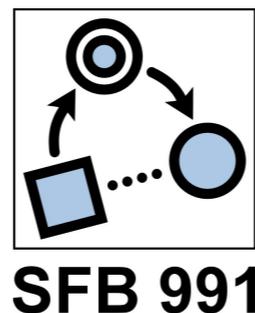


# On the status of lexical categories in RRG

Robert D. Van Valin, Jr.

University at Buffalo  
Heinrich Heine University Düsseldorf



This research was supported in part by the German Research Foundation [DFG] through SFB/CRC 991.

## Lexical and syntactic categories revisited

- Most syntactic theories assume some variety of endocentric syntax, where phrases are projections of lexical or functional/grammatical categories.
- Moreover, the projecting categories are assumed to be universally valid, i.e. N, V, and A for lexical categories (Baker 2003), and aspect, negation, tense, force, etc. for functional categories.
- RRG, on the other hand, rejects endocentrism as a fundamental feature of constituent structure.
- Rather, the two key notions are **NUCLEUS**, which is a category-neutral unit housing the predicating element in the clause (Foley & Van Valin 1984), and which, viewed from a cross-linguistic perspective, is not restricted to any particular lexical category or even to being a head (it can be phrasal), and **REFERENCE PHRASE** [RP], which is a potentially referring expression, which may be headed in principle by a range of lexical categories (Van Valin 2008).

## Lexical and syntactic categories revisited

—The appropriateness of NUCLEUS and RP as opposed to VP and NP are illustrated in the following examples from Tagalog (Schachter 1985), Nootka (Swadesh 1939) and Lakhota.

- (1) a. [CORE [NUC Nagtrabaho] [RP ang babae]]. Tagalog  
worked NOM woman  
‘The woman worked.’
- b. [CORE [NUC Babae] [RP ang nagtrabaho]].  
woman NOM worked  
‘The one who worked is/was a woman.’
- (2) a. [CORE [NUC Wała:k-ma] [RP qo:ʔas-ʔi]]. Nootka  
go-3sgPRES man-DEF  
‘The man is going.’
- b. [CORE [NUC Qo:ʔas-ma] [RP wała:k-ʔi]].  
man-3sgPRES go-DEF  
‘The one who is going is a man.’

## Lexical and syntactic categories revisited

(3) a. [CORE [RP Wíŋyaŋ kiŋ hená] [NUC wačhí-pi-kte-šni]]. Lakhota  
          woman the those           dance-PL-FUT-NEG

‘Those women will not dance.’

b. [CORE [RP Wačhí-pi kiŋ hená] [NUC wíŋyaŋ-pi-kte-šni]].  
          dance-PL the those           woman-PL-FUT-NEG

