CROSS-LINGUISTIC PATTERNS OF LINKING

Robert D. VAN VALIN, Jr.
State University of New York at Buffalo

RESUMO

Este artigo apresentará um número de generalizações inter-lingüísticas sobre a conexão de argumentos, tomando por base uma teoria conceptualmente baseada na gramática de papéis, referência e conexão [linking, Role and Reference Grammar [RRG], embora não estruturalmente baseada. Será mostrado que há aspectos universais da conexão que são prontamente capturados neste arcabouço e que a variação específica de cada língua é bastante limitada e restringe-se a um conjunto limitado de possibilidades. O fenômeno da conexão (dos argumentos) a ser discutido neste artigo inclui: alternações de voz, verbos com estrutura direta (direct verbs) e com estrutura inversa (inverse verbs) (ex. German haben 'ter' vs. gehören 'pertencer a') e construções aplicativas em línguas acusativas e ergativas. Argumentar-se-á que é necessário fazer uma distinção entre aspectos sintáticos e lexicais da conexão (de argumentos). A fase lexical da conexão é virtualmente universal. A fase sintática é o lócus do contraste entre sistemas ergativo-absolutivos, nominativo-acusativos e outros sistemas. Mostrar-se-á que as propriedades das diferentes construções de voz, – ex. passivas "simples' vs passivas adversativas, como por exemplo no japonês – são funções condicionadas pelo fato de essas propriedades pertencerem ora a uma fase lexical ora a uma fase sintática da conexão. Será também mostrado que as construções aplicativas nas línguas sintaticamente ergativas funcionam do mesmo modo que em suas contrapartes sintaticamente acusativas.

ABSTRACT

This paper will present a number of cross-linguistic generalizations about argument linking framed within a conceptually based, rather than structurally based, theory of linking, Role and Reference Grammar [RRG] (Van Valin 1993, 2005, Van Valin & LaPolla 1997). It will be shown that there are a universal aspects of linking which are readily captured in this framework and that the language-specific variation is quite limited and falls within a narrow range of possibilities. The linking phenomena to be discussed in the paper include: voice alternations, direct vs. inverse verbs (e.g. German haben 'have' vs. gehören 'belong to') and applicative constructions in syntactically accusative and ergative languages. It will be argued that it is necessary to make a distinction between lexical and syntactic aspects of linking. The lexical phase of the linking is virtually universal. The syntactic phase is the locus of the contrast between ergative, accusative and other syntactic systems. Properties of different voice constructions, e.g. Japanese 'plain' vs. adversative passives, will be shown to be a

function of whether they are part of the lexical or the syntactic phase of the linking. It will also be shown that applicative constructions in syntactically ergative languages work the same way as their counterparts in syntactically accusative languages.

PALAVRAS-CHAVE
Construções aplicativas, ergatividade sintática, gramática de referência e papel temático, conexão, voz, papéis semânticos, decomposição lexical

KEYWORDS
Applicative constructions, syntactic ergativity, role and reference grammar, linking, voice constructions semantic roles, lexical decomposition

Introduction

The relationship between the semantic argument structure of a verb (or other predicate) and the syntactic realization of those arguments is a major topic of interest in contemporary linguistic theory. This paper presents a number of cross-linguistic generalizations about argument linking within a conceptually based, rather than structurally based, theory of linking, Role and Reference Grammar [RRG] (Van Valin, 2005; Van Valin and LaPolla, 1997). It will be shown that there are universal aspects of linking which are readily captured in this framework and that the language-specific variation is quite limited and falls within a narrow range of possibilities.

The discussion will proceed as follows. In section 2 the relevant aspects of RRG will be summarized. In section 3, an analysis of voice in both ergative and accusative syntactic systems will be given. In section 4 universal vs. language-specific aspects of linking will be investigated. Conclusions will be presented in section 5.

1. The Linking System in Role and Reference Grammar

The general organization of RRG is presented in (1).

SYNTATIC REPRESENTATION

↑

Lynking Algorithm

↓

SEMANTIC REPRESENTATION

Discourse Pragmatics
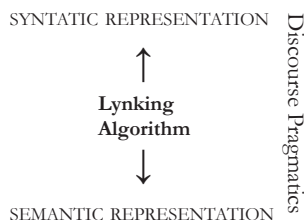
(1) Organization of Role and Reference Grammar

RRG posits a direct mapping between the semantic representation of a sentence and its syntactic representation; there are no intermediate levels of representation such as 'D-structure' or syntactic argument structure. It is a truly 'minimalist' theory. The main focus in this paper will be on the semantic representation and the linking algorithm; the nature of the syntactic representation and the role of discourse-pragmatics in linking not be discussed. See Van Valin (2005) for a detailed presentation.

1.1. The semantic representation of sentences

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

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

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

Syntactic and semantic tests determine the *Aktionsart* of a clause (see VV §2.2.1). As the sentences in (3e-f́) show, a single verb, e.g. *walk*, can have more than one *Aktionsart* interpretation. This verb would be listed in the lexicon as an activity verb, and lexical rules would derive the other uses from the basic activity use (see Van Valin (in press a)).

The system of lexical decomposition builds on the one proposed in Dowty (1979). Unlike Dowty's scheme, the RRG system treats both state and activity predicates as basic. The lexical representation of a verb or other predicate is termed its LOGICAL STRUCTURE [LS]. State predicates are represented simply as **predicate´**, while all activity predicates contain **do´.** Accomplishments, which are durative, are distinguished

from achievements, which are punctual. Accomplishment LSs contain BECOME, while achievement LSs contain INGR, which is short for 'ingressive'. Semelfactives contain SEML. In addition, causation is treated as an independent parameter which crosscuts the six basic and derived *Aktionsart* classes, hence the twelve classes in (3). It is represented by CAUSE in LSs. The lexical representations for each type of verb in (3) are given in (4).

