

# A Brief Overview of Role and Reference Grammar

Robert D. Van Valin, Jr.  
(State University of New York at Buffalo)

## 1. Introduction

Role and Reference Grammar [RRG] (Van Valin 1993, Van Valin & LaPolla 1997, Yang 1998) takes language to be a system of communicative social action, and accordingly, analyzing the communicative functions of grammatical structures plays a vital role in grammatical description and theory from this perspective. Language is a system, and grammar is a system in the traditional structuralist sense; what distinguishes the RRG conception of language is the conviction that grammatical structure can only be understood and explained with reference to its semantic and communicative functions. In terms of the abstract paradigmatic and syntagmatic relations that define a structural system, RRG is concerned not only with relations of cooccurrence and combination in strictly formal terms but also with semantic and pragmatic cooccurrence and combinatory relations. It is a monostratal theory, positing only one level of syntactic representation, the actual form of the sentence (cf. fn. 5). With respect to cognitive issues, RRG adopts the criterion of psychological adequacy formulated in Dik (1991), which states that a theory should be "compatible with the results of psycholinguistic research on the acquisition, processing, production, interpretation and memorization of linguistic expressions" (1991:248). It also accepts the related criterion put forth in Bresnan & Kaplan (1982) that theories of linguistic structure should be directly

relatable to testable theories of language production and comprehension. The RRG approach to language acquisition, sketched in Van Valin (1991a, 1994), rejects the position that grammar is radically arbitrary and hence unlearnable, and maintains that it is relatively motivated (in Saussure's sense) semantically and pragmatically. Accordingly, there is sufficient information available to the child in the speech to which it is exposed to enable it to construct a grammar.<sup>1</sup>

## 2. Historical background

RRG grew out of an attempt to answer two basic questions: (i) what would linguistic theory look like if it were based on the analysis of Lakhota, Tagalog and Dyirbal, rather than on the analysis of English?, and (ii) how can the interaction of syntax, semantics and pragmatics in different grammatical systems best be captured and explained? These questions reflect issues that were prominent in the mid-1970's in some strands of American linguistics. Dixon's grammar of Dyirbal and Schachter & Otnes' grammar of Tagalog had been published in 1972, and the implications of these languages for linguistic theories were just being recognized. Furthermore, the Prague School and Hallidayan ideas regarding the role of discourse-pragmatics in grammar were being explored from a number of different perspectives. Many of the typological issues, e.g. the universality of the notion of 'subject', and theoretical issues, e.g. the relation between 'subject' and 'topic' in grammatical systems, were central in the initial conceptualization of RRG, and this is reflected in the early work on the theory (Foley & Van Valin 1977, Van Valin 1977a,b, 1980, 1981 Van Valin & Foley 1980).

The theory from which RRG is most directly descended is Fillmore's

---

1. In Van Valin (1991a) it is argued that assuming that grammar is learned is the only empirically falsifiable position, following the subset principle.

(1968) Case Grammar. As in Fillmore's model, there is a semantic representation employing semantic case roles which is mapped into the syntactic surface structure, without any intervening level of syntactic representation. The details of the mapping differ substantially, however, and one prime difference is that in RRG discourse-pragmatic factors may play a role in the mapping. The RRG theory of clause structure follows the insight of Fillmore's division of the clause into "modality" and "proposition" by treating predicates, arguments and their modifiers distinctly from grammatical categories such as tense, aspect, modality and mood. Finally, RRG, like Fillmore, does not assume grammatical relations to be universal.

### 3. Central concepts of the theory

The basic organization of RRG is given in Figure 1.

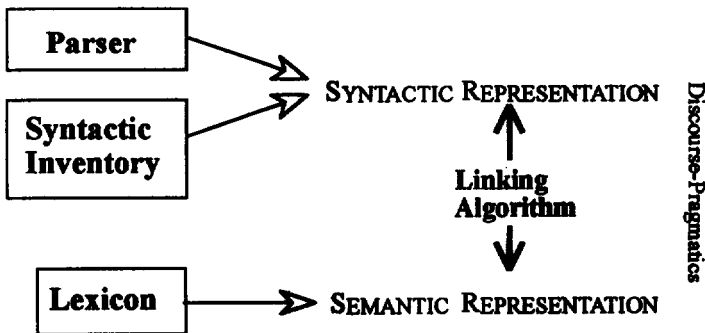


Figure 1: Organization of Role and Reference Grammar

As mentioned earlier, RRG is a monstratal theory, and there is a direct mapping or linking between the semantic representation of a sentence and its syntactic representation. Each of the aspects of this figure will be explicated in the course of the discussion.

### 3.1 Clause structure

RRG rejects the standard formats for representing clause structure (grammatical relations, X-bar syntax), because they are not universal and hence necessarily impose aspects of structure on at least some languages where it is not appropriate. This follows from the assumptions regarding a theory of clause structure in (1).

(1) General considerations for a theory of clause structure:

- a. A theory of clause structure should capture all of the universal features 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.

The RRG conception of clause structure (originally proposed in Foley & Van Valin 1984 and further developed in Van Valin 1993), is known as the LAYERED STRUCTURE OF THE CLAUSE [LSC]. It is made up of the NUCLEUS, which contains the predicate(s), the CORE, which contains the nucleus plus the arguments of the predicate(s), and the PERIPHERY, which contains adjunct temporal and locative modifiers of the core. The semantic basis of the LSC is summarized in Table 1.

These aspects of the LSC are universal. Some language have a PRE-CORE SLOT [PRCS], which is the position of WH-words in languages like English and Malagasy, and a LEFT-DETACHED POSITION, [LDP], which is the position of the pre-clausal element in a left-dislocation construction or the NP marked by *-(n)un* in Korean or *wa* in Japanese (Yang 1984). In addition, some verb-final languages have a POST-CORE SLOT [POCS] (e.g. Japanese; Shimojo 1995), and some languages also have a RIGHT-DETACHED POSITION, [RDP],

<u>Semantic Element(s)</u>	<u>Syntactic Unit</u>
Predicate	Nucleus
Argument in Semantic representation of predicate	Core argument
Non-arguments	Periphery
Predicate + Arguments	Core
Predicate + Arguments + Non-arguments	Clause (= Core+Periphery)

Table 1: Semantic Units Underlying the Syntactic Units of the Layered Structure of the Clause

which is the position of the post-clausal element in a right-dislocation construction. Each of the major layers (nucleus, core, clause) is modified by one or more OPERATORS, which include grammatical categories such as tense, aspect, modality and evidentiality. The LSC applies equally to fixed word-order and free word-order languages, to head-marking and dependent-marking languages, to languages with and without grammatical relations.<sup>2</sup> In the formal representation of the LSC (proposed in Johnson 1987), operators are represented in a distinct projection of the clause from the predicates and arguments (the constituent projection). This is presented in Figures 2-5.

---

2. It is assumed that noun phrases and adpositional phrases have a comparable layered structure; see Van Valin (1993), §2.3. Operators in the NP include determiners, quantifiers and adjectival and nominal modifiers.

