neurocognitive processing of language?" The final chapter of Van Valin and LaPolla (1997) (henceforth VVLP) is devoted to the first question, and the tentative conclusion is that the acquisition of quite a few core grammatical phenomena can be explained in RRG terms, including some for which it has been argued that there is no evidence available to the child regarding them, e.g., subjacency.[1] In the last few years there has been increasing work on applying RRG to language processing, both in computational and neurolinguistic terms. Computational implementation of RRG is in its infancy (Kailuweit et al. 2003; Butler 2004; Nolan 2004; Guest and Brown 2007; Guest 2008), but the results so far are very promising; in particular, Guest (2008) has developed a parser based on RRG and has used it to successfully parse a large corpus of English sentences as well as a small corpus of Dyirbal sentences, including ones with discontinuous constituents. With respect to sentence processing, one of the distinctive features of RRG, to be discussed in more detail below, is the bidirectionality of the mapping between syntax and semantics. The RRG linking algorithm maps from semantics to syntax and from syntax to semantics. This is an idealization of what a speaker does (semantics to syntax) and what a hearer does (syntax to semantics). Hence the design of the theory makes it readily amenable to psycho- and neurolinguistic sentence processing models. Recently there has been experimental neurolinguistic research involving RRG as the grammar component of a sentence comprehension model (Bornkessel, Schlesewesky and Van Valin 2004; Bornkessel and Schlesewsky 2006*a*, 2006*b*; Van Valin 2006; Bornkessel-Schlesewsky and Schlesewsky 2008). The success and explanatory power of this model can be taken as support for the RRG linking system.

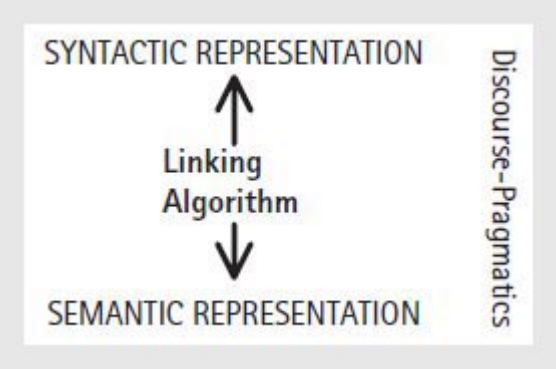


Figure 30.1 The organization of RRG

An important issue with respect to the questions RRG seeks to answer is the nature of the evidence that can be brought to bear on them, and in this regard the theory is particularly eclectic. Obviously, data from grammatical descriptions from a wide range of languages is central to the evaluation of analyses and theoretical claims. However, other types of evidence are used as well. The analysis of a number of grammatical phenomena has been influenced by evidence from language acquisition and (p. 709) neurolinguistic experiments. Sociolinguistic and conversational data have also played a role in RRG analyses; the data investigated in Belloro (2007), for example, include the results of a sociolinguistic survey of use of clitic pronouns in Buenos Aires Spanish, and Shimojo (1995) presents an account of *wa* and *ga* in Japanese based on the analysis of transcripts of Japanese TV talk shows. His 2005 book explores argument encoding in Japanese conversation. Thus, RRG takes a variety of evidence into account.

The vast majority of work in RRG has been synchronically oriented, although as far back as Foley and Van Valin (1984) and Ohori (1992) it was argued that there were diachronic implications of RRG analyses. In the last few years there has been more diachronic work, e.g., Matasović (2002), Wiemer (2004), Eschenberg (2004, 2005), and the most recent results are presented in a collection of papers on applying RRG to issues in language change and grammaticalization (Kailuweit et al. 2008).

The presentation of RRG will proceed as follows. In the next section, the organization of theory will be discussed, and the basics of the representations posited in the theory will be introduced. In the subsequent section, the linking system will be described and applied to some much discussed grammatical phenomena. The final section contains a brief summary.

30.2 Organization and Representations

The organization of RRG is given in Figure 30.1; it will be elaborated further in § 30.3.

There is a direct mapping between the semantic and syntactic representations of a sentence, unmediated by any kind of abstract syntactic representations; this excludes not only derivational representations, as in, for example, the Minimalist Program, but also the use of abstract f-structures as in LFG. There is only a single syntactic representation for a sentence, and it corresponds to the actual form of the sentence. RRG does not allow any

phonologically null elements in the syntax; if there's nothing there, there's nothing there.[2] RRG posits a very concrete syntactic representation, and this (p. 710) constrains the theory significantly; as noted above, this rules out derivations and there-with movement rules, however they are formulated, and also abstract representations like relational networks in Relational Grammar or f-structures in LFG. Many of the descriptive and theoretical devices in theories with abstract syntactic representations are not available in RRG. The syntactic representation will be described in §30.2.1.

The syntactic representation is linked via the linking algorithm (§30.3) to the semantic representation. It consists of a lexical decomposition representation of the meaning of the predicator along with its arguments. This will be discussed further in §30.2.2.

The last component in Figure 30.1 is labeled "discourse-pragmatics", and it is parallel to the linking algorithm. What this depicts is the fact that discourse-pragmatics, primarily as realized in information structure, plays a role in the linking between syntax and semantics. Crucially, however, exactly what role it plays can vary across languages, and this variation is the source of important cross-linguistic differences among languages. Information structure and its representation will be introduced in §30.2.3.

There is no RRG-related theory of phonology. Work on an RRG theory of morphology is in its initial stages (e.g. Everett 2002; Martin Arista 2008; Van Valin 2012), and the representation of the internal structure of words would be part of the syntactic representation. The role of the lexicon will be discussed in §30.2.2 and §30.3.

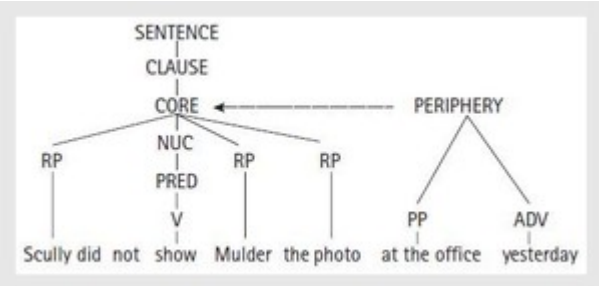
30.2.1 The Syntactic Representation of a Sentence

The RRG theory of syntactic representation strives to satisfy the two conditions in (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.



Click to view larger

Figure 30.2 The layered structure of the clause in English

Clause structure is not represented in RRG in terms of X-bar syntax or even traditional immediate constituency structure;[3] rather, it is captured in a semantically-based model known as the "layered structure of the clause". The essential components of this model of the clause are (i) the NUCLEUS, which contains the predicate, (ii) the CORE, which contains the nucleus plus the arguments of the predicate in the nucleus, and (iii) a PERIPHERY for each layer, which contains adjunct modifiers. These aspects of the layered structure are universal. The structure of a simple English clause is given in Figure 30.2; the structure in Figure 30.2 is the constituent projection of the clause.[4] (p. 711) In Table 30.1 the semantic units underlying the layered structure of the clause are summarized.

Table 30.1 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)

The distinctions in Table 30.1 are universal and follow from the fact that language is used to refer and predicate. There is no verb phrase in the layered structure because it is not universal. Hence the layered structure of the clause meets the condition in (1a).

The predicate in the nucleus need not be a head, nor is it restricted to a particular category. While the most common category for the predicate in the nucleus is undoubtedly verb, adjectives, as in *Pat is **tall*** (*be* is analyzed as a tense-carrying auxiliary, not as a predicate), nominal phrases, as in *Mary is **a very good lawyer***, and adpositional phrases, as in *Sam is **at the office***, can all serve as the predicate in the nucleus. Hence the RRG theory of constituent structure is non-endocentric.[5]

Some languages have a "pre-core slot", which is the position of WH-words in languages like English and Icelandic, and a "left-detached position", which is the (p. 712) position of the pre-clausal element in a left-dislocation construction. In addition, some verb-final languages have a "post-core slot" (e.g., Japanese; Shimojo 1995), and some languages

also have a “right-detached position”, which is the position of the post-clausal element in a right-dislocation construction.

Table 30.2 Operators

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

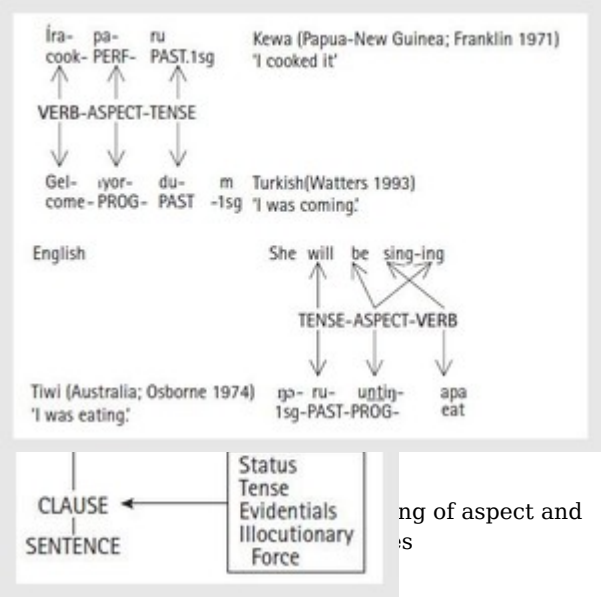
A second important component of the RRG theory of clause structure is the theory of OPERATORS. Operators are closed-class grammatical categories like aspect, negation, tense, and illocutionary force. An important property of operators is that they modify specific layers of the clause. This is summarized in Table 30.2.

Languages normally do not have all of these operators as grammatical categories; the absolutely universal ones are illocutionary force and negation. Operators are represented in a separate projection of the clause, which is the mirror image of the constituent projection. This is exemplified in Figure 30.3.