(4) Lexical representations for *Aktionsart* classes

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

Examples of simple English sentences with the LS of the predicate are presented in (5).

(5) a. STATES
    The window is shattered.
      shattered′ (window)
    Fred is at the house.
      be-at′ (house, Fred)
    John saw the picture.
      see′ (John, picture)

b. ACTIVITIES

The children cried.

    do′ (children, [**cry′** (children)])

John ate fish.

    do′ (John, [**eat′** (John, fish)])

c. ACHIEVEMENTS

The window shattered.

    INGR **shattered′** (window)

The balloon popped

    INGR **popped′** (balloon)

d. SEMELFACTIVES

The light flashed.

    SEML **do′** (light, [**flash′** (light)])

John glimpsed the picture.

    SEML **see′** (John, picture)

e. ACCOMPLISHMENTS

The snow melted.

    BECOME **melted′** (snow)

Mary learned French.

    BECOME **know′** (Mary, French)

f. ACTIVE ACCOMPLISHMENTS

John ate the fish.

    do′ (John, [**eat′** (John, fish)]) & BECOME **eaten′** (fish)

The dog walked to the park.

    do′ (dog, [**walk′** (dog)]) & BECOME **be-at′** (park, dog)

g. CAUSATIVES

Mary showed John the picture.

    [**do′** (Mary, Ø)] CAUSE [BECOME **see′** (John, picture)]

The boy shattered the window.

    [**do′** (boy, Ø)] CAUSE [INGR **shattered′** (window)]

The sun melted the snow.

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

The girl walked the dog.
  [**do´** (girl, Ø)] CAUSE [**do´** (dog, [**walk´** (dog)])]
h. CAUSATIVE ACTIVE ACCOMPLISHMENTS
  The girl walked the dog to the park.
  [**do´** (girl, Ø)] CAUSE [**do´** (dog, [**walk´** (dog)]) & INGR **be-at´** (park, dog)]

Full semantic representations of sentences also contain lexical representations of the NPs, adjuncts, and grammatical operators like tense and aspect; see Van Valin (2005, §2.2, 2.3).

1.2 Semantic macroroles and lexical entries for verbs

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

(6) a. Maria fechou a janela.
      'Maria [Actor] closed the window [Undergoer].'
    b. A janela foi fechada por Maria.
      'The window [Undergoer] was closed by Maria [Actor].'
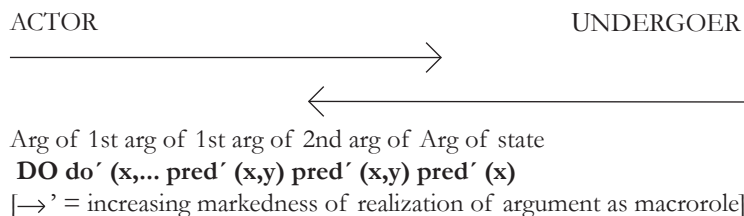   c. Maria cantou.
      'Maria [Actor] sang.'

    d. Maria morreu.
       'Maria [Undergoer] died.'

In (6a), *María* is the actor and *a janela* 'the window' is the undergoer of the transitive verb *fechar* 'sell'; in the sentences with intransitive verbs, *María* is an actor with the activity verb *cantar* 'sing' and an undergoer with the accomplishment verb *morrer* 'die'. Actor is not equivalent to syntactic subject, nor is undergoer equivalent to syntactic direct object, as the examples in (6c) and crucially (6d) show: in both of these sentences the syntactic subject is an undergoer, and in the passive sentence in (6d) the actor is an oblique adjunct. In an English clause with an active voice transitive verb, the actor is the initial NP (the traditional subject) and the undergoer is the direct NP immediately following the verb.

    Actor and undergoer are generalizations across specific semantic argument types, as defined by LS positions. This is illustrated in (7).

(7) Macroroles as generalizations over specific argument types

kill	**[do´** (x, Ø)] CAUSE [BECOME **dead´** (y)]
ee	**see´** (x, y)
put	**[do´** (x, Ø)] CAUSE [BECOME **be-LOC´** (y, z)]
present	**[do´** (x, Ø)] CAUSE [BECOME **have´** (y, z)]
	↓ ↓
	Actor Undergoer

The x argument of all of these verbs functions as the actor, regardless of whether it is the first argument of the generalized activity verb **do´** (conventionally labeled 'effector'), as with *kill*, *put* and *present*, or the first argument of a two-place state predicate, as with *see*. With two-place transitive verbs like *kill* and *see*, the y argument is the undergoer. With three-place verbs like *put* and *present*, the situation is potentially more complex. *Put* allows only the z argument to be undergoer (in (5a, e) with locational predicates the first argument is the location and the second

argument is the located entity), as in *Chris* [*x*, actor] *put the book* [*z*, undergoer] *on the table* [*y*]. *Present*, on the other hand, permits either the *y* or the *z* argument to function as undergoer; these two possibilities are given in (8); (7) depicts (8b), with the *y* argument as undergoer.

(8) a. Chris [*x*, actor] presented the award [*z*, undergoer] to Pat [*y*].
    b. Chris [*x*, actor] presented Pat [*y*, undergoer] with the award [*z*].

The relationship between LS argument positions and macroroles is captured in the Actor-Undergoer Hierarchy [AUH] in (9).