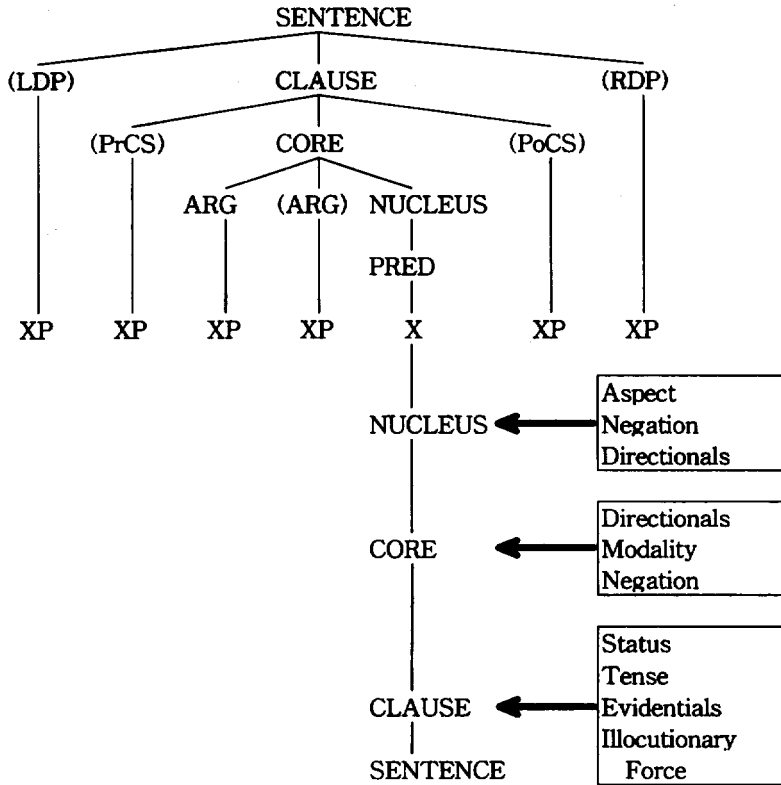


Figure 2: The Layered Structure of the Clause<sup>3</sup>

3. The periphery has been omitted from this diagram for the sake of simplicity.

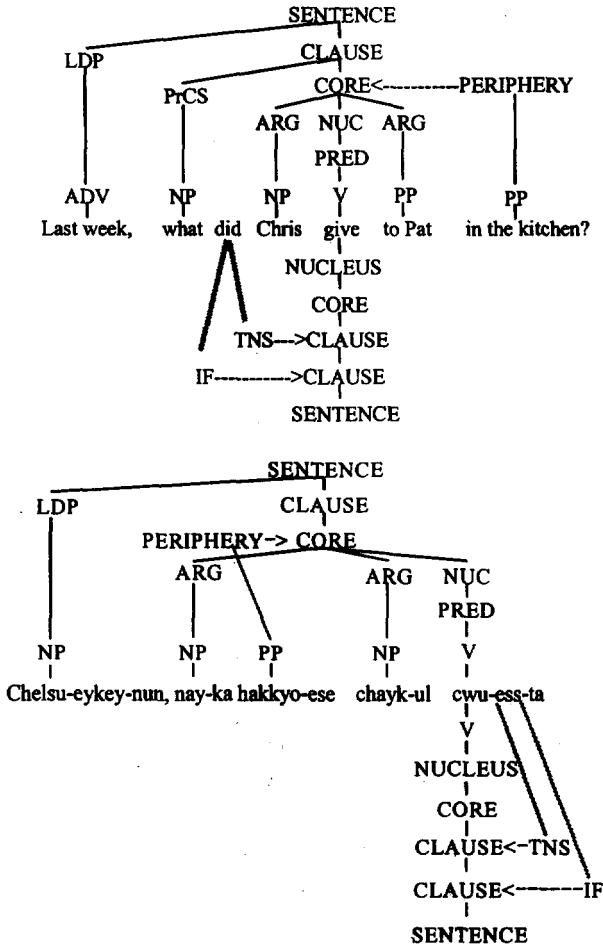


Figure 3: The LSC in English<sup>4</sup> and Korean

4. Did is labelled 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, core-initial (pre-core) tense signals interrogative IF, and the absence of tense in a matrix core signals imperative IF.

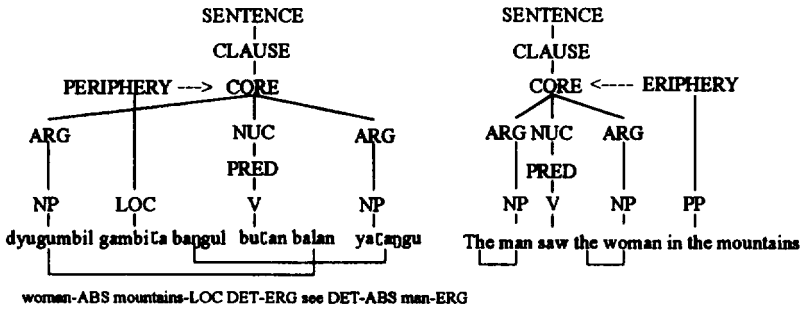


Figure 4: The LSC in Dyirbal and English

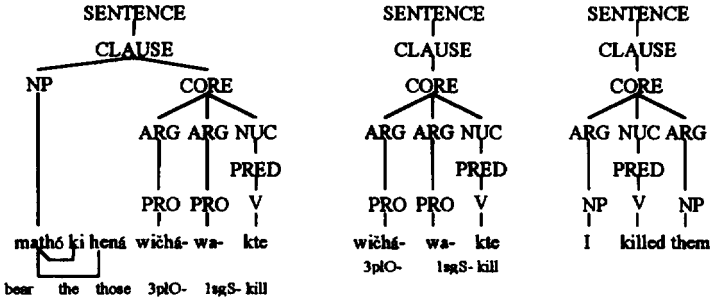


Figure 5: The LSC in Lakhota(Head-marking) and English(Dependent-marking)

Dyirbal (Australia; Dixon 1972) and Lakhota (Siouan, North America) represent free-word-order and head-marking languages, respectively; Dyirbal is also dependent-marking. The operator projections have been omitted in the Dyirbal and Lakhota examples. The lines connecting the determiners to the head nouns are the operator projection within the NP, analogous to the operator projection within the clause, as in Figures 2-3. In head-marking languages like Lakhota, the bound pronominals on the verb are considered to be the core arguments; overt NPs are within the clause in apposition to them (Van Valin 1985, 1987). Note



that despite the differences between the three languages in Figures 4-5, comparable structural relations, e.g. core argument, peripheral adjunct, are represented in the same way. It should be noted that these representations are not abstract, unlike relational networks or functional structures; they are intended to be concrete, in the sense that they should represent the actual form of the sentence, including the linear sequence of its constituent elements and their morphological properties.<sup>5</sup>

Representations of constituent projections such as these should be viewed as constructional templates, the inventory of which in a language constitutes an important component of its grammar. It may be termed the SYNTACTIC INVENTORY and complements the lexicon (cf. Figure 1).

The three layers of the LSC are also the three basic 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, or clauses with clauses. These are called levels of JUNCTURE in RRG, i.e. nuclear juncture, core juncture and clausal juncture. Clausal junctures, as the name implies, involve sentences containing multiple clauses. Examples of nuclear junctures from French, English and Mandarin are given in (2) and their representations are in Figure 6. Justifications for these structures can be found in Foley & Van Valin (1984), Van Valin (1993), Van Valin & LaPolla (1997).

- (2) a. Je ferai manger les gâteaux à Jean.  
           [two nuclei, *faire* and *manger*, in a single core]  
       1sg make.FUT eat the cakes to John  
       'I will make John eat the cakes.'
- b. John forced open the door.  
           [two nuclei, *push* and *open*, in a single core]

---

5. The representation may be abstract with respect to phonology or morphophonology, e.g. the output could be in terms of abstract morphophonological units rather than concrete phonetic ones. RRG is not concerned with the issue of (morpho)phonological representation.

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

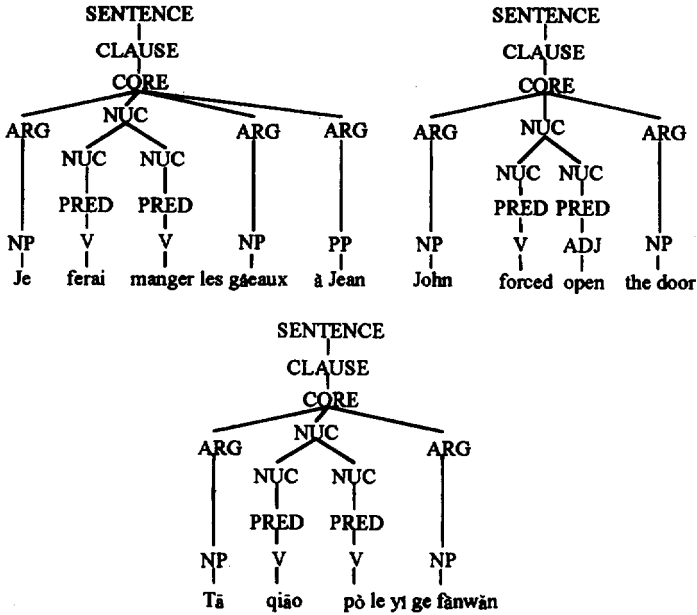


Figure 6: Nuclear junctures in French, English and Mandarin

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 (3) and their structures in Figure 7. 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 (cf. §3.4).

