A Brief Overview of Role and Reference Grammar

Robert D. Van Valin, Jr.
Department of Linguistics & Center for Cognitive Science
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 ôcompatible with the results of psycholinguistic research on the acquisition, processing, production, interpretation and memorization of linguistic expressionsö(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 SaussureÆs 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.¹

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-1970ÆEs in some strands of American linguistics. DixonÆEs grammar of Dyirbal and Schachter & OtanesÆs 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 æsubjectÆ, and theoretical issues, e.g. the relation between

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

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.

<table>
<thead>
<tr>
<th>Semantic Element(s)</th>
<th>Syntactic Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate</td>
<td>Nucleus</td>
</tr>
<tr>
<td>Argument in semantic representation of predicate</td>
<td>Argument in Core argument</td>
</tr>
<tr>
<td>Non-arguments</td>
<td>Periphery</td>
</tr>
<tr>
<td>Predicate + Arguments</td>
<td>Core</td>
</tr>
<tr>
<td>Predicate + Arguments + Non-arguments</td>
<td>Clause (= Core + Periphery)</td>
</tr>
</tbody>
</table>

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

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], 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. In the formal representation of the LSC (proposed in Johnson

---

2It is assumed that noun phrases and adpositional phrases have a comparable layered structure; see Van Valin (1993), *1.7, Van Valin & LaPolla (1997), *2.3. Operators in the NP include determiners, quantifiers and adjectival and nominal modifiers.
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.

Figure 2: The Layered Structure of the Clause

---

3The periphery has been omitted from this diagram for the sake of simplicity.
Did is labelled both as tenseÆ and as IFÆ 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.
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.5

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).

---

5The 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 be concerned with the issue of (morpho)phonological representation.
(2)  

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

- b. John forced open the door. [two nuclei, \textit{push and open}, in a single core]

- c. Taì qiao pō le yì ge faênwaên. [two nuclei, \textit{qiao ðhit and p ðbreak}, in a single core] (Hansell 1993)
  3sg hit break PRFV one CL bowl
  \textit{He broke (by hitting) a ricebowl.}

---

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; \textit{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)  

  1sg let.FUT John eat the cakes
  \textit{I will let John eat the cakes.}

- b. I ordered Fred to force the door open.
c.  Taı́ jiáo woË xìÉe zìÈ.
3sg teach 1sg write characters
æShe teaches me to write characters.Æ

The RRG theory of complex sentences is equally concerned with the set of possible syntactic and semantic relations between the units in a juncture. 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 OlsonÆs (1981) analysis of clause linkage in Barai (a Papuan language), postulates three nexus types: coordination, subordination, and COSUBORDINATION, which is in essence 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 (2c), aspect obligatorily has scope over both nuclei, and therefore the nexus is cosubordination. This is represented as in Figure 8.

---

6The semantic relations will be discussed in §3.2 below.
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  g÷r-meli-yiz. Core cosubordination  
\( \text{go-CMPL see-MODAL-1pl} \)  
\( \text{æWe ought to go and see.Æ} \)
b. Mzik dinle-yerek, uyu-yabil-i r-im. Core coordination  
\( \text{music listen-CMPL sleep-MODAL-AOR-1sg} \)  
\( \text{æListening to music, I can sleep.Æ} \)
In (4a), the modal operator -mElI- æoughtÆ has scope over both cores, and accordingly the nexus is cosubordinate; in (4b), on the other hand, the modal operator -yAbil- æableÆ has scope only over the final core, hence coordinate nexus. The following examples from Kewa (Franklin 1971) are a minimal triple for the three nexus types at the level of clausal juncture.