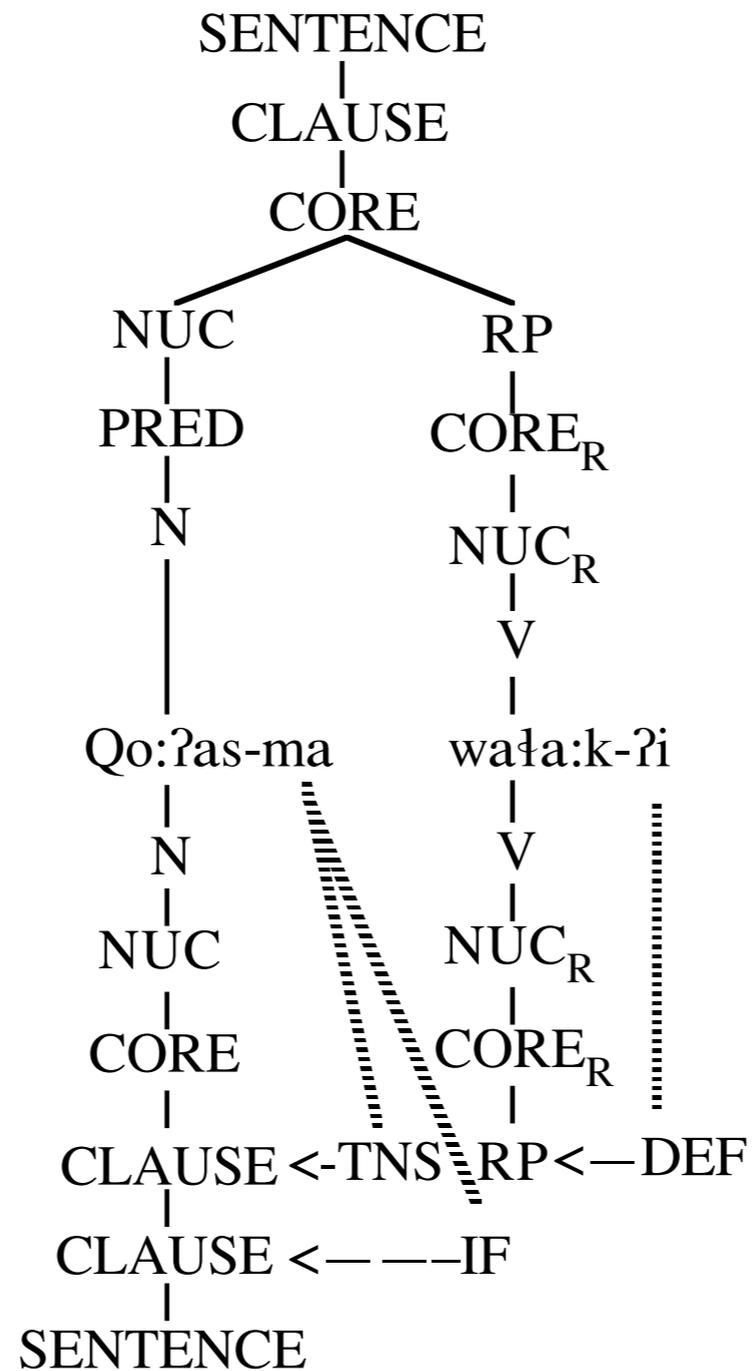
‘Those dancers will not be women.’

—In the (a) examples the predicate in the nucleus is a verb, and the head of the RP is a noun.

—In the (b) examples, on the other hand, the predicate in the nucleus is a noun, and the head of the RP is a verb, without derivational morphology or any special morphosyntactic treatment, unlike the corresponding elements in the English translation.

—Thus, RP and NUCLEUS are not projections of a particular lexical category but rather are category-neutral.

# Lexical and syntactic categories revisited



Structure of (2b) from Nootka

## Lexical and syntactic categories revisited

—The other crucial property of the NUCLEUS, that its ‘head’ can be phrasal, is essential for the analysis of Wari’ ‘intentional state constructions’ [ISC](Everett 2008), exemplified in (4b).

(4) a. *Mi'*      *non<sub>i</sub> -on<sub>j</sub>*                      *con*      *hwam*      *hwijima'<sub>i</sub>*      *mon tarama'<sub>i</sub>*.  
give      3plS.RP/P-3plO.M      PREP.3sgM      fish      children      COLL man  
‘The men gave the children fish.’

b. [*Ma'<sub>i</sub>*                      *co*                      *mao*      *na<sub>i</sub> -in<sub>j</sub>*                      *Guajará<sub>j</sub>*]  
that.PROX.hearer      M/F.RP/P      go(sg)      3sg.RP/P-3N      Guajará (Brazilian city)

*na<sub>k</sub> -nam<sub>l</sub>*      *'oro*      *narima'<sub>l</sub>*      *taramaxicon<sub>k</sub>*.  
3sg.RP/P-3plF      COLL      woman      chief