(9) The Actor-Undergoer Hierarchy[1]

ACTOR   UNDERGOER

$\longrightarrow$

$\longleftarrow$

Arg of 1st arg of 1st arg of 2nd arg of Arg of state
DO do′ (x,... pred′ (x,y) pred′ (x,y) pred′ (x)
[→' = increasing markedness of realization of argument as macrorole]

The basic idea of the AUH is that in a LS the leftmost argument in terms of the hierarchy will be the actor and the rightmost will be the undergoer. This was true for *kill*, *see* and *put* in (7). It was not true for *present*, however, and this reflects a fundamental asymmetry in the AUH: the leftmost argument in a LS (in terms of the AUH) is always the actor, but the rightmost argument is only the default choice for undergoer. There are languages in which undergoer selection follows a different principle, so-called 'primary object languages'; in such languages, the second highest ranking argument is chosen as the undergoer. See Guerrero and Van Valin (2004) for detailed discussion.

Transitivity in RRG is defined semantically in terms of the number of macroroles a predicate takes. This is termed 'M-transitivity' in RRG, following Narasimhan (1998), in order to distinguish it from the number of syntactic arguments a predicate takes, its 'S-transitivity. The three

M-transitivity possibilities are: transitive (2 macroroles), intransitive (1 macrorole), and atransitive (0 macroroles). The principles determining the M-transitivity of verbs are given in (10).

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

These principles correctly predict the transitivity of the verbs from Brazilian Portuguese and English discussed above. (The LSs for the Brazilian Portuguese verbs are [**do´** (x, Ø)] CAUSE [BECOME **closed´** (y)] for *fechar* 'close', **do´** (x, [**sing´** (x)]) for *cantar* 'sing', and BECOME **dead´** (x) for *morrer* 'die'.) *Fechar* 'close', like English *kill, see, put* and *present*, has at least two arguments in its LS, and therefore it is transitive, following (10a1). *Cantar* 'sing' and *morrer* 'die' both have only one argument in their LSs, and accordingly they are intransitive, following (10a2). If a verb takes only one macrorole, then the principles in (10b) come into play. *Cantar* 'sing' has an activity predicate in its LS, and therefore its macrorole is actor, following (10b1). *Morrer* 'die', on the other hand, has no activity predicate in its LS, and consequently its macrorole is undergoer, following (10b2). If a verb is irregular and has exceptional transitivity, it will be indicated in its lexical entry by '[MR<]', where '<' is a variable for the number of macroroles. Examples of lexical entries for some English verbs are given in (11).

(11) a. *kill*   [**do**´ (x, Ø)] CAUSE [BECOME **dead**´ (y)]

     b. *receive*   BECOME **have**´ (x, y)

     c. *o*wn   have´ (x, y)

     d. belong (to)   have´ (x, y) [MR1]

     e. see   see´ (x, y)

     f. watch   do´ (x, [see´ (x, y)])

     g. show   [do´ (w, Ø)] CAUSE [BECOME see´ (x, y)]

     h. run   do´ (x, [run´ (x)])

     i. drink   do´ (x, [**drink**´ (x, y)])

A major claim in RRG is that no syntactic subcategorization information of any kind is required in the lexical entries for verbs. For regular verbs, all that is required is the LS and nothing more, as in all except (11d). For most irregular verbs, only the macrorole number needs to be specified.[2] The prepositions that mark oblique arguments with verbs like *show* are predictable from general principles and need not be listed in the lexical entry (see Jolly, 1993; Van Valin and LaPolla (1997), §7.3.2). All of the major morphosyntactic properties of verbs and other predicates follow from their LS together with the linking system.

1.3 Syntactic functions, case and agreement

The linking between semantics and syntax depicted in (1) has two phases: first, the determination of semantic macroroles based on the LS of verb or other predicate in the clause, and second, the mapping of the macroroles and other arguments into syntactic functions. The traditional grammatical relations have no theoretical status in RRG; rather, RRG posits a single, construction-specific grammatical relation, which is termed the PRIVILEGED SYNTACTIC ARGUMENT [PSA] of the construction.[3] It subsumes the notion of 'syntactic pivot' used in Dixon (1972, 1979, 1994) and in earlier work in RRG. The non-PSA syntactic arguments in the clause are referred to as DIRECT or OBLIQUE CORE ARGUMENTS. The PSA for most (but not all) English constructions is the

traditional subject. Languages have selection hierarchies to determine the PSA; the ones for syntactically accusative and ergative languages are given in (12).

(12) Privileged Syntactic Argument [PSA] Selection Hierarchies
    a. Syntactically accusative languages: Highest ranking direct core argument in terms of the AUH (taking 'ARG of DO' as highest) is default.
    b. Syntactically ergative languages: Lowest ranking direct core argument in terms of the AUH is default.

For a language like English, (12a) captures the fact that in an active voice clause with a transitive verb, the actor is the PSA, whereas for a language like Dyirbal (Dixon 1972), in an active voice clause with a transitive verb the undergoer is the PSA, following (12b). Languages also differ with respect to whether the PSA must be a macrorole: German, Italian, Dyirbal, Jakaltak (Mayan) and Sama (Philippine) restrict PSA selection to actors and undergoers only, while Icelandic, Georgian, Japanese, Korean and Kinyarwanda allow non-macrorole direct core arguments to function as PSA (see Van Valin, 1991; Van Valin (2005), §4.2). The linking system is summarized in (13).