(5) a. Nipu‰ ï‰pu-la             pare nï‰  paala‰  na-pï‰a. Coordination
   3sg come-3sgPRES but 1sg afraid NEG-be.1sgPRES
   æHe is coming, but I am not afraid.Æ

   b. (Ni‰) EÂpo la‰-ri           e‰pa-wa. Cosubordination
      1sg whistle say-SIM.SS come-1sgPAST
      æI whistled while I came,Æ or æI came whistling.Æ

   c. (Ni‰) EÂpo la‰-lo-pulu                   irikai e‰pa-lia. Subordination
      1sg whistle say-1sgPRES-CAUSAL dog come-3sgFUT
      æBecause I am whistling, the dog will come.Æ

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 LisaÆs losing her job shocked everyone instantiate core subordination in English. The nine juncture-nexus types may be

---

7The term æcoordinationÆ 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 æX conj YÆ, which may be one of the formal instantiations of coordinate nexus.
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 VendlerÆEs (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).\(^8\)

(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  \(\text{predicate} \ (x) \text{ or } (x,y)\)
   b. Activity  \(\text{do} \ (x, \ [\text{predicate} \ (x) \text{ or } (x, y)])\)
   c. Achievement  \(\text{INGR(ESSIVE) predicate} \ (x) \text{ or } (x, y), \text{ or}\)
   \(\text{INGR do} \ (x, \ [\text{predicate} \ (x) \text{ or } (x, y)])\)
   d. Accomplishment  \(\text{BECOME predicate} \ (x) \text{ or } (x,y), \text{ or}\)
   \(\text{BECOME do} \ (x, \ [\text{predicate} \ (x) \text{ or } (x, y)])\)
   e. Active accomplishment  \(\text{do} \ (x, \ [\text{predicate}_{1} \ ((x, (y)))] \text{ & BECOME predicate}_{2} \ (z, x) \text{ or } (y))\)
   e. Causative  \(\alpha \text{ CAUSE } \beta, \text{ where } \alpha, \beta \text{ are representations of any type}\)

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 is 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.  \(\text{be} \ (Pat, [\text{lawyer }])\)
   The glass is shattered.  \(\text{shattered} \ (\text{glass})\)
   Chris is at the house.  \(\text{be-at} \ (\text{house, Chris})\)

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

**b. ACTIVITIES**

The children cried.  

The wheel squeaks.  

Dana ate pizza.

**c. ACHIEVEMENTS**

The glass shattered.  

The bubble popped.  

Kim noticed the message.

**d. ACCOMPLISHMENTS**

The snow melted.  

The sky reddened.  

Leslie learned Korean.

**e. ACTIVE ACCOMPLISHMENTS**

Dana ate the pizza.  

Chris ran to the house.  

**f. CAUSATIVES**

The dog frightens the boy.

Kim showed Pat the message.

The girl shattered the glass.

The sun melted the snow.

The girl rolled the soccer ball.

Chris fed Dana the 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.  

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 & Wilkins (1993), Van Valin & Wilkins (1996).

---

9. Edo (x, )E indicates that the nature of the causing activity is unspecified.

10. For further development of this decompositional system, see Van Valin & Wilkins (1993), Van Valin & Wilkins (1996).
Valin (1977b) and have no exact analog in other theories, although JackendoffÆs æaction tierÆ and DowtyÆs 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.

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 do (x, [sew (x, (y))]) x = EFFECTOR, y = LOCUS
      5. Directed perception do (x, [see (x, (y))]) x = OBSERVER, y = STIMULUS
      6. Use do (x, [use (x, (y))]) 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, æexperiencerÆ stands for æthe first argument of a two-place state predicate of internal experienceÆ, 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 æDO (x, ...Æ, 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 æmacrorolesÆ 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.

\[
\begin{array}{c}
\text{ACTOR} \\
\text{—————>} \\
\text{DO} \\
\text{do}^\circ (x,...) \\
\text{pred}^\circ (x,y) \\
\text{pred}^\circ (x) \\
\end{array}
\begin{array}{c}
\text{UNDERGOER} \\
\text{<—————} \\
\text{Arg of state} \\
\text{Arg of 2nd arg of} \\
\text{1st arg of} \\
\text{1st arg of} \\
\end{array}
\]