An example of an English sentence with constituent and operator projections is given in Figure 30.4.

The sentence in Figure 30.4 involves a left-detached position as well as a pre-core slot housing a WH-expression. Note that there is no empty argument position in the core corresponding to the WH-word in the PrCS; see fn. 2. In this example, *did* is labeled both “tense” and “IF” in the operator projection, because the position of the tense operator signals illocutionary force in English: core-medial tense signals declarative IF, pre-core

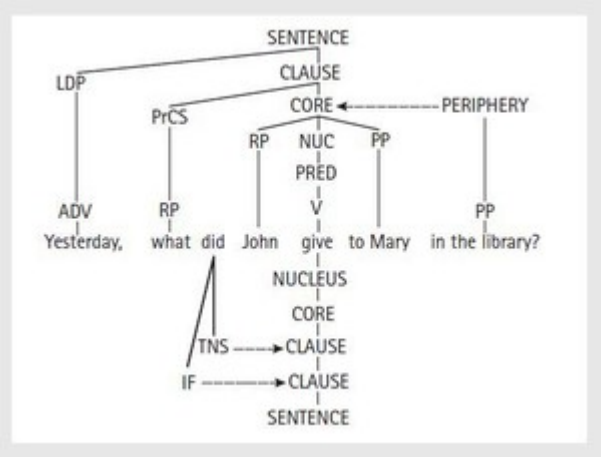
tense signals interrogative IF, and the absence of tense in a matrix core signals imperative IF.



ng of aspect and
s

Click to view larger

Figure 30.3 The operator projection in the layered structure of the clause

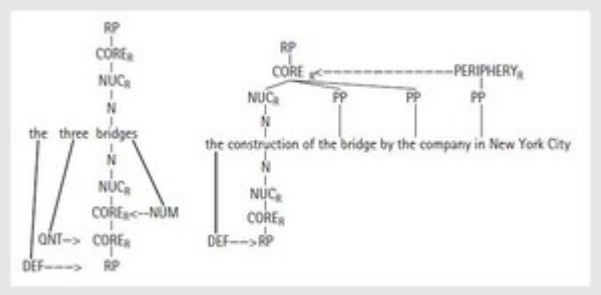


Click to view larger

Figure 30.4 An English sentence with both constituent and operator projections

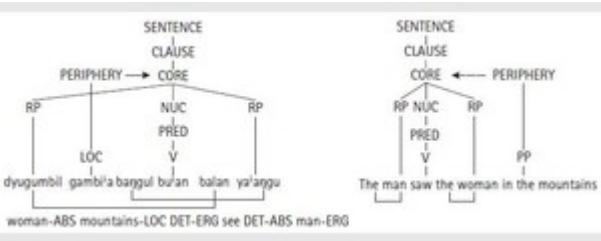
An important claim made in Foley and Van Valin (1984) was that the linear order of the morphemes expressing the operators is a function of their scope.[6] That is, morphemes expressing nuclear operators appear closer to the nucleus than morphemes (p. 713) (p. 714) expressing core operators, and these in turn occur closer to the nucleus than mor-phemes expressing clausal operators, when an ordering relationship among them can be established, i.e., they all occur on the same side of the nucleus. This is illustrated for two left-branching and two right-branching languages in Figure 30.5.

Other types of phrases also have a layered structure analogous to the clause, and they may have both constituent and operator projections, as appropriate. In Figures 30.2 and 30.4 the nominal phrases are labeled “RP” instead of “NP”. “RP” stands for “reference phrase”, and unlike “NP” but like the clause, it is a non-endocentric construct. The nucleus of an RP is neither restricted to nominals, nor is it restricted to lexical heads. The first point is particularly important, given the much discussed issues raised by languages like Nootka (Swadesh 1939; Jakobsen 1979) and Tagalog (Schachter 1985; Himmelmann 2008), in which the heads of referring expressions need not be nominal in nature. See Van Valin (2008*a*) for detailed discussion.



Click to view larger

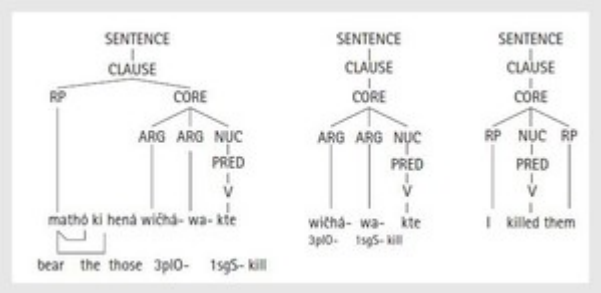
Figure 30.6 The layered structure of the RP with operator projection



Click to view larger

Figure 30.7 The layered structure of the clause in Dyirbal and English

The idea of a layered structure is applied to other categories as well, especially RPs and PPs. RPs have both constituent and operator projections, with the operator projection containing categories such as definiteness, deixis, quantification, and number. Examples of RPs and their layered structure are given in Figure 30.6. This approach to the structure of RPs makes possible an analysis of discontinuous constituency that satisfies principle (1b). The Dyirbal and English sentences in Figure 30.7 are translations of each other, and the Dyirbal example involves discontinuous RPs. The lines connecting the determiners to the head nouns are the operator projection within the RP, as in Figure 30.6. In head-marking languages like Lakhota, the bound argument (p. 715) markers on the verb are considered to be the core arguments; overt RPs are within the clause in apposition to them (Van Valin 1977*b*, 1985, 2012a). Hence, in the first diagram in Figure 30.8, which means “I killed the bears”, the bound argument markers plus the nucleus, all one phonological word, constitute the core of the clause, while the independent RP is clause-internal (it can be the focus of a question, and therefore is within the scope of the IF operator, namely the clause). In the second one, which means “I killed them”, the verb word constitutes the core, clause, and sentence by itself. It contrasts with its English translation only in terms of the order of morphemes and the fact that the English pronominals are free rather than bound morphemes.



Click to view larger

Figure 30.8 The LSC in Lakhota (head-marking) and English (dependent-marking)

Note that despite the differences between the three languages in Figures 30.7 and 30.8, comparable structural relations, e.g., core argument, peripheral adjunct, are represented in the same way. (p. 716)

Representations of constituent projections such as these are analyzed as "syntactic templates", the inventory of which in a language constitutes an important component of its grammar. It is termed the "syntactic inventory" and complements the lexicon. There are template selection principles, based on the semantic representation of a sentence, that determine the selection of the proper template(s). The syntactic structure for a sentence may be made up of a combination of multiple templates.

30.2.2 The semantic representation of a sentence

The semantic representation of a sentence is based on the lexical representation of the verb or other predicating element. It is a decompositional representation based on Vendler's (1967) theory of *Aktionsart*. The four basic classes (state, achievement, accomplishment, and activity) are augmented by two additional classes, semelfactives (punctual events; Smith 1997) and active accomplishments (telic uses of activity verbs, e.g., *devour, run to the store*); in addition, there are causative versions of each. Examples of the six classes are given in (2), and sentences illustrating the classes plus their causative counterparts are given in (3).

(2)

a. States: *be sick, be tall, be dead, love, know, believe, have*
b. Activities: *march, swim, walk* (— goal PP); *think, eat* (+ mass noun/bare plural RP)
c. Semelfactives: *flash, tap, burst* (the intransitive versions), *glimpse*
d. Achievements: *pop, explode, shatter* (all intransitive)
e. Accomplishments: *melt, freeze, dry* (the intransitive versions), *learn*
f. Active accomplishments: *walk* (+ goal PP), *eat* (+ quantified RP), *devour*

(p. 717)

(3)

a.	State:	The boy fears the dog.
a′.	Causative state:	The dog frightens/scares the boy.
b.	Achievement:	The balloon popped.
b′.	Causative achievement:	The cat popped the balloon.
c.	Semelfactive	The light flashed.
c′.	Causative semelfactive	The conductor flashed the light.
d.	Accomplishment:	The ice melted.
d′.	Causative accomplishment:	The hot water melted the ice.
e.	Activity:	The dog walked in the park.
e′.	Causative activity:	The girl walked the dog in the park.
f.	Active accomplishment	The dog walked to the park.
f′.	Causative active accomplishment:	The girl walked the dog to the park.

Syntactic and semantic tests determine the *Aktionsart* of a clause (see VVLP §3.2.1; Van Valin (2005) (henceforth VV_{05}); §2.1.1). As the sentences in (3e–f t) show, a single verb, e.g., *walk*, can have more than one *Aktionsart* interpretation. This verb would be listed in the lexicon as an activity verb, and lexical rules would derive the other uses from the basic activity use (see VVLP §4.6; Van Valin 2012b).

The system of lexical decomposition builds on the one proposed in Dowty (1979). Unlike Dowty's scheme, the RRG system treats both state and activity predicates as basic. The lexical representation of a verb or other predicate is termed its LOGICAL STRUCTURE [LS]. State predicates are represented simply as **predicatet**', while all activity predicates contain **do**'. Accomplishments, which are durative, are distinguished from achievements, which are punctual. Accomplishment LSs contain BECOME, while achievement LSs contain INGR, which is short for "ingressive". Semelfactives contain SEML. In addition, causation is treated as an independent parameter which crosscuts the six *Aktionsart* classes, hence the twelve classes in (3). It is represented by CAUSE in LSs. The lexical representations for each type of verb in (3) are given in Table 30.3. (p. 718)

Table 30.3 Lexical representations for *Aktionsart* classes

Verb Class	**Logical Structure**
STATE	**predicate**′ (x) or (x, y)
ACTIVITY	**do**′ (x, [**predicate**′ (x) or (x, y)])
ACHIEVEMENT	INGR **predicate**′ (x) or (x, y), or

	INGR **do**′ (x, [**predicate**′ (x) or (x, y)])
SEMELFACTIVE	SEML **predicate**′ (x) or (x, y), or
	SEML **do**′ (x, [**predicate**′ (x) or (x, y)])
ACCOMPLISHMENT	BECOME **predicate**′ (x) or (x, y), or
	BECOME **do**′ (x, [**predicate**′ (x) or (x, y)])
ACTIVE ACCOMPLISHMENT	**do**′ (x, [**predicate**′$_1$ (x, (y))]) & BECOME **predicate**′$_2$ (z, x) or (y)
CAUSATIVE	α CAUSE β, where α, β are LSs of any type

Examples of simple English sentences with the LS of the predicate are presented in (4).