(13) System linking semantic and syntactic representations in RRG

SYNTATIC FUNCTIONS: PSA Direct Core Arguments Oblique Core Arguments

Privileged Syntactic Argument [PSA] Selection:
Highest ranking MR = default (e.g. English)
Lowest ranking MR = default (e.g. Dyirbal)

SEMANTIC MACROROLES:

        ACTOR   UNDERGOER

    Arg of 1st arg of 1st arg of 2nd arg of Arg of state
    DO do′ (x,... pred′ (x,y) pred′ (x,y) pred′ (x)

Transitivity = No. of Macroroles [MRα]
Transitive = 2
Intransitive = 1
Atransitive = 0

        Argument Positions in **LOGICAL STRUCTURE**

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

(Right margin: Language-specific / Universal)

The technical details of the linking algorithm are developed in Van Valin (2005), chapter 5 presents the linking algorithm for simple sentences, while chapter 9 presents the one for complex sentences. Both (1) and (13) contain double-headed arrows; this means that the linking system not only maps semantic representations into syntactic representations, but it also maps syntactic representations into semantic representations. This is, after all, part of what language users must do when they are producing and comprehending speech.[4] The emphasis in this paper is solely on the semantics to syntax mapping.

Case and agreement rules are formulated with reference to the linking system. Examples of case and agreement rules are given in (14) and (15).

(14) Case marking rules
    a. General Case Marking Rules for Accusative Languages:
        1. Highest ranking core macrorole takes nominative case.
        2. Other core macrorole takes accusative case.
        3. Non-macrorole direct core arguments take dative as their default case.
    b. Case Marking Rules for Dyirbal:
        1. Third-person NPs
            a. Lowest ranking core macrorole takes absolutive case.
            b. Other core macrorole takes ergative case.
            c. Same as (14a3)
        2. First- and second-person NPs: same as (14a)

(15) Agreement rules
    a. Finite Verb Agreement in English, Icelandic and German
       Finite verb agrees with highest ranking core macrorole in person and number.
    b. Predicate Adjective and Passive Participle Agreement in Icelandic
       Predicate adjectives and passive participles agree with the undergoer of the predicate of which they are a part in number, case and gender.

The operation of these rules can be illustrated using the following examples from German in (16a) and Dyirbal in (16b); the LS for verbs meaning 'give' is presented in (16c).[5]

(16) a. Die   Frauen hab-en   dem   Mann einen Hut gegeben.
      the.NOM women have-3pl the.DAT man   a.ACC hat   given
      'The women gave a hat to the man.'

b. Balam   mirañ   bagul ya≠a-gu   wuga-nbaÑgun dyugumbi-≠u.
   NM beans.ABS   NM   man-DAT give-TNS   NM   woman-ERG
   'The woman gave some beans to the man.'

c. [**do´** (x, Ø)] CAUSE [BECOME **have´** (y, z)]

In the German example, x = *Frauen* 'women', y = *Mann* 'man', and z = *Hut* 'hat'; in the Dyirbal sentence, x = *dyugumbil* 'woman', y = *ya≠a* 'man', and z = *mirañ* 'beans'. The first step in the linking is to assign macroroles. The x argument is the leftmost in terms of the AUH, and therefore the nouns meaning 'woman/women' will be the actor in each clause. The z argument is the rightmost, and therefore *Hut* 'hat' and *mirañ* 'beans' will be the undergoer. The third argument, the y argument meaning 'man', will be a non-macrorole core argument. Following the PSA selection hierarchy in (12), the actor *Frauen* 'women' will be the PSA in German, since it is a syntactically accusative language, and the undergoer *mirañ* 'beans' will be the PSA in Dyirbal, since it is a syntactically ergative language. The case rules in (14) are relevant now. For German, the rules in (14a) require that the actor, the highest ranking core macrorole, appear in the nominative case, that the undergoer, the other macrorole, appear in the accusative case, and that the non-macrorole core argument, *Mann* 'man', appear in the dative case. For Dyirbal, the rules in (14b1) are relevant, because all of the arguments are third person. The lowest ranking macrorole, the undergoer, must appear in the absolutive case, while the other macrorole, the actor, appears in the ergative case; finally, the third argument, *ya≠a* 'man', which is not a macrorole, appears in the dative case. The single macrorole argument of an intransitive verb is both the highest ranking and the lowest ranking, and therefore it will appear in the nominative case in German and in the absolutive case in Dyirbal. German, but not Dyirbal, has finite verb agreement, and the rule in (15a) states that the finite verb must agree with the highest ranking macrorole, which is the actor in this sentence. Thus, the basic morphosyntactic properties of the German and Dyirbal

sentences in (16) are readily captured by the lexical representation of the verbs in (16c) and the principles in (12)-(15) from the RRG linking system.
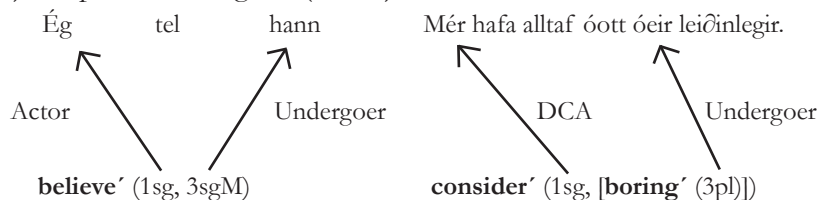
An important problem is posed by verbs which are very similar in meaning or fall into the same semantic class but which exhibit very different morphosyntactic properties, e.g. English *own* (e.g. *I own the house*) vs. *belong to* (e.g. *The house belongs to me*), or German *mögen* 'like' (e.g. *Ich mag das Haus* 'I [NOM] like the house [ACC]') vs. *gefallen* 'please' (e.g. *Das Haus gefällt mir* 'The house [NOM] pleases me [DAT]'). The RRG analysis of *belong (to)* and *gefallen* 'please' is that they are M-intransitive; they are exceptional in that despite having two arguments in their LS, they take only one macrorole. Hence this would have to be marked in their lexical entry, as in (11c) for *belong (to)*. This is all that needs to be stated regarding these verbs, as all of their other properties fall out from the linking system. In order to illustrate this, comparable examples from will be used; they involve the verbs *telja* 'believe' and *óykja* 'think, consider', as in (17).