- (3) 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 xiě zì  
 3sg teach 1sg write characters  
 'She teaches me to write characters.'

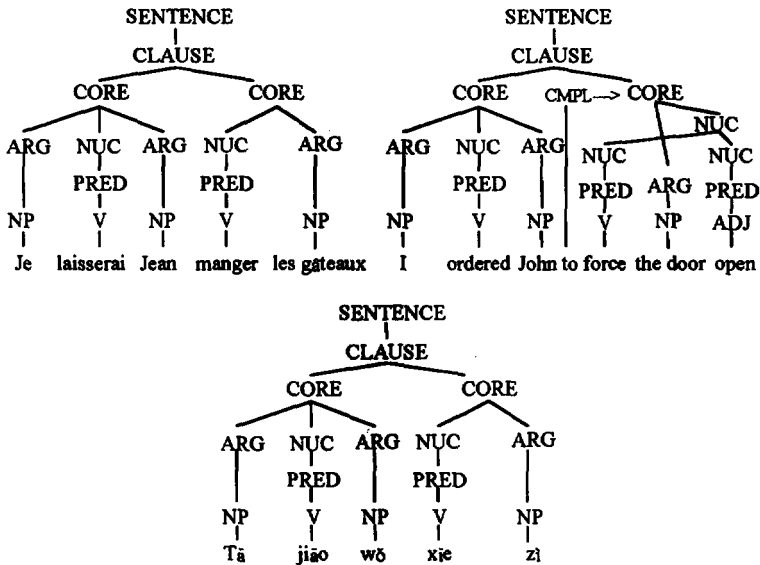


Figure 7: Core Junctures in French, English and Mandarin

The RRG theory of complex sentences is equally concerned with the set of possible syntactic and semantic relations between the units in a juncture.<sup>6</sup> The syntactic relations between units are termed NEXUS relations in RRG. Traditionally, only two basic nexus relations are recognized, coordination and subordination, but RRG, following Olson's (1981) analysis of clause linkage in *Burai* (a Papuan language), postulates three nexus types: coordination, subordination, and COSUBORDINATION, which is in essence dependent coordination. The dependence is operator

6. The semantic relations will be discussed in §3.2 below.

dependence; that is, in cosubordination, the units obligatorily share one or more operators at the level of juncture. In the Mandarin example in (2c), aspect obligatorily has scope over both nuclei, and therefore the nexus is cosubordination. This is represented as in Figure 8.

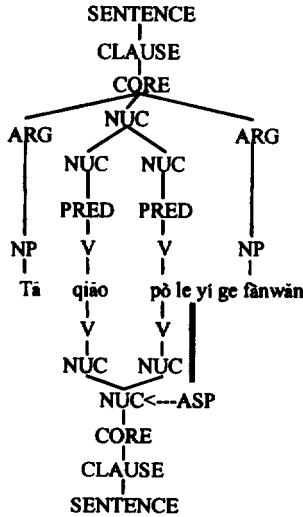


Figure 8: Nuclear cosubordination in Mandarin

The following examples from Turkish (Watters 1993) exemplify obligatory operator sharing and the lack of it in Turkish core cosubordination and coordination, respectively.

- (4) a. *Gid-ip gör-meli-yiz.* Core cosubordination  
 go-CMPL see-MODAL-1pl  
 'We ought to go and see.'
- b. *Müzik dinle-yerek, uyu-yabil-ir-im.* Core coordination  
 music listen-CMPL sleep-MODAL-AOR-1sg  
 'Listening to music, I can sleep.'

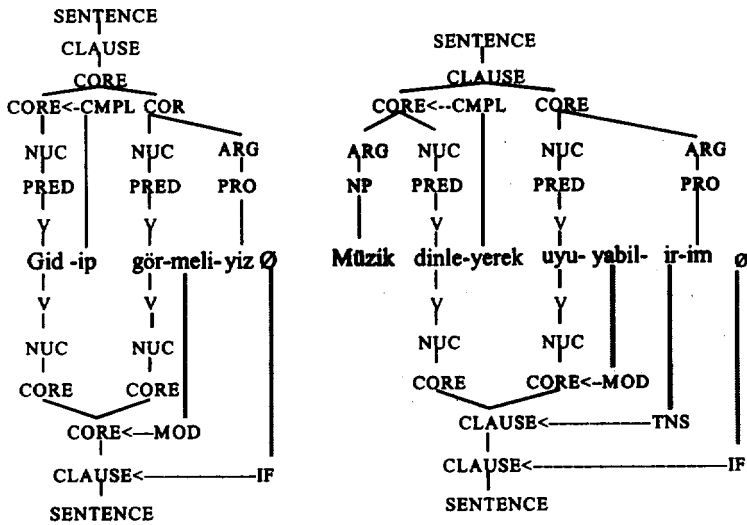


Figure 9: Turkish Core Junctures

In (4a), the modal operator *-meli-* 'ought' has scope over both cores, and accordingly the nexus is cosubordinate; in (4b), on the other hand, the modal operator *-yabil-* 'able' has scope only over the final core, hence coordinate nexus.<sup>7</sup> The following examples from Kewa (Franklin 1971) are a minimal triple for the three nexus types at the level of clausal juncture.

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

7. The term 'coordination' here is being used for an abstract linkage relation referring to a relationship of equivalence and 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.

- 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. Subordination  
 1sg whistle say-1sgPRES-CAUSAL dog come-3sgFUT  
 'Because I am whistling, the dog will come.'

The three levels of juncture together with the three nexus types create nine possible complex sentence types. Not all of them are instantiated in every language; for example, Korean appears to have all nine (Yang 1994), while English appears to have six and Jacaltec seven. 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 nine juncture-nexus types may be ordered into a hierarchy in terms of the tightness of the syntactic link between the units (see the hierarchy in Figure 11 in §3.2).

### 3.2 Semantic structure

The semantic representation in Figure 1 is based on a system of lexical representation and semantic roles. The system of lexical representation is based on Vendler's (1967) Aktionsart classification of verbs into states, activities, achievements and accomplishments. There is an additional class, called active accomplishments, which are telic uses of activity verbs. Examples of each class and their formal representation are given in (6)-(7).<sup>8</sup>

---

8. This system differs in important ways from the one proposed in Foley & Van Valin (1984) and Van Valin (1990, 1991b, 1993).

- (6) a. State:                   The teacher is upset about the school situation.  
 a'. Causative state: The school situation upsets the teacher.  
 b. Achievement:                The bubble popped.  
 b'. Causative achievement:     The baby popped the bubble.  
 c. Accomplishment:            The snow melted.  
 c'. Causative accomplishment:  The hot sun melted the snow.  
 d. Activity:                    The soccer ball rolled around the field.  
 d'. Causative activity:  
                                   The girl rolled the soccer ball around the field.  
 e. Active accomplishment:  
                                   The soldiers marched to the barracks.  
 e'. Causative active accomplishment  
                                   The sergeant marched the soldiers to the barracks.

- (7) a. State                        **predicate'** (x) or (x,y)  
 b. Activity                        **do'** (x, [**predicate'** (x) or (x, y)])  
 c. Achievement                  **INGR(ESSIVE) predicate'** (x) or (x,y), or  
                                   **INGR do'** (x, [**predicate'** (x) or (x, y)])  
 d. Accomplishment               **BECOME predicate'** (x) or (x,y), or  
                                   **BECOME do'** (x, [**predicate'** (x) or (x, y)])  
 e. Active accomplishment  
                                   **do'**(x, [**predicate<sub>1</sub>'**(x, (y))] & **BECOME predicate<sub>2</sub>'**(z, x) or (y))  
 f. Causative    α CAUSE β, where α, β are representations of any type

Achievements are punctual, and accomplishments are durative, as are their causative counterparts.

A crucial component of this system is a set of syntactic and semantic tests for determining the class membership of a verb in a particular sentence, since the class of the verb determines its lexical representation or LOGICAL STRUCTURE [LS] (see Van Valin 1993:35, Van Valin & LaPolla 1977, §3.2.2). Examples of English verbs with their LSs are

given in (8).