(4)

a. STATES	
Leon is a fool.	be′ (Leon, [fool′])
The window is shattered.	shattered′ (window)
Fred is at the house.	be-at′ (house, Fred)
John saw the picture.	see′ (John, picture)
b. ACTIVITIES	
The children cried.	do′ (children, [cry′ (children)])
The wheel squeaks.	do′ (wheel, [squeak′ (wheel)])
Carl ate snails.	do′ (Carl, [eat′ (Carl, snails)])
c. SEMELFACTIVES	
The light flashed.	SEML do′ (light, [flash′ (light)])
John glimpsed Mary.	SEML see′ (John, Mary)
d. ACHIEVEMENTS	
The window shattered.	INGR shattered′ (window)
The balloon popped.	INGR popped′ (balloon)
e. ACCOMPLISHMENTS	
The snow melted.	BECOME melted′ (snow)
The sky reddened.	BECOME red′ (sky)
Mary learned French.	BECOME know′ (Mary, French)
f. ACTIVE ACCOMPLISHMENTS	
Carl ate the snail.	do′ (Carl, [eat′ (Carl, snail)]) & BECOME eaten′ (snail)
Paul ran to the store.	do′ (Paul, [run′ (Paul)]) & BECOME be-at′ (store, Paul)
g. CAUSATIVES	
The dog scared the boy.	[do′ (dog, Ø)] CAUSE [feel′ (boy, [afraid′])]
The sun melted the ice.	[do′ (sun, Ø)] CAUSE [BECOME melted′ (ice)]
The cat popped the balloon.	[do′ (cat, Ø)] CAUSE [INGR popped′ (balloon)]
Bill flashed the light.	[do′ (Bill, Ø)] CAUSE [SEML do′ (light, [flash′ (light)])]
Felix bounced the ball.	[do′ (Felix, Ø)] CAUSE [do′ (ball, [bounce′ (ball)])]
The girl walked the dog to the park.	[do′ (girl, Ø)] CAUSE [do′ (dog, [walk′ (dog)]) & BECOME be-at′ (park, dog)]

Full semantic representations of sentences also contain lexical representations of the RPs, adjuncts, and grammatical operators like tense and aspect; see VVLP §4.4, 4.7; VV_{05} §2.2–2.3.

(p. 719) 30.2.2.1 Semantic Macroroles and Lexical Entries for Verbs

The semantic interpretation of an argument is a function of its position in the LS of the predicate, and, as will be seen below, the linking system refers to an element's LS position. Thematic relations as such play no role in the theory; the traditional thematic role labels are used only as mnemonics for the LS argument positions, e.g., "theme" is the mnemonic for the second position (y) in a two-place locational LS like **be-at**' (x, y). RRG posits two generalized semantic roles or SEMANTIC MACROROLES, which play a crucial role in the linking system. The two macroroles are ACTOR and UNDERGOER, and they are the two primary arguments of a transitive predication; the single argument of an intransitive predicate can be either an actor or an undergoer, depending upon the semantic properties of the predicate. The basic distinction is illustrated in the following German examples.

(5)

a. Der Junge [SUBJ, ACTOR] hat den Kuchen [OBJ, UNDERGOER] aufgegessen.
'The boy [SUBJ, ACTOR] ate the cake. [OBJ, UNDERGOER]'

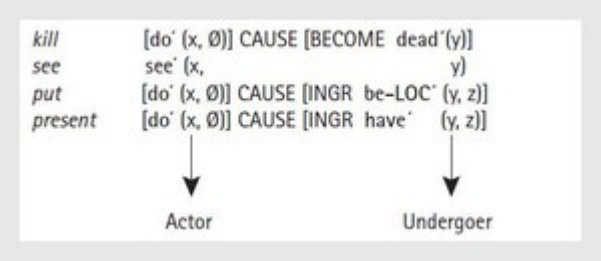
b. Der Hund [SUBJ, ACTOR] ist um das Haus herumgelaufen.
'The dog [SUBJ, ACTOR] ran around the house.'

c. Der Hund [SUBJ, UNDERGOER] ist gestorben.
'The dog [SUBJ, UNDERGOER] died.'

d. Der Kuchen [SUBJ, UNDERGOER] wurde vom Jungen [ACTOR] aufgegessen.
'The cake [SUBJ, UNDERGOER] was eaten by the boy [ACTOR].'

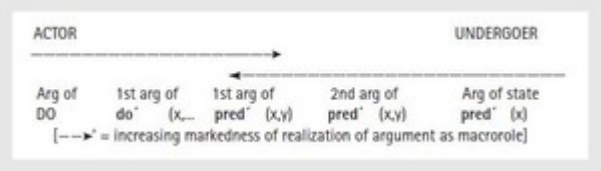
In (5a), *der Junge* "the boy" is the actor and *den Kuchen* "the cake" is the undergoer of the transitive verb *aufessen* "eat up"; in the sentences with intransitive verbs, *der Hund* is an actor with the activity verb *herumlaufen* "run around" and an undergoer with the accomplishment verb *sterben* "die". Actor is not equivalent to syntactic subject, nor is undergoer equivalent to syntactic direct object, as the examples in (5c) and crucially (5d) show: in both of these sentences the syntactic subject is an undergoer, and in the passive sentence in (5d) the actor is an oblique adjunct. In an English clause with an active voice transitive verb, the actor is the initial RP (the traditional subject) and the undergoer, when it occurs, is always the direct RP immediately following the verb. In an English passive construction, the undergoer is the subject and the actor, if it occurs, is in an adjunct PP in the $periphery_{CORE}$.

Actor and undergoer are generalizations across specific semantic argument types, as defined by LS positions. This is illustrated in Figure 30.9.



Click to view larger

Figure 30.9 Macroroles as generalizations over specific thematic relations



Click to view larger

Figure 30.10 The Actor–Undergoer Hierarchy[7]

The *x* argument of all of these verbs functions as the actor, regardless of whether it is the first argument of the generalized activity verb **do**′ (conventionally labeled "effector"), as with *kill, put* and *present*, or the first argument of a two-place state predicate, as with *see*. With two-place transitive verbs like *kill* and *see*, the *y* argument is the undergoer. With three-place verbs like *put* and *present* (as in *Bill presented Mary* (p. 720) *with the flowers*), on the other hand, the situation is potentially more complex, and this will be discussed in §§ 30.3 and 30.4.

The relationship between LS argument positions and macroroles is captured in the Actor–Undergoer Hierarchy [AUH] in Figure 30.10. The basic idea of the AUH is that in a LS the leftmost argument in terms of the hierarchy will be the actor and the rightmost will be the undergoer. This was true for *kill, see*, and *put* in Figure 30.9. It was not true for *present*, however, and this reflects a fundamental asymmetry in the AUH: the leftmost argument in a LS (in terms of the AUH) is always the actor, but the rightmost argument is only the default choice for undergoer. This possible variation in the selection of the undergoer is the basis of the RRG analysis of dative shift and related phenomena (see §30.3).

Transitivity in RRG is defined semantically in terms of the number of macroroles a predicate takes. This is termed "M-transitivity" in RRG, following Narasimhan (p. 721) (1998), in order to distinguish it from the number of syntactic arguments a predicate takes, its "S-transitivity". The three M-transitivity possibilities are: transitive (2 macroroles), intransitive (1 macrorole), and atransitive (0 macroroles). It is important to point out in the context of this discussion of three-place predicates that there is no third macrorole; there is nothing in RRG corresponding to Primus' (1999) notion of "proto-recipient". From theoretical and empirical perspectives, there are no grounds for positing a third macrorole; see Van Valin (2004), VV_{05}: 64–6, for detailed discussion). The theoretical label for the third argument in a ditransitive predication; e.g., *the picture*, in the English sentence *Sam showed Sally the picture*, is "non-macrorole direct core argument".

The principles determining the M-transitivity of verbs are given in (6).

(6)
Default Macrorole Assignment Principles

a. Number: the number of macroroles a verb takes is less than or equal to the number of arguments in its LS.
   1. If a verb has two or more arguments in its LS, it will take two macroroles.
   2. If a verb has one argument in its LS, it will take one macrorole.
b. Nature: for predicates which have one macrorole,
   1. If the verb LS contains an activity predicate, the macrorole is actor.
   2. If the predicate has no activity predicate in its LS, it is undergoer.

If a verb is irregular and has exceptional transitivity, it will be indicated in its lexical entry by "[MRα]", where "α" is a variable for the number of macroroles. Examples of lexical entries for some English verbs are given in (7).

	a.	*kill*	[do′ (x, Ø)] CAUSE [BECOME dead′ (y)]
	b.	*receive*	BECOME have′ (x, y)
	c.	*own*	have′ (x, y)
	d.	*belong (to)*	have′ (x, y) [MR1]
(7)	e.	*see*	see′ (x, y)
	f.	*watch*	do′ (x, [see′ (x, y)])
	g.	*show*	[do′ (w, Ø)] CAUSE [BECOME see′ (x, y)]
	h.	*run*	do′ (x, [run′ (x)])
	i.	*drink*	do′ (x, [drink′ (x, y)])

A major claim in RRG is that no syntactic subcategorization information of any kind is required in the lexical entries for verbs. For regular verbs, all that is required is the LS and nothing more, as in all except (7d). For most irregular verbs, only the macrorole number needs to be specified. The prepositions that mark oblique arguments with verbs like *show* are predictable from general principles and need not be listed in the lexical entry (see below, also Jolly 1993; VVLP §7.3.2). All of the major morphosyntactic (p. 722) properties of verbs and other predicates follow from their LS together with the linking system.