(17) a. Ég   tel   hann
    1sgNOM believe.1sgPRES 3sgM.ACC
    'I$_{ACT}$ believe him$_{UND}$.'
  b. Mér   hafa   alltaf óótt   óeir   lei∂inlegir.
    1sgDAT have.3pl always thought 3plNOM boring.NOMpl
    'I$_{DCA}$ have always considered them$_{UND}$ boring.'

Telja 'believe' is a regular transitive verb with the LS **believe´** (x, y); the *x* argument is the actor, and the *y* argument is the undergoer. The actor is the PSA, following (12a), and it appears in the nominative case, following (14a); the finite verb also agrees with it, following (15a). The undergoer, the other macrorole, appears in the accusative case. In contrast, *óykja* 'consider' is M-intransitive and has the LS **consider´** (x, y) [MR 1]. This specification says only that there is one macrorole but not which one it is; that follows from (10b), which assigns undergoer, due to the lack

of an activity predicate in the LS. The y argument is the undergoer, following the AUH, leaving the x argument as a non-macrorole core argument. Icelandic is one of the languages which does permits non-macrorole core arguments to function as PSA, and following (12a), the x argument is the PSA. This is the correct assignment, because the dative NP in (17b) is the argument that can be 'raised', can control a reflexive, etc.; the nominative NP cannot. With respect to case and agreement, the undergoer is the only (and therefore the highest ranking) macrorole, and consequently it appears in the nominative case; the non-macrorole core argument appears in the dative case, following (14a3). The finite verb agrees with the highest ranking macrorole, hence with *óeir* 'they', and the predicate adjective agrees with the undergoer, which the nominative NP, according to (15b). The linking in these two sentences is presented in a simplified form in (18).

(18) Simplified linking for (17a, b)

Ég   tel   hann   Mér hafa alltaf óott óeir lei∂inlegir.

Actor   Undergoer   DCA   Undergoer

believe´ (1sg, 3sgM)   **consider´** (1sg, [**boring´** (3pl)])

Hence all of the morphosyntactic properties of (17b) follow from the LS plus the linking system; verbs like *óykja* 'consider' in Icelandic, *gefallen* 'please' in German and *belong (to)* in English are completely regular except for their transitivity.[6]

2. Voice

All of the discussion in the previous section concerned active voice clauses with transitive verbs, and the PSA selection hierarchies in (12) specify only the default selections. In many languages (probably most),

the default linkings are the only linkings possible, but in some languages the default can be overridden and a different argument can be selected to function as PSA. In an accusative language like German, it is possible for the other macrorole argument of a transitive verb, the undergoer, to function as PSA in a passive construction; similarly, in a syntactically ergative language like Dyirbal, it is possible for the actor of a transitive verb to function as PSA in an antipassive construction. This is illustrated in (19).

(19) a. Der   Hut   wurde dem   Mann von der   Frauen gegeben.
        the.NOM hat was the.DAT man   by   the.DAT women given
        'The hat was given to the man by the women.'
     b. Balan dyugumbil baÑgum mirañ-dyu wugal-Ña-ñu
        bagul ya≠a-gu.
        NM woman.ABS NM beans-INST give-ANTI-TNS
        NM man-DAT
        'The woman gave some beans to the man.'

In (19) the default PSA selections have been overridden, yielding passive and antipassive clauses, respectively. The case and agreement rules apply as before. In these constructions there is only one core macrorole, and therefore it is the highest ranking (German) or the lowest ranking (Dyirbal). Hence the undergoer *Hut* 'hat' appears in the nominative in German, and the actor *dyugumbil* 'woman' appears in the absolutive in Dyirbal. The case marking of 'man' is the same as in (16), because it is still a non-macrorole core argument. The default choices for PSA are treated as oblique adjuncts in these constructions, appearing in a PP headed by *von* in German or in the instrumental case in Dyirbal. The finite verb agrees with *Hut* 'hat' in German, as it is the highest ranking core macrorole, following the rule in (15a). The situation with respect to PSA selection and voice in these two systems is summarized in (20).[7]

(20) Default and marked PSA selections in accusative and ergative systems

Syntactic system	Default choice for PSA	Choice for PSA requiring special construction
Accusative	Actor	Undergoer [Passive]
Ergative	Undergoer	Actor [Antipassive]

3. Universal vs. Language-specific Aspects of Linking

The summary of the RRG linking system in (13) is divided into two parts, one of which is labeled 'universal' and the other 'language-specific'. A major claim of RRG is that what may be called the 'lexical' phase of the linking, the determination of macrorole assignments based on the LS of the verb, is virtually universal and subject to only very limited cross-linguistic variation, while the second phase, which may be termed the 'syntactic' phase, the mapping of the macroroles and other arguments into the syntax, is subject to considerable cross-linguistic variation, e.g. the contrast between accusative and ergative syntactic systems. The variation in the lexical phase concerns (i) whether languages allow variable undergoer assignment of the type discussed in §2.2, and (ii) the role of animacy in macrorole assignment (see Van Valin and LaPolla (1997), §7.4.1).