(8) a. STATES

Pat is a lawyer.	<b>be'</b> (Pat, [ <b>lawyer'</b> ])
The glass is shattered.	<b>shattered'</b> (glass)
Chris is at the house.	<b>be-at'</b> (house, Chris)
Kim saw the message.	<b>see'</b> (Kim, message)

b. ACTIVITIES

The children cried.	<b>do'</b> (children, [ <b>cry'</b> (children)])
The wheel squeaks.	<b>do'</b> (wheel, [ <b>squeak'</b> (wheel)])
Dana ate pizza.	<b>do'</b> (Dana, [ <b>eat'</b> (Dana, pizza)])

c. ACHIEVEMENTS

The glass shattered.	INGR <b>shattered'</b> (glass)
The bubble popped.	INGR <b>popped'</b> (bubble)
Kim noticed the message.	INGR <b>see'</b> (Kim, message)

d. ACCOMPLISHMENTS

The snow melted.	BECOME <b>melted'</b> (snow)
The sky reddened.	BECOME <b>be'</b> (sky, [ <b>red'</b> ])
Leslie learned Korean.	BECOME <b>know'</b> (Leslie, Korean)

e. ACTIVE ACCOMPLISHMENTS

Dana ate the pizza.  
**do'** (Dana, [**eat'** (Dana, pizza)]) & BECOME **eaten'** (pizza)

Chris ran to the house.  
**do'** (Chris, [**run'** (Chris)]) & BECOME **be-at'** (house, Chris)

f. CAUSATIVES

The dog frightens the boy.  
**[do'** (dog, Ø) CAUSE [**afraid.of'** (boy, dog)]<sup>9</sup>

Kim showed Pat the message.  
**[do'** (Kim, Ø) CAUSE [**INCH see'** (Pat, message)]

---

9. 'do'(x, Ø)' indicates that the nature of the causing activity is unspecified.



The girl shattered the glass.

[do' (girl, Ø) CAUSE [INCH shattered' (glass)]]

The sun melted the snow.

[do' (sun, Ø) CAUSE [BECOME melted' (snow)]]

The girl rolled the soccer ball.

[do' (girl, Ø) CAUSE [do' (ball, [roll' (soccer ball)]) ]]

Chris fed Dana the pizza.

[do' (Chris, Ø) CAUSE [ do' (Dana, [eat' (Dana, pizza)]) ] &  
BECOME eaten' (pizza)]

Examination of the verbal systems of a number of languages had led to the conclusion that this set of distinctions is one of the fundamental organizing principles of verbal systems in human language.<sup>10</sup>

The RRG theory of semantic roles is rather different from that of other theories, in that it posits two types of semantic roles. The first are specific thematic relations, the traditional (since Fillmore 1968 and Gruber 1965) notions of agent, theme, patient, experiencer, etc. The second are generalized semantic roles called SEMANTIC MACROROLES; they were introduced in Van Valin (1977b) and have no exact analog in other theories, although Jackendoff's 'action tier' and Dowty's proto-roles bear some resemblance (see Van Valin 1998a for more discussion). Following the ideas of Gruber (1965) and Jackendoff (1976), RRG defines thematic relations in terms of argument positions in LSs such as those in (8)-(9). All thematic relations are defined in terms of argument positions in state and activity LSs; all other LS types are composed of them plus elements like BECOME, INGR and CAUSE, as shown in (7). This is summarized in Table 1, which should not be considered exhaustive.

---

10. or further development of this decompositional system, see Van Valin & Wilkins (1993), Van Valin & Wilkins (1996).

## I. STATE VERBS

### A. Single Argument

1. State or condition **broken'**(x) x=PATIENT
2. Existence **exist'**(x,y) x=ENTITY

### B. two Arguments

1. Pure location **be-LOC'**(x,y) x=LOCATION, y=THEME
2. Perception **hear'**(x,y) x=PERCEIVER, y=STIMULUS
3. Cognition **know'**(x,y) x=COGNIZER, y=CONTENT
4. Desire **want'**(x,y) x=WANTER, y=DESIRE
5. Propositional Attitude  
**consider'**(x,y) x=JUDGER, y=JUDGMENT
6. Possession **have'**(x,y) x=POSSESSOR, y=POSSESSED
7. Internal Experience  
**feel'**(x,y) x=EXPERIENCER, y=SENSATION
8. Emotion **love'**(x,y) x=EMOTER, y=TARGET
9. Attrib/Identificational  
**be'**(x,y) x=ATTRIBUTANT, y=ATTRIBUTE

## II. ACTIVITY VERBS

### A. Single Argument

1. Unspecified action **do'**(x,∅) x=EFFECTOR
2. Motion **do'**(x,[**walk'**(x)]) x=MOVER
3. Static motion **do'**(x,[**spin'**(x)]) x=ST-MOVER
4. Light emission **do'**(x,[**shine'**(x)]) x=L-EMITTER
5. Sound emission **do'**(x,[**gurgle'**(x)]) x=S-EMITTER

### B. One or Two Arguments

1. performance **do'**(x,[**sing'**(x,(y))])  
x=PERFORMER, y=PERFORMANCE
2. Consumption **do'**(x,[**eat'**(x,(y))])  
x=CONSUMER, y=CONSUMED
3. Creation **do'**(x,[**write'**(x,(y))])  
x=CREATOR, y=CREATION

4. repetitive action	<b>do'(x,[sew'(x,(y))])</b> x=EFFECTOR, y=LOCUS
5. Directed perception	<b>do'(x,[see'(x,(y))])</b> x=OBSERVER, y=STIMULUS
6. Use	<b>do'(x,[use'(x,(y))])</b> x=USER, y=IMPLEMENT

Table 1: Definitions of Thematic Relations in terms of LS Argument Positions

Since thematic relations have no independent status, they are really just mnemonics for the argument positions in LSs. That is, 'experiencer' stands for 'the first argument of a two-place state predicate of internal experience', for example. It looks like there is a large number of thematic relations, but in fact there are really only five distinct argument positions which contrast with each other. In verbs that lexicalize agency, e.g. murder, agent is represented by 'DO (x, ...', following Dowty (1979). However, in most cases agent is an implicature related to human effectors with certain types of activity predicates and would not be represented in the LS of the verb. See Holisky (1987), Van Valin & Wilkins (1996).

The second type of semantic role plays a central role in the theory; macroroles acts as the **primary interface** between the LS and syntactic representations. There are **only two macroroles**, ACTOR and UNDERGOER, corresponding to the two primary arguments in a prototypical transitive relation. They are called 'macroroles' because each subsumes a number of **specific thematic relations**; the relationship between the macroroles and the argument positions in LS is captured in the Actor-Undergoer Hierarchy in Figure 10.

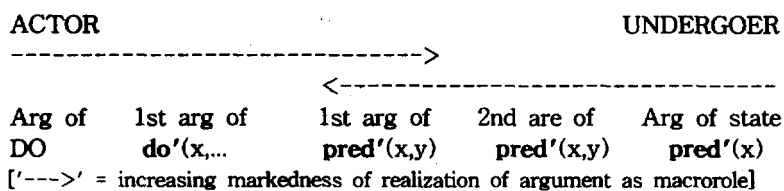


Figure 10: The Actor-Undergoer Hierarchy

Given the LS of a verb, the most agent-like argument will be actor, the most patient-like undergoer, in the default case. Macroroles are not equivalent to grammatical relations, as shown in (9).

- (9) a. Chelswu-nun wuntongcang-kkaci ttwiessta  
       'Chulsoo [actor] ran to the playground.'
- b. Ku os-i mallassta.  
       'The clothes [undergoer] dried.'
- c. Chelswu-ka Swoonhi-lul capessta  
       'Chulsoo [actor] caught Soonhi [undergoer].'
- d. Swoonhi-ka Chelswu-eyuihayeykey caphiessta  
       'Soonhi [undergoer] was caught by Chulsoo [actor].'

The exact role of macroroles in the mapping (or linking) between semantic and syntactic representations will be sketched in 3.4 and summarized in Figure 14.