30.2.3 The Information Structure Representation of a Sentence

The morphosyntactic means for expressing the discourse-pragmatic status of elements in a sentence is called "focus structure", and the approach to focus structure used in RRG is based on Lambrecht (1994). He proposes that there are recurring patterns of the organization of information across languages, which he calls "focus types". The three

main types are presented in (8), with data from English and Italian; focal stress is indicated by all caps.

(8) Focus structure in English and Italian (Lambrecht 1994)

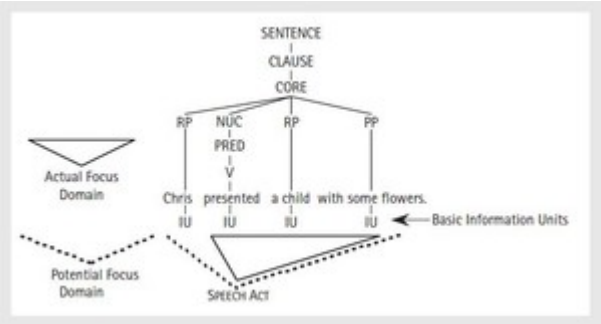
a. Q:	What happened to your car?	Predicate Focus
A:	i. My car/It broke DOWN.	English
	ii. (La mia macchina) si è ROTTA.	Italian
b. Q:	What happened?	Sentence Focus
A:	i. My CAR broke down.	English
	ii. Mi si è rotta la MACCHINA.	Italian
c. Q:	I heard your motorcycle broke down.	Narrow Focus
A:	i. My CAR broke down.	English
	ii. Si è rotta la mia MACCHINA./ È la mia MACCHINA che si è rotta.	Italian (Lit: 'broke dow my car'/'it's my car which broke down')

Predicate focus corresponds to the traditional topic-comment distinction, with a topical subject RP and a focal predicate phrase which receives the focal stress. It is universally the least marked or default focus structure. In English, the subject would most likely be an unstressed pronoun, while in Italian it would most likely not occur at all; if it were overt, it would be preverbal in Italian. Sentence focus is a topicless construction in which the entire sentence is focal. In English, the subject receives the focal stress, while in Italian the subject appears postverbally and with focal stress. Narrow focus involves focus on a single constituent, in these examples, the subject. In English this is signaled by focal stress on the element or by a cleft, e.g., *It was my* CAR *that broke down*. Italian likewise has two options: postposing the subject, when it is the focused element, or a cleft.

There is an important distinction between unmarked and marked narrow focus. All languages have an unmarked focus position in the clause; in English it is the last constituent of the core, whereas in verb-final languages it is the position immediately before the verb. Consider the following English sentence with different focal stress options. (p. 723)

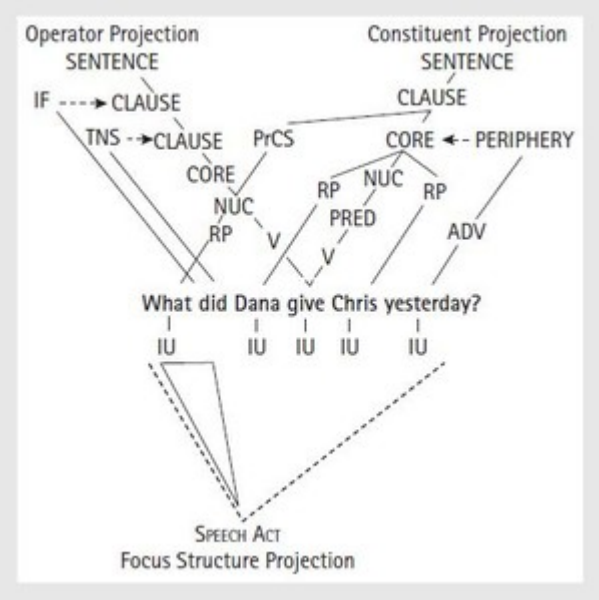
(9)
a. Dana sent the package to LESLIE yesterday.
b. Dana sent the package to Leslie YESTERDAY.
c. Dana sent THE PACKAGE to Leslie yesterday.
d. Dana SENT the package to Leslie yesterday.
e. DANA sent the package to Leslie yesterday.

Focal stress on *Leslie* in (a) is a case of unmarked narrow focus, while focal stress on any other constituent of the clause, as in (b)–(e), yields marked narrow focus. The most marked narrow focus is on the subject, as in (e).



Click to view larger

Figure 30.11 Focus structure projection of an English predicate-focus construction



Click to view larger

Figure 30.12 English sentence with all three projections represented

Information structure is represented by an additional projection of the clause, the focus structure projection. It is illustrated in Figure 30.11 for a predicate focus construction in English. There are three main components of this projection. Basic information units correspond to the information content captured by a simple WH-word like *who*, *what*, or *where*. In simple sentences this notion may seem redundant with syntactic phrases, but it plays an important role in the analysis of information structure in complex sentences. The second component is the actual focus domain, which is what is actually in focus in a given context; the elements in small caps in (9) are in the actual focus domain in those examples. The third component, which was introduced in RRG and is not part of Lambrecht's original account, is the potential focus domain. Languages differ as to constraints on where the actual focus domain can be in a clause. In some like English, it can fall on any word or phrase, as (9) shows. In others, e.g., Italian, it is excluded from the preverbal core position and can only include the nucleus and what follows (see VVLP §5.4, Van Valin 1999 for detailed discussion). The potential focus domain is a feature of the grammar of the language, while the actual focus domain is contextually determined. In Van Valin (2005) formal (p. 724) representations of context based on Discourse Representation Theory (Kamp and Reyle 1993; von Heusinger 1999) are incorporated into the theory, in order to derive the different focus types. They

can also play an important role in linking in some languages (see Van Valin 2005, §5.4.1; Shimojo 2008, 2009). A very new development in the theory is an explicit representation of prosody (O'Connor 2008), which will be added as an additional projection.

It is possible to represent all three projections in a single tree, as in Figure 30.12. It should be noted that these are not three separate representations of the sentence; rather, they are representations of three types of information which are simultaneously present in the sentence.

30.2.4 Grammatical Relations

As noted in §30.1, in the earliest work on RRG it was argued that grammatical relations like subject and direct object are not universal and cannot be taken as the basis for adequate grammatical theories. In place of these notions, RRG employs the notion of "privileged syntactic argument" [PSA], which is a construction-specific relation and is defined as a restricted neutralization of semantic roles and pragmatic functions for syntactic purposes. The other arguments in a clause are characterized as direct or oblique core arguments; there is nothing in RRG corresponding to direct or indirect object (see Van Valin 2005, chapter 4).

(p. 725) Languages have selection hierarchies to determine the PSA; the two main ones are given in (11).

(10)
Privileged syntactic argument selection hierarchy:
Arg of DO > 1st arg of do′ > 1st arg of pred′ (x,y) > 2nd arg of pred′ (x,y) > pred′ (x)

(11)
Privileged Syntactic Argument Selection Principles

a. Accusative construction: Highest ranking direct core argument in terms of (10)-default
b. Ergative constructions: Lowest ranking direct core argument in terms of (10)-default
c. Restrictions on PSA in terms of macrorole status:
   1. Languages in which only macrorole arguments can be PSA: German, Italian, Dyirbal, Jakaltek, Sama, etc.
   2. Languages in which non-macrorole direct core arguments can be PSA: Icelandic, Georgian, Japanese, Korean, Kinyarwanda, etc.

The PSA selection hierarchy in (10) is the actor part of the AUH. For a language like English, (11a) captures the fact that, in an active voice clause with a transitive verb, the actor is the PSA, whereas for a language like Dyirbal, in an active voice clause with a transitive verb the undergoer is the PSA, following (11b). These are the default choices; it

is possible for an undergoer to serve as PSA in a passive construction in an accusative language like English or German, and it is likewise possible for an actor to serve as PSA in an antipassive construction in syntactically ergative languages like Dyirbal and Sama (Philippines; Walton 1986). Languages also differ with respect to whether the PSA must be a macrorole: German, Italian, Dyirbal, Jakaltak (Mayan), and Sama restrict PSA selection to actors and undergoers only, while Icelandic, Georgian, Japanese, and Kinyarwanda allow non-macrorole direct core arguments to function as PSA (see VVLP §7.3.1.1; Van Valin 2005 §4.2).

An aspect of (11a) with significant typological consequences is whether it is a default rule, as in English, German, and many other languages, or whether it is an absolute rule, as in Lakhota, Warlpiri, and many other languages. That is, in a language like Lakhota the highest-ranking argument in the LS is always the PSA; there is no other choice, as the language lacks a voice opposition. With a transitive or other multi-argument verb, the speaker has no choice as to which argument serves as the PSA. The contrast between PSAs in English-type languages and Lakhota-type languages can be captured in a distinction between "variable PSAs" [most English constructions] vs. "invariable PSAs" [all Lakhota constructions].[8] In languages with variable PSAs, in (p. 726) particular constructions, e.g., those like (12) below, one of the factors affecting which argument may be selected as PSA is information structure. It has long been recognized that there is a strong tendency for the RP selected as PSA to be the most topical in the particular context. In RRG variable PSAs in such constructions are analyzed as "pragmatically-influenced PSAs"; it should be noted that not all variable PSAa are pragmatically influenced. There is no pragmatic influence on PSA selection with invariable PSAs.