Most linking theories go directly from thematic relations to grammatical relations and posit universal constraints on the linking like 'highest ranking thematic relation links to subject'. Approaches such as this ignore ergative systems and split-intransitive systems like Acehnese.[8] The RRG system, in contrast, has two steps, (1) LS to macroroles and (2) macroroles to syntactic functions.[9] By factoring the linking in this way, it is possible to capture significant generalizations that the first approach cannot. In order to illustrate the kind of generalization that the RRG approach can capture, applicative constructions in syntactically

accusative and syntactically ergative languages will be examined. Applicative constructions have been traditionally characterized as involving the appearance as direct object of an argument that would normally appear as an indirect or oblique object. Applicatives in two accusative languages, Indonesian (Dardjowidjojo, 1971) and Swahili (Driever, 1976; Vitale 1981) are presented in (21) and (22); analogous constructions from two syntactically ergative languages, Sama (Walton, 1986) and Dyirbal, are given in (23) and (24).

(21) a. Ali meng-(k)irim   surat itu   kepada Hasan.   Indonesian
        ACV-send   letter DEF   to
     'Ali sent the letter to Hasan.' Direct object[10] = theme (UND)
   a´. Surat itu   di-kirim   kepada Hasan   oleh Ali.
      letter DEF   PASS-send   to   by
      'The letter was sent to Hasan by Ali.'
   a´´. *Hasan di-kirim surat itu   (kepada)   oleh Ali.
              PASS-send   letter DEF (to)   by
       'Hasan was sent the letter (to) by Ali.'
   b. Ali meng-(k)irim-kan   Hasan   surat itu
         ACV-send-APL   letter DEF
      'Ali sent Hasan the letter.' Direct object = recipient (UND)
   b´. Hasan di-kirim-kan   surat   itu   oleh Ali.
            PASS-send-APL letter   DEF   by
      'Hasan was sent the letter by Ali.'
   b´´. *Surat   itu   di-kirim-kan   Hasan   oleh Ali.
        letter   DEF   PASS-send-APL   by
        'The letter was sent Hasan by Ali.'

(22) a. Badru   a-li-andik-a   barua   (%kwa Juma).[11]   Swahili
                 1-PAST-write-IND   letter to
      'Badru wrote a letter (to Juma).' Direct object = theme (UND)

a´. Barua i-li-andik-w-a   na Badru   (%kwa Juma).
   letter   9-PAST-write-PASS-IND   by   to
   'The letter was written (to Juma) by Badru.'
b. Badru   a-li-mw-andik-i-a   Juma   barua.
            1-PAST-1-write-APL-IND   letter
   'Badru wrote Juma a letter.'   Direct object = recipient (UND)
b´. Juma   a-li-andik-i-w-a   barua   na Badru.
            1-PAST-write-APL-PASS-IND   letter   by
   'Juma was written a letter by Badru.'
b´´. *Barua   i-li-mw-andik-i-w-a   Juma   na Badru.
       letter   9-PAST-1-write-APL-PASS-IND   by
   'The letter was written Juma by Badru.'

Standard (i.e. thematic relations map directly to grammatical relations) analyses of applicative constructions postulate that they involve an alternative choice for direct object. The theme is the default choice in (21a) and (22a), and only the direct object of an active verb can function as the subject in a passive construction in these languages, as shown in (21a´, a´´) and (22a´). The recipient is the direct object in the applicative constructions in (21b) and (22b), and it, not the theme, can serve as the subject of the passive versions of these sentences, as shown in (21b´, b´´) and (22b´, b´´). The RRG analysis in terms of macroroles would claim that what is involved here is not an alternative choice for 'direct object' but rather an alternative choice for undergoer. In these languages, undergoer is the default choice for 'direct object' and is the only argument that can serve as the subject of a passive construction.

With respect to syntactically accusative languages like Indonesian and Swahili, these two analyses seem to be equivalent, but they make very different claims with regard to syntactically ergative languages like Sama and Dyirbal. (In Sama only pronouns are case-marked; full NPs are marked as oblique by the preposition *ma*.)

(23) a. B'lli   ku   taumpa' ma   si   Andi. Sama
       buy   1sgERG shoes   OBL PM Andy
       'I bought the shoes for Andy.' Subject/PSA = theme (UND)
    b. B'lli-an   ku   si   Andi   taumpa'.
       buy-APL   1sgERG PM Andy   shoe
       'I bought Andy some shoes.' Subject/PSA = benefic. (UND)

(24) a. Bayi ya≠a   yugu-Ñga   dyana-ñu.   Dyirbal
       NM   man.ABS wood-LOC   stand-TNS
       'The man is standing in/at/on some wood.' Subject/PSA = theme (ACT)
    b. Bala yugu   baÑgul ya≠a-Ñgu   dyanay-ma-n.
       NM wood.ABS NM   man-ERG   stand-APL-TNS
       'The man is standing on/with some wood.' Subj/PSA = location (UND)

The absolutive NP is the syntactic subject for many constructions in these languages,[12] and the addition of the applicative morpheme *(-an* in Sama[13] and *-mal* in Dyirbal) results in an alternative choice for the absolutive NP, hence an alternative choice for *subject*, not direct object as in (21) and (22). For the thematic-relations-to-grammatical-relations theories, it appears that two accounts are required for the languages in (21)-(24): in syntactically accusative languages, applicative constructions yield an alternative choice for direct object, while in syntactically ergative languages, they result in an alternative choice for subject. The RRG analysis, on the other hand, applies equally to all four languages. Applicative constructions involve an alternative undergoer choice (or, in the case of Dyirbal, adding an undergoer argument to the clause), and this is true of each of the languages. The differences among them arise from the linking between macroroles and syntactic functions summarized in (20): in Indonesian and Swahili the undergoer is the 'direct object' in an active-voice clause and PSA in a passive clause, whereas in

Sama and Dyirbal it is the PSA in an active-voice (i.e. non-antipassive) clause. Thus, the RRG analysis captures an important generalization about applicative constructions that analyses positing only thematic relations and grammatical relations miss.