As mentioned in 3.1, an important component of the theory of complex sentences is the semantic relations that obtain between units in a juncture. These include causality, psych-action, direct perception, cognition, propositional attitude, conditional, and varieties of temporal sequence. These may be ordered into a hierarchy in terms of whether the units in the juncture express facets of a single event, state or action or distinct events, states or actions. This semantic hierarchy interacts with the syntactic hierarchy of juncture-nexus types as

follows: there is an iconic relation between the semantics and syntax of clause linkage, such that the tightness of the syntactic linkage directly reflects the semantic integration of the units in the linkage (cf. Silverstein 1976, Givón 1980, Foley & Van Valin 1984). This is expressed in the Interclausal Relations Hierarchy in Figure 11.

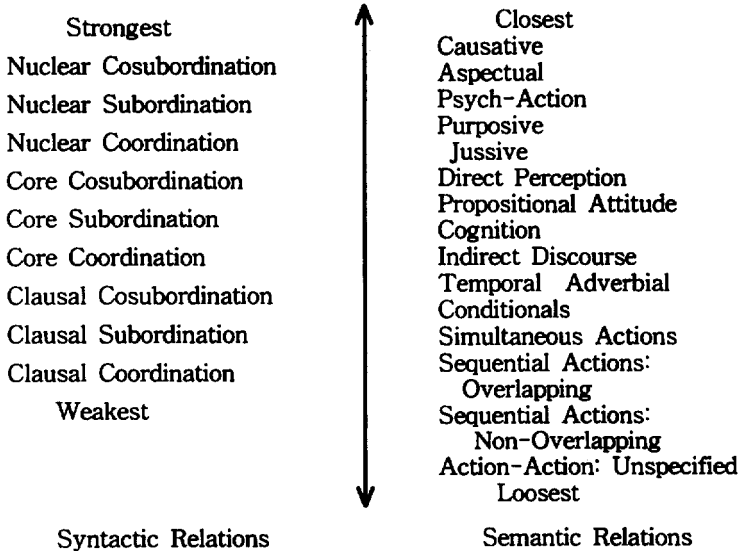


Figure 11: Interclausal Relations Hierarchy

Van Valin & Wilkins (1993) employ this hierarchy, together with an enriched version of the system of lexical representation introduced above, to show how it is possible to predict the syntactic form of certain types of complex sentences from their semantic representations.

### 3.3 The lexicon

RRG is a lexicalist theory, and therefore the lexicon plays a very

important role in it. Lexical entries for verbs are based on LSs; the lexical representation of nouns follows the theory of nominal qualia proposed in Pustejovsky (1995). RRG takes the position that lexical entries for verbs should contain only idiosyncratic information, with as much as possible derived from general lexical principles or rules. Information about transitivity is very important, and RRG defines transitivity in terms of the number of macroroles that a verb takes: transitive = 2, intransitive = 1, atransitive = 0. The general principles in (10) predict the transitivity of regular verbs.

(10) 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 logical structure
  - 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 verbs which take one macrorole,
  - 1. If the verb has an activity predicate in its LS, the macrorole is actor.
  - 2. If the verb has no activity predicate in its LS, the macrorole is undergoer.

In RRG, no syntactic subcategorization information is included in lexical entries; all of the relevant information is derivable from the LS of the verb plus information about its transitivity. Thus these principles have the effect of predicting the syntactic subcategorization of a verb from its semantic representation. See Van Valin (1990, 1991b) for application of this to the analysis of syntactic issues in Italian, Georgian and Icelandic. All theories must stipulate the transitivity of exceptional verbs, and this is done in RRG by specifying their transitivity in terms of [MRa], where 'a' is 0, 1 or 2. Sample lexical entries for some

English verbs are given in (11).<sup>11</sup>

- |         |             |                                      |
|---------|-------------|--------------------------------------|
| (11) a. | kill        | [do' (x, Ø) CAUSE [BECOME dead' (y)] |
| b.      | receive     | INGR have' (x,y)                     |
| c.      | own         | have' (x, y)                         |
| d.      | belong (to) | have' (x, y) [MR1]                   |
| e.      | arrive      | INGR be-at' (x,y) [MR1]              |
| f.      | seem        | seem' (x,y) [MR0]                    |
| g.      | see         | see' (x,y)                           |
| h.      | watch       | do' (x, [see' (x,y)])                |
| i.      | show        | [do' (w, Ø) CAUSE [INGR see' (x,y)]  |
| j.      | run         | do' (x, [run' (x)])                  |
| k.      | drink       | do' (x, [drink' (x, y)])             |
| l.      | melt        | BECOME melted' (x)                   |

The prepositions that mark oblique core arguments can in many instances be predicated from the LS of the verb and therefore need not be listed in the lexical entry (cf. Jolly 1993).

RRG distinguishes lexical from syntactic phenomena in terms of the linking scheme, as will be discussed below. Basically, any process which affects LSs or the arguments therein or the mapping between LSs and macroroles is considered to be lexical. Examples include causativization, regardless of whether it is morphologically unmarked (as in English) or marked (as in Turkish and Chicewa), noun incorporation, the 'dative alternation' (which is analyzed as variable linking to undergoer; cf. Van Valin 1993, Van Valin & LaPolla 1997),

---

11. These are intended as lexical representations. It is well-known that verbs can have constructionally derived *Aktionsart* properties, e.g. *run*, an activity verb, behaves like an accomplishment in *run to the store*. Hence the semantic representation of a core containing a verb may well have a derived LS expanding the LS from the lexical entry of the verb. See Van Valin (1990, 1993), Van Valin & LaPolla (1997) for detailed discussion.

and some types of passivization and antipassivization. Syntactic phenomena involve the mapping between macroroles and the syntactic representation, e.g. some types of passivization and antipassivization, WH-question formation in languages like English, Icelandic and Malagasy, and 'raising' constructions (cf. Van Valin 1993).

### 3.4 Focus structure

The issue of the distribution of information in clauses and sentences was not addressed in Foley & Van Valin (1984), and in Van Valin (1993). Lambrecht's (1986, 1987, 1994) theory of FOCUS STRUCTURE is integrated into RRG. Focus structure is the grammatical system which serves to indicate the scope of the assertion in an utterance in contrast to the pragmatic presupposition, and it is vital to the RRG analysis of many grammatical phenomena. An innovation in RRG is the distinction between the *potential focus domain* [PFD] i.e. the syntactic domain in the sentence where focus may fall, and the *actual focus domain*, i.e. the part that is focussed in a particular sentence. Languages vary in terms of how the PFD is restricted, both in simple sentences and in complex sentences, and this variation underlies important grammatical differences across languages (cf. Van Valin 1993, 1995). The focus structure of an utterance is represented in a distinct projection of the clause from the operator and constituent projections; this is exemplified in Figure 12 for a predicate-focus construction in English. 'Predicate focus' is Lambrecht's term for the traditional 'topic-comment' structure with a topical subject and a focal predicate.



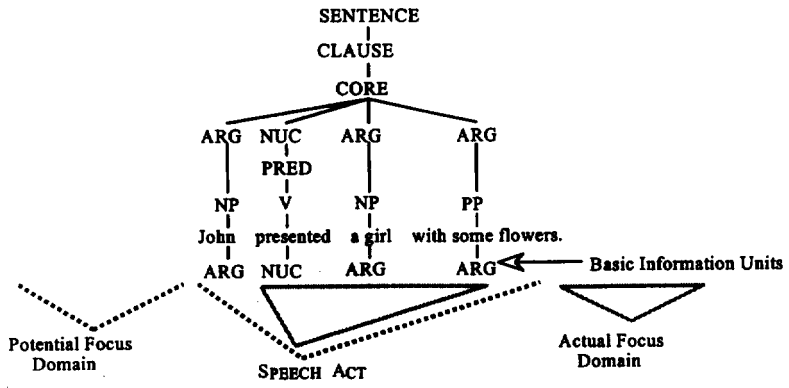


Figure 12: Predicate Focus Construction in English

It is possible to represent all three projections in a single representation, as in Figure 13.

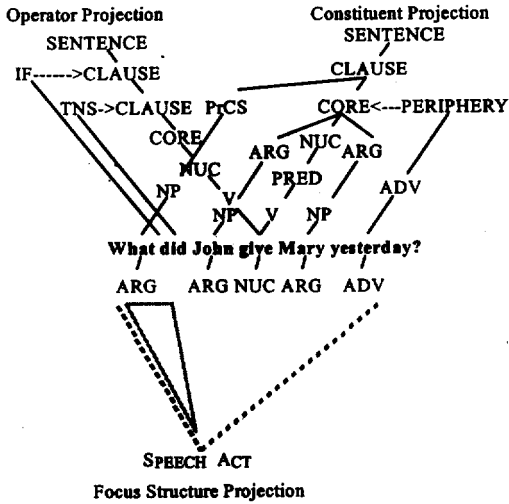


Figure 13: Clause Structure with Constituent, Operator and Focus Structure Projections

### 3.5 Grammatical relations and linking

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 & LaPolla (1997) for detailed discussion.