Two typological points need to be mentioned. First, invariable ergative PSAs, i.e., (11b), are extremely rare but not non-existent. Second, the most common type of PSA cross-linguistically seems to be invariable accusative PSAs. VVLP §6.5 presents extensive arguments for both of these claims and proposes an explanation for these facts.

PSAs may be characterized functionally as controllers or pivots. These two functions are exemplified in (12) and (13).

(12)

a. The tall man$_i$ hit William$_j$ and then $\_\_{}_{i/*j}$ ran away.
CONTROLLER PIVOT

b. William$_j$ was hit by the tall man$_i$ and then $\_\_\_{}_{*i/j}$ ran away.
CONTROLLER PIVOT

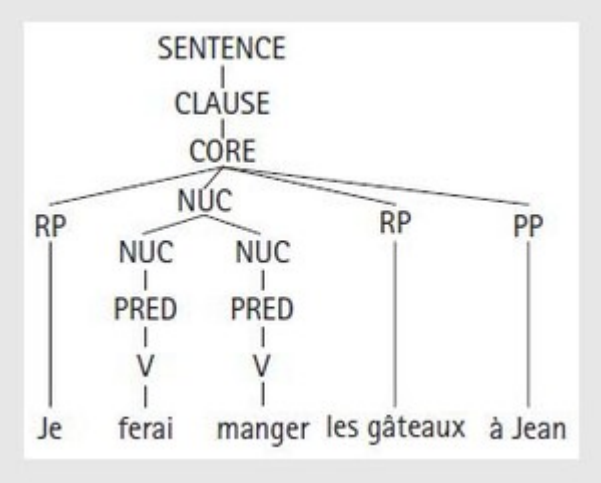
(13)

a. Bill$_i$ persuaded the tall man$_j$ [$\_\_\_{}_{*i/j}$ to visit Leslie].
CONTROLLER PIVOT

b. The tall man$_j$ was persuaded by Bill$_i$ [$\_\_\_{}_{*i/j}$ to visit Leslie].
CONTROLLER PIVOT

Pivots are canonically the missing argument in a construction, as in (12) and (13), while controllers prototypically supply the interpretation for a pivot. It should be noted that there can be pivots without controllers, e.g., the extracted element in an extraction

construction, and controllers without pivots, e.g., reflexive controllers. A further contrast is highlighted in these examples, the contrast between syntactic and semantic pivots and controllers. In the construction in (12), the controller is the first RP in the core, the traditional "subject", regardless of its semantic function, whereas in the construction in (13), the controller is the undergoer argument, regardless of its syntactic status. Hence the controller in (12) is a syntactic controller, while the controller in (13) is a semantic controller. The types of pivots and controllers that the constructions of a language have are typologically very significant.



Click to view larger

Figure 30.13 The structure of (14a)

The RRG position on the universality of grammatical relations can be summarized as follows. For a language to have grammatical relations in the usual sense of syntactic relations which are not reducible to semantic roles, it must have at least one construction with a syntactic pivot and/or a syntactic controller. There are languages, e.g., Acehnese (Durie 1985, 1987), which have only semantic pivots and controllers, and therefore they lack grammatical relations in the purely syntactic sense. Furthermore, in languages with syntactic pivots and controllers, there is variation in terms of whether they pattern accusatively, ergatively, or some other pattern, and whether they have variable and invariable PSAs or only invariable PSAs. The grammatical relations in the (p. 727) former type of language are not the same as those in the latter. Hence even among the vast majority of languages with syntactic pivots and controllers, there is no uniformity as to the nature of the PSAs.

30.2.5 The Structure of Complex Sentences

The three central components of the LSC also turn out to be the three fundamental building blocks of complex sentences in human language. The unmarked pattern for the construction of complex sentences involves combining nuclei with nuclei, cores with cores, clauses with clauses, or sentences with sentences. These are called levels of "juncture" in RRG, i.e., nuclear juncture, core juncture, clausal juncture, and sentential juncture. Sentential junctures are complex constructions made up of multiple sentences, while clausal junctures involve sentences containing multiple clauses. Examples of nuclear junctures from French, English, and Mandarin are given in (14) and the

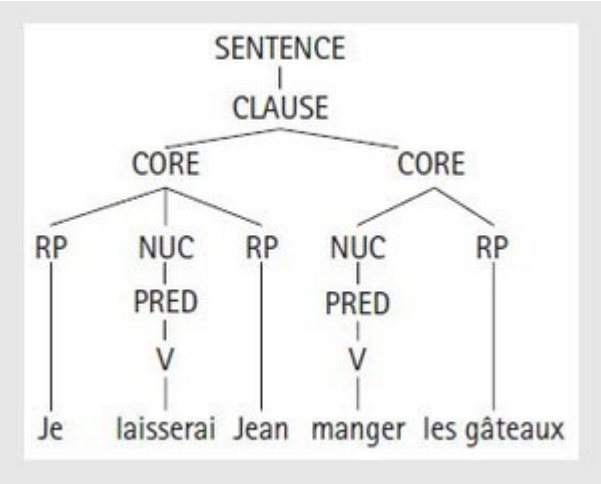
representation of (14a) is in Figure 30.13. Justifications for these structures can be found in Van Valin (2005).

(14)

a. Je ferai manger les gâteaux à Jean.
1sg make.FUT eat the cakes to John
'I will make John eat the cakes.'
[two nuclei, *faire* and *manger*, in a single core]

b. John forced open the door.
[two nuclei, *force* and *open*, in a single core]

c. Tā qiāo pò le yí ge fànwǎn.
3sg hit 'break' PRFV one CL bowl
'He broke (by hitting) a ricebowl.'
[two nuclei, *qiāo* 'hit' and *pò* 'break', in a single core] (Hansell 1993)



Click to view larger

Figure 30.14 The structure of (15a)

Core junctures involve two or more cores (which may themselves be internally complex) in a clause. Examples from French, English, and Mandarin are given in (15), (p. 728) and the structure of (15a) is presented in Figure 30.14. In this type of core juncture, the two cores share a core argument; "sharing a core argument" is defined formally in terms of the linking algorithm mapping syntactic and semantic representations into each other.

(15)

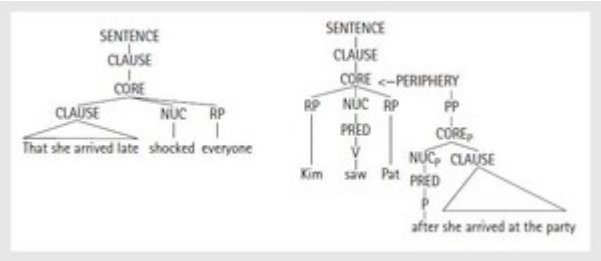
a. Je laisserai Jean manger les gâteaux.
1sg let.FUT John eat the cakes
'I will let John eat the cakes.'

b. I ordered Fred to force the door open.

c. Tā jiāo wǒ xǐe zì.
3sg teach 1sg write characters
'She teaches me to write characters.'

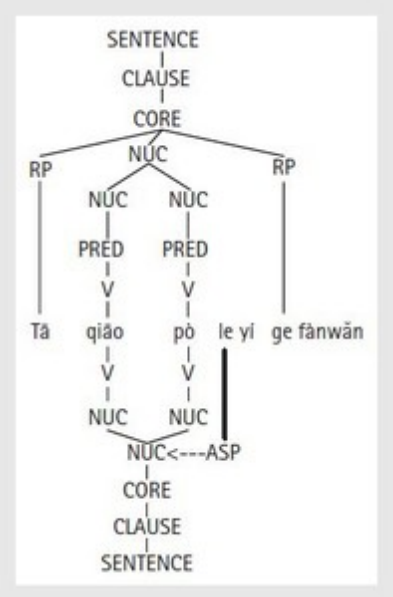
Of equal importance in the RRG theory of complex sentences is the set of possible syntactic and semantic relations between the units in a juncture; the semantic relations are discussed below. The syntactic relations between units are called “nexus” relations in RRG. Traditionally, only two basic nexus relations are recognized, coordination and subordination. Subordination is divided into two subtypes, daughter subordination and peripheral subordination. They are illustrated in Figure 30.15.

The embedded clause in the first sentence is a daughter of the core node, while in the second the embedded clause is an adjunct in the periphery modifying the core.



Click to view larger

Figure 30.15 Daughter and peripheral subordination at the core level in English



Click to view larger

Figure 30.16 Nuclear cosubordination in Mandarin

In addition to distinguishing two types of subordination, RRG, following Olson’s (1981) analysis of clause linkage in Barai (a Papuan language), posits a third nexus type: “cosubordination”, which is essentially tight, dependent coordination. The dependence is operator dependence; that is, in cosubordination, the units obligatorily share one or more operators at the level of juncture. In the Mandarin example in (14c), aspect obligatorily has scope over both nuclei, and therefore the nexus is cosubordination. This is represented as in Figure 30.16. (p. 729)

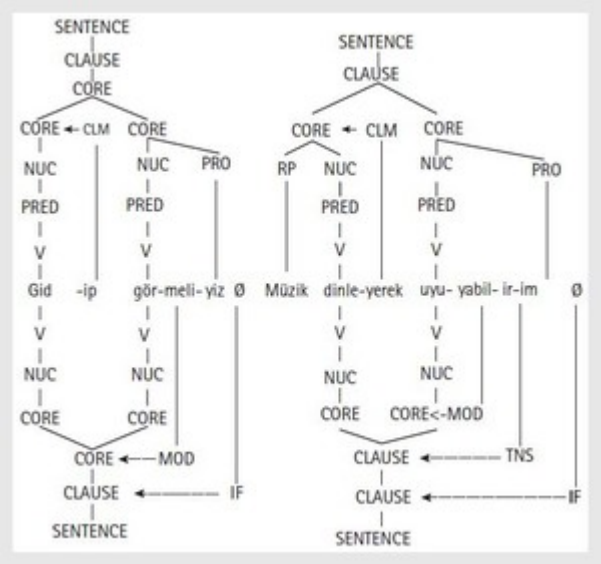
The following examples from Turkish (Watters 1993) illustrate obligatory operator sharing and the lack of it in Turkish core cosubordination and coordination, respectively. The term “coordination” here is being used for an abstract linkage relation referring to a relationship of equivalence and operator independence at the level of juncture. It is distinct from conjunction, which is a construction type of the general form “X conj Y”, which may be one of the formal instantiations of coordinate nexus. (p. 730)

a. Core cosubordination
Gid-ip gör-meli-yiz.
go-CMPL see-MODAL-1pl
'We ought to go and see.'