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. Chris [SUBJ-ACTOR] ran to the store.
    c. Pat [SUBJ-ACTOR] grabbed Dana [DIRECT OBJECT-UNDERGOER].
    d. Dana [SUBJECT-UNDERGOER] was grabbed by Pat [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, Givn 1980, Foley & Van Valin 1984). This is expressed in the Interclausal Relations Hierarchy in Figure 11.
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.

\[
\text{(10) Default Macrorole Assignment Principles}
\]

- **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 $[\text{MR}\alpha]$, where $\alpha\in\{0, 1, 2\}$. Sample lexical entries for some English verbs are given in (11).\footnote{These are intended as lexical representations. It is well-known that verbs can have constructionally derived \textit{Aktionsart} properties, e.g. run, an activity verb, behaves like an accomplishment in \textit{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.}

\begin{enumerate}
\item \textit{kill} $[\text{do} \ (x,\ )]\text{CAUSE} [\text{BECOME dead (y)}]$
\item \textit{receive} \hspace{1cm} \text{INGR have (x,y)}
\item \textit{own} \hspace{1cm} \text{have (x, y)}
\item \textit{belong (to)} \hspace{1cm} \text{have (x, y) [MR1]}
\item \textit{arrive} \hspace{1cm} \text{INGR be-at (x,y) [MR1]}
\item \textit{seem} \hspace{1cm} \text{seem (x,y) [MR0]}
\item \textit{see see} (x,y)
\item \textit{watch} \hspace{1cm} \text{do (x, [see (x,y)])}
\item \textit{show} \hspace{1cm} \text{do (w, )\text{CAUSE} [\text{INGR see (x,y)}]}
\item \textit{run do} (x, [\text{run (x)})]
\item \textit{drink} \hspace{1cm} \text{do (x, [drink (x, y)])}
\item \textit{melt} \hspace{1cm} \text{BECOME melted (x)}
\end{enumerate}

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 Chice\textsubscript{w-a}), noun incorporation, the \textit{ædative alternation} (which is analyzed as variable linking to undergoer; cf. Van Valin 1993, Van Valin & LaPolla 1997), 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 \textit{æraisingÆ} constructions (cf. Van Valin 1993).

\subsection{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) LambrechtÆs (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 \textit{potential focus domain} [PFD] i.e. the syntactic domain in the sentence where focus may fall, and the \textit{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. æPredicate focusÆ is LambrechtÆs term for the traditional ætopic-commentÆ structure with a topical subject and a focal predicate.

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

Figure 12: Predicate Focus Construction in English

It is possible to represent all three projections in a single representation, as in Figure 13.
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 $\varnothing$ 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.
Most of what counts as æsyntaxÆ 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 (æclause-boundÆ 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.
What did Mary give to John yesterday?

Figure 15: Linking syntax and semantics in a simple sentence in English
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 æextractionÆ 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.
    2. Assign accusative case to the other macrorole argument.
    3. Assign dative case to non-macrorole arguments (default\(^{13}\)).

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 sß laf-ur?
    who-F.ACC saw Olaf-NOM
    æWho did Olaf see?Æ

b. Hver-ja tk l÷greglan fast-a?
    who-F.ACC took.3pl the.police.NOM fast-FsgACC
    æWho did the police arrest?Æ

\(^{12}\)The subscripts æACVÆ, æACSÆ and æINAÆ stand for æactivatedÆ, æaccessibleÆ and æinactiveÆ, respectively, and they refer to different cognitive statuses that a referent of the element may have; cf. Lambrecht (1994).

\(^{13}\)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), \textit{lafl-} is the actor and accordingly will occur in the nominative case; the finite verb will agree with it. \textit{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); \textit{lögregl-} is the actor and will therefore occur in the nominative case, and the finite verb will agree with it. \textit{Hver-} is the undergoer and consequently it will appear in the accusative case and the predicate adjective \textit{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).

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 arole-

---

dominatedÆ (no discourse-pragmatic influence on linking) vs. æreference-dominatedÆ (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.

The interaction of the three projections of the clause with linking is represented in Figure 18.
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. subjacency) could be learned.

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