The linking system relating semantic and syntactic representations is summarized in Figure 14. Syntactic functions like PSA and direct core argument (which are structurally instantiated in the LSC) represent the syntactic pole of the system, while LSs represent the semantic pole. In every language with grammatical relations, there is an accessibility to PSA hierarchy for multiple-argument verbs: in syntactically accusative languages like English and German, it is highest ranking macrorole in terms of the Actor-part of the hierarchy in Figure 10 is the default choice for PSA, whereas in syntactically ergative languages like Dyirbal and Sama (Austronesian, Philippines; Walton 1986), the lowest ranking macrorole is the default choice. That is, in a syntactically accusative language the unmarked choice for the PSA of a transitive verb is the actor, with the undergoer being a marked choice possible only in a passive construction. On the other hand, in a syntactically ergative language, the unmarked choice for the PSA of a transitive verb is the undergoer, with the actor being a marked choice possible only in an antipassive construction. With an intransitive verb, the hierarchy is irrelevant, as the single macrorole functions as PSA regardless of whether it is actor or undergoer.

The overall linking system is summarized in Figure 14. We have

discussed logical structures, macroroles and the hierarchy linking them. This part of the system is universal, in that there is very little cross-linguistic variation; this is the domain of lexical processes, as mentioned in 3.3. Where languages differ substantially is how macroroles and other arguments link into the syntax. The reason the arrows in Figure 14 are double-headed is that the linking system works both from semantics to syntax and from syntax to semantics. In 1 I mentioned the criterion of psychological adequacy and in particular the point made by Bresnan & Kaplan (1982) that theories of linguistic structure should be directly relatable to testable theories of language production and comprehension. 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 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 semantics linking algorithm is required. The details of the linking algorithm are given in Van Valin & LaPolla (1997). It is constrained by the Completeness Constraint, given in (12).

- (12) Completeness Constraint: All of the arguments explicitly specified in the semantic representation of a sentence must be realized syntactically in the sentence, and all of the non-predicate elements in the syntactic representation of a sentence must be linked to an argument position in a logical structure in the semantic representation of the sentence.

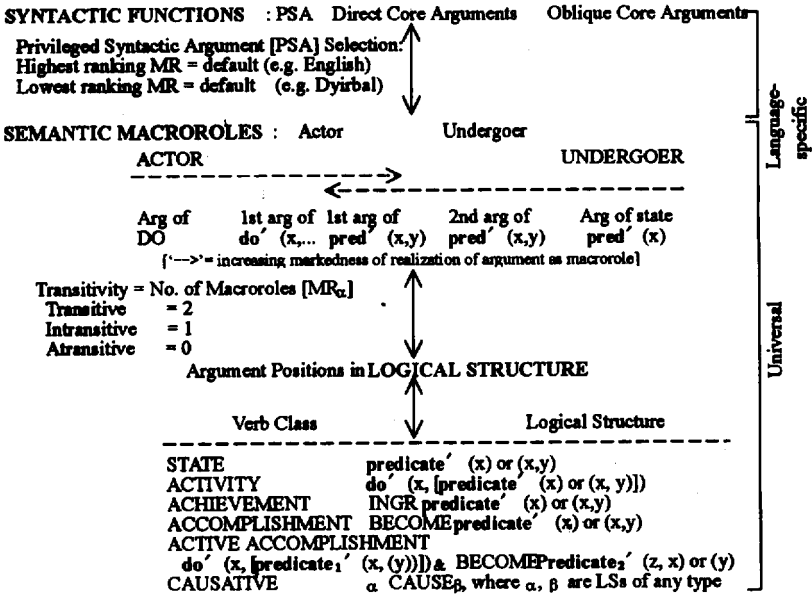


Figure 14: RRG Linking System

Most of what counts as 'syntax' in many theories, e.g. case assignment, agreement, WH-movement, and reflexivization, is handled in RRG in terms of the syntactic phase of the linking. The analysis of reflexivization in RRG follows the approach in Jackendoff (1992) and states the constraints for core-internal ('clause-bound' in other theories) reflexivization at the LS level, not with respect to the syntactic representation. The linking in a WH-question in English is illustrated in Figure 15.

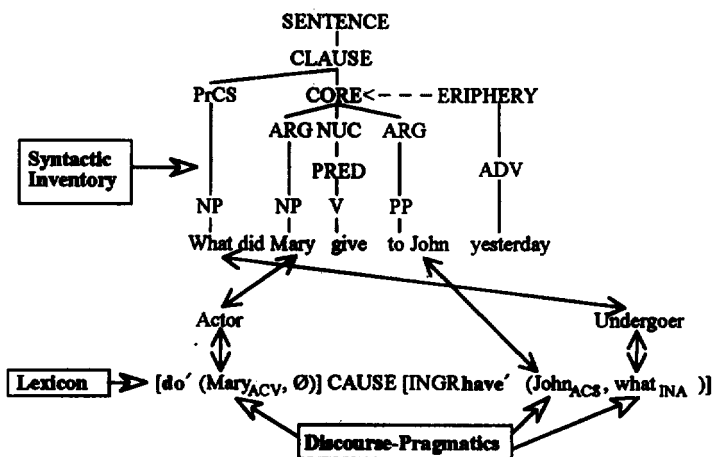


Figure 15: Linking syntax and semantics in a simple sentence in English<sup>12</sup>

The role of discourse-pragmatics in linking will be discussed below. Note that there is a direct linking between the WH-word in the precore slot and the semantic representation. Constraints on WH-question formation and other 'extraction' constructions is explained in terms of the interaction of focus structure and syntax, in particular in terms of restrictions on the potential focus domain (Van Valin 1993, 1995). The case assignment and agreement rules for Icelandic proposed in Van Valin (1991b) are given in (12) as an example; they presuppose an accusative PSA selection hierarchy (cf. Figure 14).

(12) a. Case assignment rules for Icelandic

1. Assign nominative case to the highest ranking macrorole argument.

---

12. The subscripts 'ACV', 'ACS' and 'INA' stand for 'activated', 'accessible' and 'inactive', respectively, and they refer to different cognitive statuses that a referent of the element may have; cf. Lambrecht (1994).

2. Assign accusative case to the other macrorole argument.
  3. Assign dative case to non-macrorole arguments (default<sup>13</sup>).
- b. Finite verb and passive participle/predicate adjective agreement in Icelandic:
1. The finite verb agrees with the highest ranking macrorole argument.
  2. Passive participles and predicate adjectives agree with the undergoer of the predicate of which they are a part in gender, number and case.

These rules account for case marking in simple and complex sentences, including WH-questions, and the agreement rules account for both local and long-distance agreement; see Van Valin (1991b, 1993). The linking in the two Icelandic WH-questions in (13) is illustrated in Figure 16.

- (13) a. Hver-ja            sá            \_laf-ur?  
           who-F.ACC    saw            Olaf-NOM  
           'Who did Olaf see?'
- b. Hver-ja            tók            lögreglan            fast-a?  
           who-F.ACC    took.3pl    the.police.NOM    fast-FsgACC  
           'Who did the police arrest?'

---

13. The idea of dative case as the default case for non-macrorole direct core arguments in languages with morphological case systems is derived from Silverstein (1976, 1981, 1990). Dative is the default case for non-macrorole direct core arguments, and as a default case it may be overridden with certain verbs. See Van Valin (1991b), Michaelis (1993).