(16) b. Core coordination
Müzik dinle-yerek, uyu-yabil-ir-im.
music listen-CMPL sleep-MODAL-AOR-1sg
'While listening to music, I can sleep.'
(Not, 'while I am able to listen to music, I am able to sleep.')

In (16a), the modal operator *-meli-* "ought" has scope over both cores, and therefore the nexus is cosubordinate; in (16b), on the other hand, the modal operator *-yabil-* "able" has scope only over the final core, hence coordinate nexus. The structural representations for (16a, b) are given in Figure 30.17.

The following sentences from Kewa (Franklin 1971) are a minimal triple for the three nexus types at the clause level.



Click to view larger

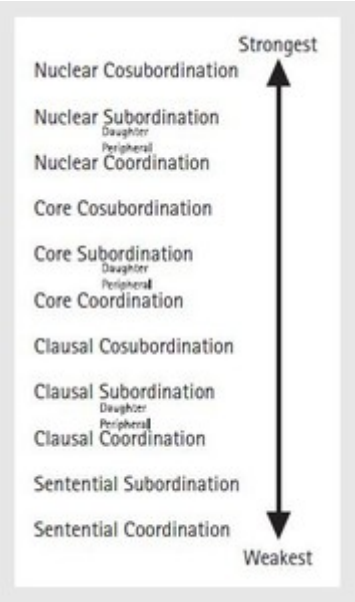
Figure 30.17 Turkish Core Junctures

(17)

a. Nipú ípu-la pare ní paalá na-pía. Coordination
3sg come-3sgPRES but 1sg afraid NEG-be.1sgPRES
'He is coming, but I am not afraid.'

b. (Ní) Épo lá-ri épa-wa. Cosubordination
(1sg) whistle say-SIM.SS come-1sgPAST
'I whistled while I came,' or 'I came whistling.'

c. (Ní) Épo lá-lo-pulu irikai épa-lia.
(1sg) whistle say-1sgPRES-CAUSAL dog come-3sgFUT
'Because I am whistling, the dog will come.' Subordination (peripheral)



Click to view larger

Figure 30.18 Interclausal syntactic relations hierarchy

(p. 731) The four levels of juncture combine with the three nexus types to generate eleven possible complex sentence types; there is no sentential cosubordination because there are no sentence-level operators, hence no possible operator sharing. In addition, both subtypes of subordination are possible at the clause, core, and nuclear levels. Not all of them are instantiated in every language. The juncture-nexus types found in a language may be realized by more than one formal construction type; for example, both *Mary sat playing the guitar* and *Robin tried to open the door* instantiate core cosubordination, while both *For Sam to leave now would be a mistake* and *Lisa's losing her job shocked everyone* instantiate core subordination in English. The juncture-nexus types may be ordered into a hierarchy in terms of the tightness of the syntactic link between the units, i.e., in terms of how integrated the units are into a single unit or are coded as distinct units. This is given in Figure 30.18.

(p. 732) The syntactic clause-linkage relations discussed earlier are used to express certain semantic relations between the units in the linkage, e.g., causation, purpose, and temporal sequence. The interclausal semantic relations are given in (18).

(18) Interclausal Semantic Relations

a. Causative [1]: the bringing about of one state of affairs directly by another state of affairs, usually an event or action, e.g., *Max painted the door green, Larry pushed the door open*.

b. Phase: a separate verb describes a facet of the temporal envelope of a state of affairs, specifically its onset, its termination, or its continuation, e.g., *Chris started crying, Fred kept singing, Hari finished writing the chapter*.

c. Modifying subevents

1. Manner: the manner in which a motion event is carried out, e.g., *Bill entered the room skipping*.

2. Motion: motion accompanying another action, e.g., Mparntwe Arrerente *angk-tyantye*-[speak-go.upward] ‘speak while going up– (Wilkins 1991).

3. Position: stance while doing an action, e.g., *Dana sat reading a newspaper*.

4. Means: the means by which an action is carried out, e.g., *Sam opened the box by slicing it with a knife*.

d. Psych-action: a mental disPosition regarding a Possible action on the part of a participant in the state of affairs, e.g., *Max decided to leave, Sally forgot to open the window, Tanisha wants to go to the movies*.

e. Purposive: one action is done with the intent of Realizing another state of affairs, e.g., *Juan went to the store to buy milk, Susan brought the book to read*.

f. Jussive: the expression of a command, Request, or demand, e.g., *Pat asked the student to leave, The king ordered the troops to attack the city*.

g. Causative [2]: the bringing about of one state of affairs through a distinct action or event, e.g., *Fred forced Max to paint the table*.

h. Direct perception: an unmediated apprehension of some act, event, or situation through the senses, e.g., *Rex saw the child open the door, Yolanda heard the guests arrive*.

i. Indirect perception: the deduction of some act, event, or situation from evidence of it, e.g., (looking at an empty desk) *I see that John has gone home early*.

j. Propositional attitude: the expression of a Participant–s attitude, judgment, or opinion regarding a state of affairs, e.g., *Carl believes that UFOs are a menace to the earth, Paul considers Carl to be a fool, Most fans want very much for their team to win*.

(p. 733) **k.** Cognition: an expression of knowledge or mental activity, e.g., *Aaron knows that the earth is round, George is thinking about Madeleine–s refusal to go out with him*.

l. Indirect discourse: an expression of Reported speech, e.g., *Frank said that his friends were corrupt*.

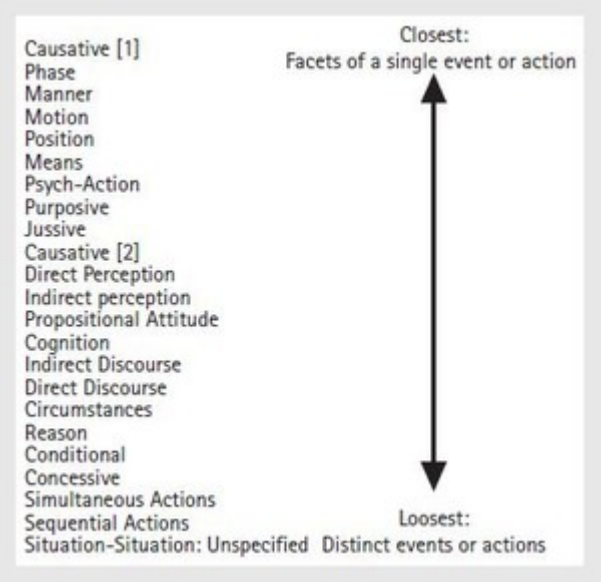
m. Direct discourse: the direct quotation of a speech event, e.g., *Frank said, ‘My friends are corrupt.’*

n. Circumstances: the spatial or temporal parameters of an event, e.g., *Sam talked to Sally at the library after work*.
o. Reason: the motivation or cause for an action or event, e.g., *The baby cried because she was hungry*.
p. Conditional: an expression of what consequence would hold, given the conditions in a Particular state of affairs, e.g., *If it rains, we won–t be able to have a picnic, Were Fred to leave now, he would look like a fool*.
q. Concessive: the content of the main clause holds unexpectedly, given the content of the subordinate clause, e.g., *Bill made it to work, even though it was snowing heavily*.
r. Temporal
1. Simultaneous states of affairs: one state of affairs is temporally coterminous with another, e.g.,*Max danced and Susan played the piano, Kim had chicken pox and at the same time Leslie had the measles*.
2. Sequential states of affairs: one state of affairs follows another temporally, with or without any temporal overlap, e.g., *Juan had finished talking, and then Carlos entered the room, Vidhu was sitting down, and the band began to play*.

s. Temporally unordered states of affairs: the temporal relation between states of affairs is unexpressed, e.g., *Tyrone talked to Tanisha, and Yolanda chatted with Kareem*.

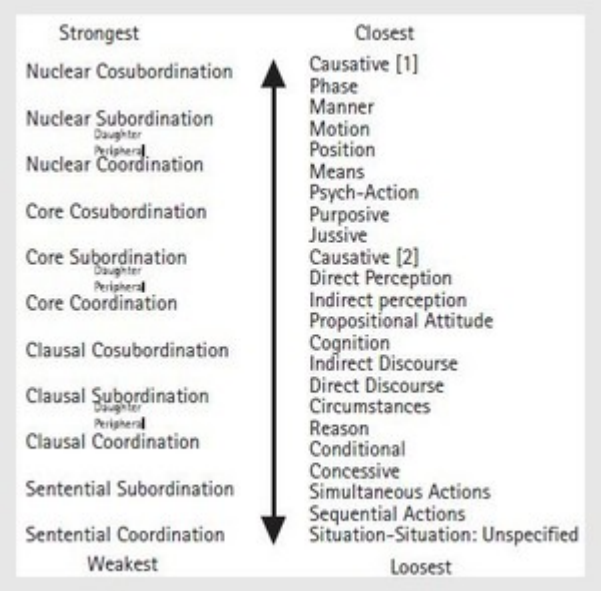
These relations may be formalized in terms of the same decomposition used for verbs (see Van Valin 2005: 207–8; also Ohori 2001, 2005).