It should be noted the Dyirbal also has an alternation like that in (8); in addition to the pattern in (16b) with *wugal* 'give', which is analogous to (8a) in that the theme is the undergoer, it is also possible to have the recipient as the undergoer, as in (25), which is analogous to (8b).

(25) Bayi ya≠a   baÑgum mirañ-dyu wuga-nbaÑgun dyugumbi-≠u.
      NM man.ABS NM beans-INST   give-TNS   NM   woman-ERG
      'The woman gave the man some beans.'

In terms of the LS in (16c), *ya≠a* 'man' is the *y* argument and is the undergoer; because (25) is active voice, the undergoer is the PSA. In RRG terms, (25) is like (8b) and English dative shift in that there is a marked linking to undergoer in terms of the AUH. Where English and Dyirbal differ, of course, is in the mapping of macroroles to syntactic functions. The identity of the alternation in the two languages would be hidden in the standard analysis, since the grammatical relation of the recipient in (25) would not be the same as the grammatical relation as in (8b) or as in English dative shift. Here again a significant generalization is captured in the RRG analysis but not in the standard account.

There is a principled reason for why the lexical phase of the linking is universal. The *Aktionsart* distinctions underlying the lexical decomposition are universal; all languages exhibit them. The notions of actor and undergoer are likewise valid for all languages, and the relationship between macroroles and LSs is governed by the principles in (10) and the AUH, which are likewise universally valid. Thus, all of the components of this phase of the linking are universal. It is in the syntactic phase that considerable cross-linguistic variation is found, and it is only because the linking has been factored into two phases, one lexical and the other syntactic, that the truly universal aspects can be identified.

4. Conclusion

In this paper it has been shown that universal patterns of linking can be identified using Role and Reference Grammar, a framework in which the linking between semantics and syntax is factored into two phases, a lexical phase and a syntactic phase. The universal patterns are found in the lexical phase of the linking, while the major cross-linguistic variation is located in the syntactic phase.[14] These generalizations could not be readily captured in frameworks which posit a direct mapping from thematic relations to grammatical functions.

The RRG linking system also demonstrates how it is possible to dispense completely with syntactic subcategorization information in the lexical entries for verbs. Given the LS of the verb and the linking principles, it is possible to derive the morphosyntactic properties of most verbs without any further information. In the case of irregular verbs, the M-transitivity must be specified in the lexical entry, but this is semantic, not syntactic information in RRG. As the discussion of pairs of verbs like *own* and *belong (to)* in English and *mögen* 'like' and *gefallen* 'please' in German showed, quite striking differences in morphosyntactic behavior can follow from a simple difference in M-transitivity. Thus, RRG provides a theoretical and descriptive framework in which linguistically significant generalizations about the relationship between the semantic argument structure of the verb and the morphosyntactic realization of the arguments can be readily captured and explained.

References

BITTNER, Maria; HALE, Ken. Ergativity: toward a theory of a heterogeneous class. *Linguistic Inquiry,* 27, p. 531-604, 1996.

DARDJOWIDJOJO, S. The *men-, men-kan* and *men-i* verbs in Indonesian. *Philippine Journal of Linguistics,* 2, p. 71-84, 1971.

DIXON, R. M. W. *The Dyirbal language of north Queensland.* Cambridge: Cambridge University Press, 1972.

DIXON, R. M. W. Ergativity. *Language,* 55, p. 59-138, 1979.

DIXON, R. M. W. *Ergativity.* Cambridge: Cambridge University Press, 1994.

DOWTY, David. *Word meaning and Montague Grammar.* Dordrecht: Reidel, 1979.

DRIEVER, Dorothea. *Aspects of a case grammar of Mombasa Swahili.* Hamburg: Helmut Buske Verlag, 1976.

DURIE, Mark. Grammatical relations in Acehnese. *Studies in Language,* 11, p. 365-399, 1987.

GUERRERO, Lilián; VAN VALIN, Robert D., Jr. Yaqui and the analysis of primary object languages. *International Journal of American Linguistics,* 70, p. 290-319, 2004.

HOLISKY, Dee A. The case of the intransitive subject in Tsova-Tush (Batsbi). *Lengua,* 71, p. 103-132, 1987.

JOLLY, Julia. Preposition assignment in English. In: VAN VALIN, R. (Ed.). *Advances in role and reference grammar.* Amsterdam/ Philadelphia: John Benjamins, 1993. p.275-310.

MANNING, Christopher. *Ergativity*: argument structure and grammatical relations. Stanford: CSLI, 1996.

NARASIMHAN, Bhuvana. A lexical semantic explanation for 'quirky' case marking in Hindi. *Studia Lingüística,* 52, p. 48-76, 1998.

ROBERTS, Linda. Pivots, voice and macroroles: from Germanic to universal grammar. *Australian Journal of Linguistics,* 15, p. 157-214, 1995.

VAN VALIN, Robert D., Jr. Semantic parameters of split intransitivity. *Language,* 66, p. 221-260, 1990.

VAN VALIN, Robert D., Jr. Another look at Icelandic case marking and grammatical relations. *Natural Language and Linguistic Theory,* 9, p. 145-194, 1991.

VAN VALIN, Robert D., Jr. A synopsis of role and reference grammar. In: _____. (Ed.). *Advances in role and reference grammar.* Amsterdam/ Philadelphia: John Benjamins, 1993. p. 1-164.

VAN VALIN, Robert D., Jr. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press, 2005.

VAN VALIN, Robert D., Jr. Lexical representation, co-composition, and linking syntax and semantics. PUSTEJOVSKY, James, *et al.* (Ed.). *New developments in the generative lexicon*. Dordrecht: Kluwer. (in press a).