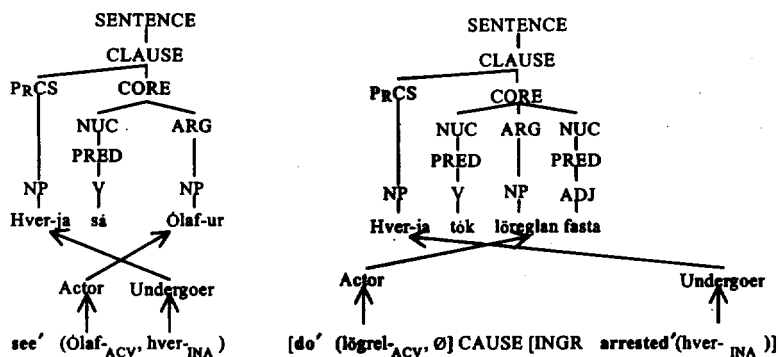


Figure 16: Linking in Icelandic WH-questions

All of the information needed for case marking and agreement is in Figure 16. For (13a), *Ólaf-* is the actor and accordingly will occur in the nominative case; the finite verb will agree with it. *Hver-* is the undergoer and therefore will occur in the accusative case; it makes no difference whether it is linked to a core-internal position or to the precore slot. The same is true with respect to (13b); *lögregl-* is the actor and will therefore occur in the nominative case, and the finite verb will agree with it. *Hver-* is the undergoer and consequently it will appear in the accusative case and the predicate adjective *fast-* will agree with it in gender, number and case. Again, it makes no difference whether it is a non-WH NP linked to a core-internal position or a WH-word linked to the precore slot; the crucial property for case marking and agreement is that it is the undergoer. Thus, the rules for case marking and agreement in (12) can handle non-local case and agreement without modification. Constraints on WH-question formation and related constructions are proposed in Van Valin (1995) and Van Valin & LaPolla (1997).<sup>14</sup>

14. For work on Korean in RRG, see Han (1997, 1998), Lee (1997, 1998), Park (1993, 1995), Song (1988, 1996, 1997), and Yang (1993, 1994, 1996a,b,c, 1997, 1998).

One of the questions which RRG asks is, when there is an option as to which arguments can be linked to PSA, what factors can affect the choice? It turns out that the answer to this question has important typological ramifications, for some languages permit discourse-pragmatic factors to play a role, whereas others do not. In Figure 15 there is input from discourse-pragmatics to the linking. It is represented by the subscripts on the referring expressions filling argument positions in the semantic representation of the sentence (cf. fn. 14). The status of a referent in the discourse context not only influences the form of the expression used to denote it, as is well known, but it may also affect how arguments may be linked into the syntax in some languages. That is, in some languages, but not all, a highly topical (activated) argument tends to appear as PSA, regardless of its semantic function. This has been much discussed in the literature on topic, subject and voice over the past two decades. This distinction is expressed in the RRG typology of 'role-dominated' (no discourse-pragmatic influence on linking) vs. 'reference-dominated' (possible discourse-pragmatic influence on linking) languages, and in the contrast between SYNTACTIC and SEMANTIC PIVOTS, on the one hand, and PRAGMATIC PIVOTS (which are found only in reference-dominated languages), on the other. English, Dyirbal, Malagasy, Sama and Icelandic all have pragmatic pivots in their grammatical system, whereas Lakhota, Warlpiri, Zapotec and Tongan do not. One of the major themes in RRG work is the important role that discourse-pragmatics plays in grammar, and the many ways in which discourse-pragmatics may affect grammatical processes is summarized in Figure 17.



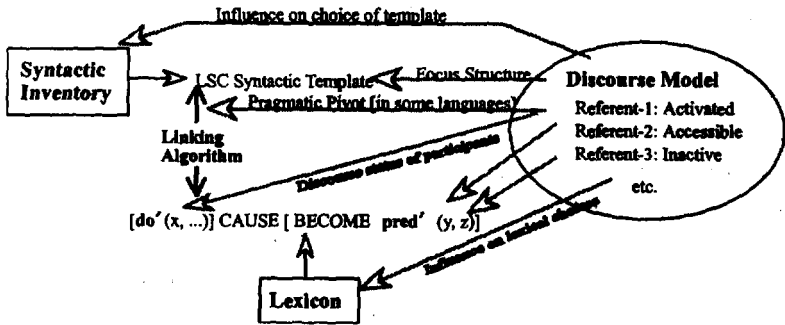


Figure 17: Interaction of discourse-pragmatics and grammar in RRG

The interaction of the three projections of the clause with linking is represented in Figure 18.

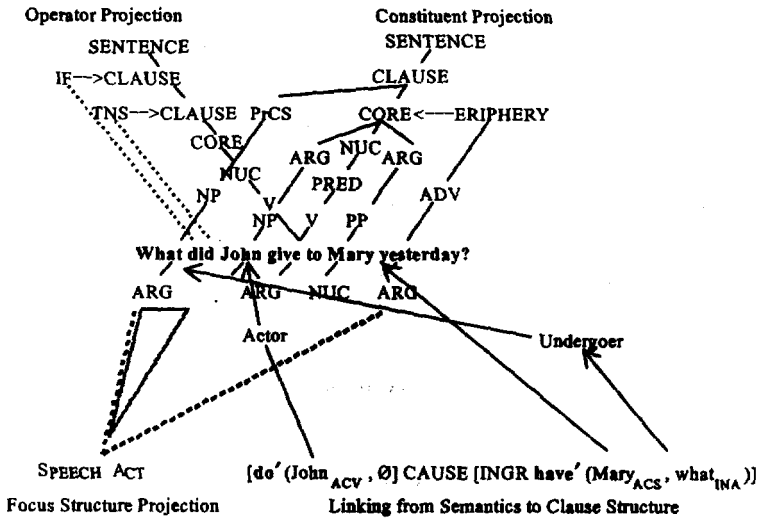


Figure 18: Interaction of linking with clause projections

#### 4. Some implications of RRG

RRG illustrates one possible answer to the questions stated at the beginning of 2, and it shows that it is possible to have a rigorous, typologically-sensitive grammatical theory which takes semantics and pragmatics as central features.

It was mentioned in 1 that Van Valin (1991a) takes a rather strong position with respect to the question of language acquisition, one that is at odds with most other theories, but there is substantial empirical work supporting this view. Braine (1992) shows how a conception of clause structure very much like the layered structure of the clause could be constructed developmentally by the child. Rispoli (1991a,b, 1994, 1995) shows how the lexical representations in 3.2 and the conception of grammatical relations in 3.4 could be learned. Bowerman (1990) provides evidence in favor of the view that rules linking syntactic and semantic representations of the type summarized in Figure 14 are learned, and Van Valin (1994, 1998b) puts forward an account of how some of the constraints on linking between syntactic and semantic representations in complex sentences (i.e. subadjacency) could be learned.

With respect to language purification-based processing models, an attempt to formalize the RRG linking system in a system is just beginning, but not enough work has been done yet for any results to be reported.

#### References

- Bowerman, Melissa. 1990. Mapping semantic roles onto syntactic functions: are children helped by innate linking rules? *Linguistics* 28: 1253-1289.
- Braine, Martin D.S. 1992. What sort of innate structure is needed to 'bootstrap' into syntax? *Cognition* 45:77-100.
- Bresnan, Joan & Ronald Kaplan. 1982. Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan (ed.), *The Mental Representation of Grammatical Relations*, 173-281. Cambridge, MA: MIT