The semantic relations form a continuum expressing the degree of semantic cohesion between the propositional units linked in the complex structure, i.e., the degree to which they express facets of a single action or event or discrete actions or events. This may be represented as in Figure 30.19.



Click to view larger

Figure 30.19 Interclausal semantic relations hierarchy



Click to view larger

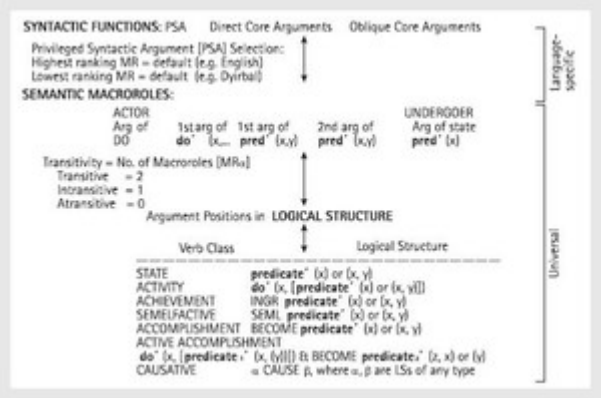
Figure 30.20 Interclausal relations hierarchy

The syntactic linkage relations are ranked hierarchically in terms of the strength of the syntactic bond between the units in Figure 30.18. The interaction of the two hierarchies is expressed in the interclausal relations hierarchy in Figure 30.20. The relationship between the syntactic and semantic relations in clause linkage is very complex; i.e., it is not one-to-one, but there are some striking regularities cross-linguistically. The primary principle governing the interaction of the two hierarchies is iconic: the closer the semantic relation between two propositions is, the stronger the syntactic (p. 734) (p. 735) link joining them (Silverstein 1976; Givón 1980). In other words, the semantic relations at the top end of the hierarchy should be realized by the linkage categories at the top as well, and the relations at the bottom of the hierarchy should be realized by the linkage categories at the bottom of the syntactic side. Moreover, while there is often more than one syntactic realization of a particular semantic relation, the tightest syntactic linkage realizing it should be tighter than the tightest syntactic linkage realizing looser semantic relations.

30.3 Linking Between Syntax and Semantics

All of the components of the RRG linking system have been introduced. This is summarized in Figure 30.21.

Logical structures, macroroles, and the hierarchy linking them are universal in that there is very little cross-linguistic variation; this is the domain of lexical processes. Where languages differ substantially is how macroroles and other arguments link into the syntax.



Click to view larger

Figure 30.21 RRG Linking System

The reason the arrows in Figure 30.21 are double-headed is that the linking system works both from semantics to syntax and from syntax to semantics. A theory which could describe the linking from semantics to syntax only could be part of a language production system, but it would not be adequate for a comprehension system. In such (p. 736) a system, the parser, as an idealization, would take the input and produce a structured syntactic representation of it, identifying the elements of the layered structure of the clause and the cases, adpositions and other grammatically relevant elements in the sentence. It is then the grammar's job to map this structure into a semantic representation, as the first step in interpreting it, and this is where the syntax to semantics linking algorithm is required. The details of the linking algorithms are given in Van Valin (2005).

The linking between syntax and semantics is governed by a very general principle called the "Completeness Constraint"; it states simply that all of the specified arguments in the semantic representation of a sentence must be realized in the syntax in some way, and conversely that all of the expressions in the syntax must be linked to something in the semantic representation of a sentence, in order to be interpreted.

An important part of the linking involves finite verb agreement, case assignment, and preposition assignment. The finite verb agreement rule for accusative languages like English, German, and Croatian is given in (19).

Finite verb agreement in Croatian, German, and Icelandic:

(19) The controller of finite verb agreement is the highest-ranking core macrorole argument (in terms of (10)).

The rule is not formulated with respect to any syntactic position or function, or with respect to any case. Case assignment rules are formulated in a similar way. The basic

rules for direct core arguments in accusative languages are given in (20) and for ergative languages in (21); these do not pertain to case assigned by adpositions.

(20)
Case marking rules for accusative languages:
a. Highest-ranking core macrorole (in terms of (10)) takes nominative case.
b. Other core macrorole takes accusative case.

(21)
Case marking rules for ergative languages:
a. Lowest-ranking core macrorole (in terms of (10)) takes absolutive case.
b. Other core macrorole takes ergative case.

In addition, there is a rule for dative case assignment, which applies to both systems.

(22) Assign dative case to non-macrorole direct core arguments (default).

Dative case is assigned only when the rules for the other cases cannot apply.[9] In a language like English without RP case marking, there are rules for preposition assignment (Jolly 1993). The rules for *to* and *from* are given in (23).[10] (p. 737)

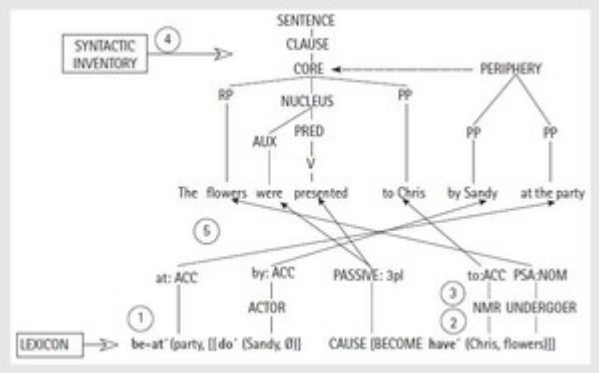
(23)
Preposition assignment rules for English
a. Assign *to* to NMR *x* argument in LS segment: . . . BECOME/INGR pred′ (x, y)
b. Assign *from* to NMR *x* argument in LS segment: . . . BECOME/INGR NOT pred′ (x, y)

The rule in (23a) is particular important for the "dative shift" verbs in English, e.g., *give*, *send*, *show*, etc. The alternation in (24) is handled in terms of variable undergoer selection; both sentences would have the same LS, given in (24c).

(24)
a. Mary sent a letter to Sally.
b. Mary sent Sally a letter.
c. [do′ (Mary, Ø)] CAUSE [INGR have′ (Sally, letter)]

In (24a) undergoer selection reflects the default choice in terms of the AUH in Figure 30.10, i.e., the rightmost argument in LS is chosen to function as undergoer. In (24b), on the other hand, the second lowest-ranking argument (which is also the second highest-ranking), *Sally*, is selected as undergoer. In (24a), the conditions for (23a) are met, and therefore Sally is marked by *to*. In (24b), however, this is not met, and therefore it does not apply. Alternations with verbs like English *present*, German *schenken* vs. *beschenken* "give as a gift", Croatian *darovati* "give as a gift", and Dyirbal *wugal* "give" are all analyzed as instances of variable undergoer selection (see Van Valin 2005, §4.4; Van Valin 2007).

Most of what counts as "syntax" in many theories is handled in RRG in terms of constraints on the semantic representation, in terms of information structure, or in syntactic phase of the linking (see the analysis of WH-questions below). The analysis of reflexivization in RRG follows the approach in Jackendoff (1992) and states the hierarchical constraints for core-internal ("clause-bound" in other theories) reflexivization at the LS level, not with respect to the syntactic representation. The principles affecting the scope and interpretation of quantifiers are related to information-structure contrasts, not phrase structure. RRG treats constructions as an important part of syntax, and they are represented in terms of constructional schemas. Cross-constructional and cross-linguistic generalizations are captured in terms of the general principles and constraints that constitute the linking algorithms, e.g., the actor–undergoer hierarchy, the layered structure of the clause, the PSA selection hierarchy. Only the idiosyncratic, language-specific features of constructions are represented in constructional schemas, which may include syntactic, morphological, semantic, and pragmatic (focus structure) information.



Click to view larger

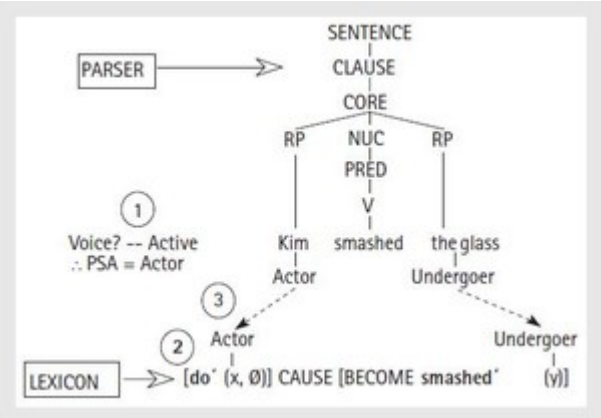
Figure 30.22 Linking from semantics to syntax in a simple sentence in English

A simple example from English illustrating the operation of the semantics-to-syntax linking algorithm is given in Figure 30.22. The numbers refer to the general steps of the algorithm: (1) constructing the semantic representation of the sentence in the lexicon; (2) assigning actor and undergoer; (3) determining PSA selection, case and adposition assignment, and agreement; (4) selecting the appropriate syntactic template from the syntactic inventory; and (5) linking the elements from the semantic representation into (p. 738) the appropriate positions in the syntactic representation. The numbers in the diagram, especially 2–3, should not be interpreted as indicating steps in a derivation. Rather, they signal steps in the linking which involve adding morphosyntactic information to the semantic representation. The output of step 3 could equally well be represented as " … [**do**´ (Sandy[Actor, *by*-ACC], Ø)] CAUSE [BECOME **have**´[passive, 3pl] (Chris[NMR, *to*-ACC], flowers[Und, PSA, NOM])] … ".

Because this sentence is a passive, the undergoer appears as the "subject", with the actor appearing in a peripheral PP marked with *by*. These language-specific details would be represented in the constructional schema for the English passive, given in Table 30.4.