VAN VALIN, Robert D., Jr. Semantic macroroles and language processing. In: BORNKESSEL; SCHLESEWSKY, Matthias (Ed.). *Semantic role universals and argument linking*: Theoretical, typological and psycho/neurolinguistic perspectives. Berlin: Mouton de Gruyter. (in press b).

VAN VALIN, Robert D., Jr.; LAPOLLA, Randy J. *Syntax*: structure, meaning & function. Cambridge: Cambridge University Press, 1997.

VAN VALIN, Robert D., Jr.; WILKINS, David P. The case for 'effector': Case roles, agents and agency revisited. In: SHIBATANI, M.; THOMPSON, S. (Ed.) *Grammatical constructions*. Oxford: Oxford University press, 1996. p. 289-322.

VENDLER, Zeno. *Linguistics in philosophy*. Ithaca, NY: Cornell University Press, 1967.

VITALE, Anthony. *Swahili syntax*. Dordrecht: Foris, 1981.

WALTON, Charles. *Sama verbal semantics*: classification, derivation and inflection. Manila: Linguistic Society of the Philippines, 1986.

Notes

1 'DO, represents the lexicalization of agentivity with certain verbs. Hence its first argument is the traditional agent. With many verbs, agentivity is an implicature rather than a lexical property; see Holisky (1987) , Van Valin and Wilkins (1996) and Van Valin and LaPolla (1997) , §3.2.3.2 for detailed discussion.

2 When systematic exceptions to the principles in (10) exist, e.g. activity verbs are universally M-intransitive as a default (Van Valin, 1990; Van Valin and LaPolla (1997), §3.2.3.3), then this can be stated as a general principle, and individual lexical entries would not be marked with this feature. See Van Valin and LaPolla (1997), §4.2 for detailed discussion. It should also be noted that syntactic

subcategorization information is a stipulation of transitivity, and most theories include this information for *all* verbs. No such information is included in the lexical entries in (11), and the macrorole number is specified *only* for irregular verbs which are not subsumable under the kind of generalization mentioned above.

3 See Van Valin (2005), Chapter 4, for a detailed presentation of the RRG theory of grammatical relations

4 See Van Valin (in press b) for discussion of how RRG's linking system functions in neurocognitive models of language processing..

5 Abbreviations used: ACT 'actor', ABS 'absolutive', ACC 'accusative', ACV 'active voice', ANTI 'antipassive', APL 'applicative', DAT 'dative', DCA '(non-macrorole) direct core argument', DEF 'definite', ERG 'ergative', IND 'indicative', INST 'instrumental', LOC 'locative', LS 'logical structure', NM 'noun marker', OBL 'oblique', PASS 'passive', PAST 'past tense', pl 'plural', PM 'proper noun marker', PRES 'present tense', PSA 'privileged syntactic argument', sg 'singular', TNS 'tense', UND 'undergoer', 1, 9 'Swahili noun classes'.

6 For a detailed comparison of this analysis with other accounts, see Van Valin (1991), Van Valin and LaPolla (1997), §7.3.1.1.

7 For analyses of other systems as well as a typology of voice functions, see Van Valin and LaPolla (1997), §§6.3-4, 7.3.1. See also Roberts (1995).

8 Acehnese lacks syntactic grammatical relations and operates with a semantically-based system in which syntactic rules and restrictions are stated in terms of actor and undergoer rather than subject and direct object; see Durie (1987), Van Valin (1993) and Van Valin (2005), §4.1 for detailed arguments.

9 Manning (1996) follows an RRG-type approach but operates with a different factorization: (1) thematic relations to argument structure [ARG-S] , and (2) ARG-S to syntactic relations. The highest ranking thematic relation is the ARG-S subject, the other argument the ARG-S object. Languages differ in terms of how ARG-S relations map into syntactic relations: in accusative systems, the ARG-S subject = syntactic subject, while in ergative systems, the ARG-S object = syntactic subject. While this appears to be similar to the RRG system, ARG-S is explicitly claimed to be syntactic, not semantic; hence ARG-S subject is not equivalent to actor and ARG-S object is not equivalent to undergoer. This difference manifests itself clearly with respect to languages like Acehnese, which Manning does not discuss; it is not at all obvious that his approach can deal with a language of this type, while it presents no problems for RRG (see references in footnote 8)

10 Strictly speaking, RRG does not employ the notion of direct object, but it will be used here to facilitate the comparison between the two analyses. See Van Valin (1993), Van Valin (2005), §4.5 for discussion of how 'direct object' phenomena are handled in RRG

11 '%' means that the PP *kwa Juma* is not acceptable to all Swahili speakers

12 Bittner and Hale (1996) argue that the actor is the subject in Dyirbal based on a single construction. This conclusion is dubious, for two reasons. First, no language is completely ergative syntactically, and consequently the existence of non-ergative constructions is to be expected. The majority of constructions in Dyirbal do operate ergatively, and this cannot be overlooked. Second, the construction in question, manner adverbials which are verbal in nature and require the same actor as the verb they modify, can be given a semantic analysis, unlike conjunction reduction (topic chaining) and relativization, the major ergative constructions, and hence it is irrelevant to the determination of grammatical relations in Dyirbal. See Dixon (1979, 1994) for arguments as to why such semantically determined constructions are irrelevant to the syntactic type of a language.

13 The applicative marker *-an* in Sama is cognate with the applicative marker in Indonesian, *-kan*.

14 For an account of this variation, and in particular why certain linking patterns are more frequent than others, see Van Valin and LaPolla (1997), §6.5