- Press.
- Dik, Simon. 1991. Functional Grammar. F. Droste & J. Joseph (eds.), *Linguistic Theory and Grammatical Description*, 247-74. Amsterdam/Philadelphia: John Benjamins.
- Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- Dowty, David. 1979. *Word Meaning and Montague Grammar*. Dordrecht: Reidel.
- Fillmore, Charles. 1968. The case for case. E. Bach & R. Harms (eds.), *Universals in Linguistic Theory*, 1-88. New York: Holt, Rinehart & Winston.
- Fillmore, Charles & Paul Kay (ms.) *Construction Grammar*. Unpublished ms., UC Berkeley.
- Foley, William & Robert Van Valin, Jr. 1977. On the viability of the notion of 'subject' in universal grammar. *BLS* 3: 293-320.
- 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Franklin, Karl. 1971. *A Grammar of Kewa, New Guinea*. (=Pacific Linguistics C-16). Canberra: Australian National University.
- Gruber, Jeffrey. 1965. *Studies in Lexical Relations*. Unpublished MIT Ph.D. dissertation.
- Han, Jeong-han. 1997. *A typology of Korean lexicalization: serial verb constructions (SVCs)*. M.A. Project, SUNY Buffalo.
- . 1998. *On Grammatical Coding of Information Structure in Korean: A Role and Reference Grammar Account*. Ph.D. dissertation, SUNY Buffalo, in preparation.
- Hansell, Mark. 1993. Serial verbs and complement constructions in Mandarin: a clause linkage analysis. In Van Valin (ed.), 197-233.
- Holisky, Dee A. 1987. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua* 71:103-132.
- Jackendoff, Ray. 1976. Toward an explanatory semantic representation. *Linguistic Inquiry* 7:89-150.
- 1992. Mme. Tussaud meets the Binding Theory. *NLLT* 10.
- Johnson, Mark. 1987. A new approach to clause structure in Role and Reference Grammar. *Davis Working Papers in Linguistics* 2:55-59.
- Jolly, Julia. 1993. Preposition assignment in English. In Van Valin (ed.), 275-310.
- Lambrecht, Knud. 1986. *Topic, Focus and the Grammar of Spoken French*. Unpublished Ph.D. dissertation, Univ. of California, Berkeley.
- 1987. Sentence focus, information structure, and the thematic-categorical distinction. *BLS* 13:366-382.
- 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge

- Univ. Press.
- Lee, Hee-chul. 1997. The Asymmetry of NOM and ACC. Paper presented at the Seventh Harvard International Symposium on Korean Linguistics. Cambridge, MA. July 13, 1997.
- . 1998. The Syntax and Semantics of Korean Periphrastic Causative Constructions. Ph.D. dissertation, SUNY Buffalo.
- Michaelis, Laura. 1993. On deviant case marking in Latin. In Van Valin (ed.), 311-73.
- Olson, Michael. 1981. Barai clause junctures: toward a functional theory of interclausal relations. Unpublished Ph.D. dissertation, Australian National Univ.
- Park, Ki-seong. 1993. Korean causatives in Role and Reference Grammar. M.A. thesis, SUNY Buffalo.
- . 1995. The semantics and pragmatics of case marking in Korean: A Role and Reference Grammar Account. Unpublished Ph.D. dissertation, SUNY Buffalo.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.
- Rispoli, Matthew. 1991a. The acquisition of verb subcategorization in a functionalist framework. *First Language* 11:41-63.
- . 1991b. The mosaic acquisition of grammatical relations. *Journal of Child Language* 18:517-52.
- . 1994. Structural dependency and the acquisition of grammatical relations. Y. Levy (ed.), *Other Children, Other Languages: Issues in the Theory of Language Acquisition*, 265-301. Hillsdale, NJ: LEA.
- . 1995. Missing arguments and the acquisition of predicate meanings. M. Tomasello & W. Merriman (eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs*. Hillsdale, NJ: LEA.
- Shimojo, Mitsuaki. 1995. Focus Structure and Morphosyntax In Japanese: Wa and Ga, and Word Order Flexibility. Ph.D. dissertation, SUNY at Buffalo.
- Silverstein, Michael. 1976. Hierarchy of Features and Ergativity. R.M.W. Dixon (ed.), *Grammatical Categories in Australian Languages*, 112-71. Canberra: Australian Institute of Aboriginal Studies.
- . 1981. Case marking and the nature of language. *Australian Journal of Linguistics* 1:227-246.
- . 1993. Of nominatives and datives: universal grammar from the bottom up. In Van Valin (ed.), 465-98.
- Song, Jae Jung. 1988. Clause Linkage in Korean Paraphrastic Causative and Purposive Constructions. *Language Research* 24:583-606.
- . 1996. Korean Periphrastic Causative Constructions in Role and Reference Grammar: A Rejoinder to Park (1993) and Yang (1994). *Language*

- Research 665-689.
- 1997. Causatives and Causation: A Universal Typological Perspective. Longman Linguistics Library. Longman Publishing Group.
- Van Valin, Robert, Jr. 1977a. Ergativity and the universality of subjects. *CLS* 13:689-706.
- 1977b. Aspects of Lakhota syntax. Unpublished Ph.D. dissertation, Univ. of California, Berkeley.
- 1980. On the distribution of passive and antipassive constructions in universal grammar. *Lingua* 50:303-27.
- 1981. Grammatical relations in ergative languages. *Studies in Language* 5:361-94.
- 1985. Case marking and the structure of the Lakhota clause. J. Nichols & A. Woodbury (eds.), *Grammar Inside and Outside the Clause*, 363-413. Cambridge: Cambridge University Press.
- 1987. The role of government in the grammar of head-marking languages. *International Journal of American Linguistics*, 53:371-397.
- 1990. Semantic parameters of split intransitivity. *Language* 66:221-260.
- 1991a. Functionalist linguistic theory and language acquisition. *First Language* 11:7-40.
- 1991b. Another look at Icelandic case marking and grammatical relations. *Natural Language and Linguistic Theory* 9:145-194.
- 1993. A synopsis of Role and Reference Grammar. In R. Van Valin (ed.), *Advances in Role and Reference Grammar*, 1-164. Amsterdam / Philadelphia: John Benjamins.
- 1994. Extraction restrictions, competing theories and the argument from the poverty of the stimulus. S. Lima et al., (eds.), *The Reality of Linguistic Rules*, 243-59. Amsterdam/ Philadelphia: John Benjamins.
- 1995. Toward a functionalist account of so-called 'extraction constraints'. B. Divriendt et al. (eds.), *Complex Structures: A Functionalist Perspective*, 29-60. Berlin: Mouton de Gruyter.
- 1998a. Generalized semantic roles and the syntax-semantics interface. To appear in the Proceedings of the Colloque de Syntaxe et Sémantique de Paris 97, in press.
- 1998b. The acquisition of WH-questions and the mechanisms of language acquisition. M. Tomasello, ed., *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, 221-249. Hillsdale, N.J.: LEA.
- Van Valin, Robert, Jr. & William Foley. 1980. Role and Reference Grammar. E. Moravcsik & J. Wirth (eds.), *Current Syntactic Theories (Syntax & semantics 13)*, 329-52. New York: Academic Press.

- Van Valin, Robert, Jr. & Randy LaPolla. 1997. *Syntax: Structure, Meaning & Function*. Cambridge: Cambridge University Press.
- Van Valin, Robert, Jr. & David Wilkins. 1993. Predicting syntactic structure from semantic representations: remember in English and its equivalents in Mparntwe Arrernte. In Van Valin (ed.), 499-534.
- 1996. The case for 'effector': Case roles, agents and agency revisited. In Shibatani & Thompson, eds., *Grammatical constructions*, 289-322. Oxford: Oxford University Press.
- Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell Univ. Press.
- Walton, Charles. 1986. *Sama Verbal Semantics: Classification, Derivation and Inflection*. Manila: Linguistic Society of the Philippines.
- Watters, James. 1993. An investigation of Turkish clause linkage. In Van Valin (ed.), 535-60.
- Yang, Byong-seon. 1993. Clause structure and NP accessibility of Korean internally headed relative clause in Role and Reference Grammar. In S. Kuno, eds., *Harvard Studies in Korean Linguistics V*, 230-38. Seoul: Hanshin Publishing Co.
- 1994. Morphosyntactic phenomena of Korean in Role and Reference Grammar: Psych-verb constructions, inflectional verb morphemes, complex sentences, and relative clauses. PhD dissertation, SUNY Buffalo. Published by Hankuk Publishers, Seoul, 1994.
- 1996a. Syntax-Semantics Interface in Psych-Verb Constructions: A Role and Reference Grammar Approach. *Modern Grammar* 7:171-207.
- 1996b. A Role and Reference Grammar Account of Unaccusativity: Split Intransitivity. *Linguistics* 4.77-93.
- 1996c. An Overview of Unaccusativity and the Issues. *Humanities Research* 2.139-161.
- 1997. A Role and Reference Grammar Account of Verb Complementation and Causations in Korean. Paper presented at the Seventh Harvard International Symposium on Korean Linguistics. Cambridge, MA. July 13, 1997.
- 1998. *Role and Reference Grammar Kaylon* [Introduction to Role and Reference Grammar]. Seoul: Hankuk Publishers.

Robert D. Van Valin, Jr.

Department of Linguistics & Center for Cognitive Science  
State University of New York at Buffalo