The information in the constructional schema is a combination of general principles (template selection principles, PSA selection principles, general characterization of non-

default PSA selection), plus language-specific information, e.g., the form of the verb and the choice of auxiliary. See Van Valin (2005) for detailed discussion and explication of all of these points.



Click to view larger

Figure 30.23 Linking from syntax to semantics in a simple sentence in English

A simple example of the linking from syntax to semantics is given in Figure 30.23. Here again the numbers refer to the general steps in the algorithm: (1) extract all of the information possible from the overt morphosyntactic form of the sentence, including the voice of the verb (if the language has voice), case marking, word order, and adpositions; (2) retrieve the LS of the predicate in the nucleus from the lexicon and assign macroroles to the extent possible; and (3) link of the information derived from steps (1) and (2). The syntactic representation is produced by the parser, which turns the acoustic input into a labeled syntactic representation. (p. 739)

Table 30.4 Constructional schema for English passive (plain)

CONSTRUCTION: English passive (plain)

SYNTAX:

Template(s): (following template selection principles; not given above)
PSA: (11a, c2), Variable [± pragmatic influence]
Linking: Undergoer to PSA; Actor omitted or in peripheral by-PP

MORPHOLOGY:

Verb: past participle
Auxiliary: be

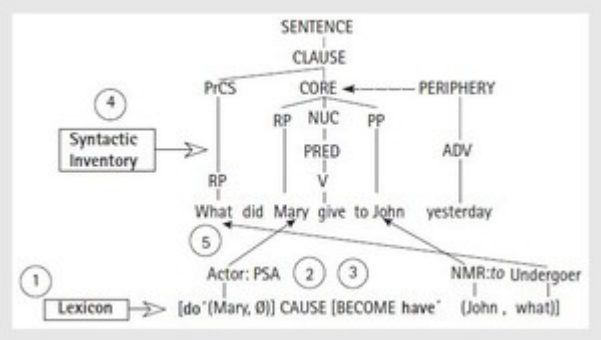
SEMANTICS:

PSA is not instigator of state of affairs but is affected by it (default)

PRAGMATICS:

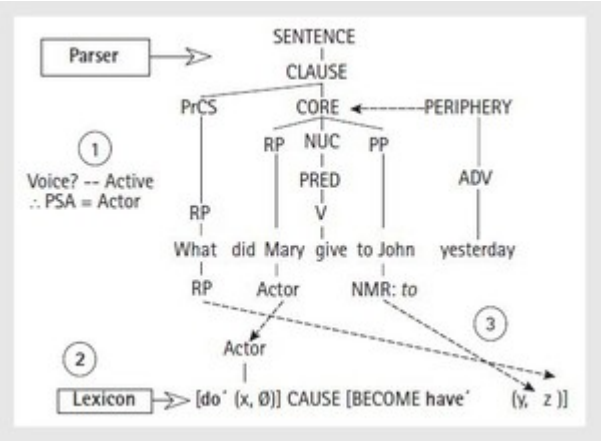
Illocutionary force: Unspecified
Focus structure: No restrictions; PSA = topic (default)

The linking in a WH-question in English, in both directions, is illustrated in Figures 30.24 and 30.25; the linking of the peripheral adjunct *yesterday* is not represented. In the linking from semantics to syntax in Figure 30.24, the undergoer *what* is linked directly to the PrCS; there is no empty argument position in the core, i.e., no trace. The rule in (23a) applies to assign *John* the preposition *to*.



Click to view larger

Figure 30.24 Linking from semantics to syntax in a WH-question in English



Click to view larger

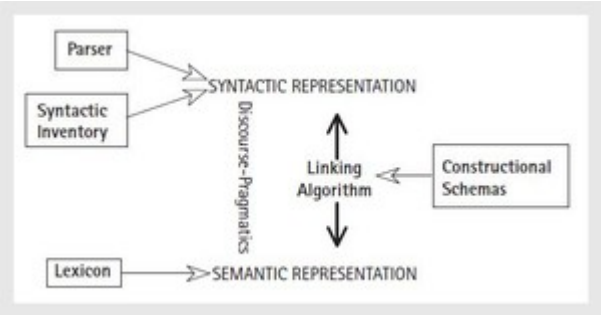
Figure 30.25 Linking from syntax to semantics in a WH-question in English

There are two important complications in the syntax to semantics linking with this sentence, as shown in Figure 30.25. First, no conclusion can be drawn from the morphosyntax regarding the function of *what*; hence in step 1 this is simply labeled "RP". Second, because *give* is a variable undergoer selection verb, it is not possible (p. 740) to assign undergoer to an argument in the LS in step 2, unlike in Figure 30.23. So the linking of *John* is determined by using the inverse of (23a) as a linking principle: since *John* is marked by *to*, it must be the first argument of **have**′. After *Mary* is linked to the *x* argument and *John* to *y*, the Completeness Constraint forces the linking of *what* to the *z* argument, which yields the correct interpretation. Constraints on WH-question formation and other "extraction" constructions are explained in terms of the interaction of information structure and syntax, in particular in terms of restrictions on the potential focus domain (Van Valin 1995, 1998, 2005).

(p. 741) 30.4 Conclusion

The more complete picture of RRG that emerges from this discussion is given in Figure 30.26.

In the linking from semantics to syntax, the source of the syntactic representation is the templates of the syntactic inventory. In the syntax to semantics linking, the source is the parser. The lexicon plays an important role in both. Discourse-pragmatics, i.e., information structure, interacts with the linking algorithm in significant ways at various steps. Constructional schemas provide the language-specific morphosyntactic information that complements the general, cross-linguistically valid principles of the theory.



Click to view larger

Figure 30.26 Organization of Role and Reference Grammar (final)

In §30.1, the motivating questions for RRG were presented: “what would a linguistic theory look like if it were based on the analysis of languages with diverse structures, such as Lakhota, Tagalog, Dyirbal, and Barai (Papua New Guinea), rather than on the analysis of English?”, and “how can the interaction of syntax, semantics, and pragmatics in different grammatical systems best be captured and explained?” This chapter has sketched out the answer to these questions, with emphasis on the first. Because of this typologically diverse starting point, the constructs presented herein differ in significant ways from those in other theories: a non-endocentric theory of phrase and clause structure, the concept of reference phrase instead of noun phrase, rejection of the universality of grammatical relations, the construction-specific notion of PSA, cosubordination as a third linkage type in addition to coordination and subordination, etc. These typologically motivated theoretical and descriptive constructs have made RRG a useful framework for linguists primarily concerned with language description, and yet it makes possible the analysis and explanation of the kind of morphosyntactic phenomena which have been the focus of much of the theoretical work of the past decades. (p. 742)

Notes:

([1]) In addition to the references in Van Valin and LaPolla (1997), see also Van Valin (1998, 2001*b*, 2002); Weist (2002); Weist et al. (2004).

([2]) RRG does allow zero morphemes in morphological paradigms; what is excluded are phonologically null pronouns (*pro*, PRO), noun phrases (traces), light verbs, adpositions, etc., in syntactic representations.

([3]) See VVLP, chapter 2, for arguments against a traditional or X-bar phrase structure analysis.

([4]) Abbreviations: ABS “absolutive”, ADV “adverb”, AUH “Actor-Undergoer Hierarchy”, DET “determiner”, ERG “ergative”, IF “illocutionary force”, LDP “left-detached position”, LOC “locative”, LS “logical structure”, NMR “non-macrorole”, NUC “nucleus”, PERF “perfect”, PrCS “pre-core slot”, PRED “predicate”, PRO “pronoun”, PROG “progressive”, PSA “privileged syntactic argument”, RP “reference phrase”, TNS “tense”.

([5]) See Everett (2008) for a discussion of intentional state constructions in Wari’, an Amazonian language, in which whole clauses can serve as the nucleus of a sentence.

([6]) This claim was supported by the results of the typological survey reported in Bybee (1985*a*).

([7]) RRG treats the notion of “agent” rather differently from other theories. The basic notion is “effector”, which is the first argument of **do**′ and is unspecified for agentivity. With many verbs, a human effector may be interpreted as an agent in certain contexts. If the verb lexicalizes agentivity, as with *murder*, then the logical structure contains “DO”, which indicates that the argument must be interpreted as an agent. See Holisky (1987), Van Valin and Wilkins (1996), VVLP §3.2.3.2, for detailed discussion. Also, primary-object languages patterns require a modified undergoer selection principle, namely that the undergoer is the second-highest ranking argument in the LS; see Guerrero and Van Valin (2004), Van Valin (2005: 123–7).

([8]) While it is normally the case that languages with variable PSAs have a voice opposition, it is not always the case. See Van Valin (2009) for discussion of two languages, Liangshan Nuoso and Barai, which have variable PSAs but lack a formal voice opposition.

([9]) There is also a rule for assigning instrumental case, but it is complex and not necessary for this discussion. See Van Valin (2005, §4.4).

([10]) There is also a rule for assigning *with*, which is very similar to the rule for instrumental case.

Robert D. Van Valin Jr.

Robert D. Van Valin, Jr. received his Ph.D. in linguistics at the University of California, Berkeley. His research is focused on theoretical linguistics, especially syntactic theory and theories of the acquisition of syntax and the role of syntactic theory in models of sentence processing. He is the co-author of Functional Syntax and Universal Grammar (Cambridge University Press, 1984), the editor of Advances in Role and Reference Grammar (Benjamins, 1993), the co-author of Syntax: Structure, Meaning and Function (Cambridge University Press, 1997), and the author of An Introduction to Syntax (Cambridge University Press, 2001). His most recent book is Exploring the Syntax-Semantics Interface (Cambridge University

Press, 2005). He is the general editor of the Oxford Surveys in Syntax and Morphology series (Oxford University Press).