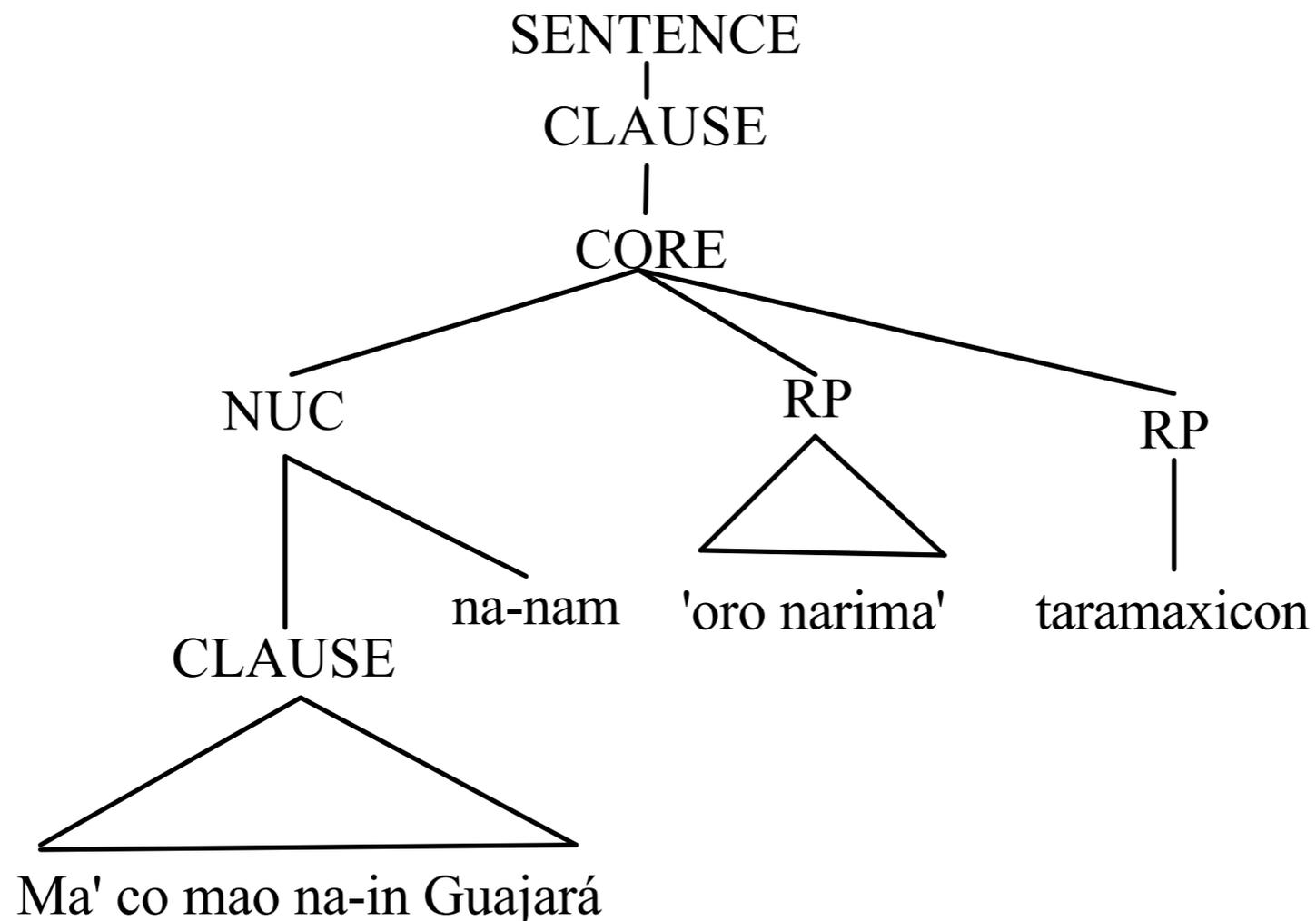
“‘Who went to Guajará?’” (said) the chief to the women.’

—Wari’ word order: (WH+Clitics) NUCLEUS+Clitics Core arguments

—What is in the NUCLEUS in the main clause in (4b)?

## Lexical and syntactic categories revisited

—The only possibility is the embedded clause ‘Who went to Guajar a?’, as Everett argues convincingly. There is no lexical verb in the main clause. The structure of (4b) is given below.



## Lexical and syntactic categories revisited

—ISCs are exocentric constructions, i.e. headless in phrase-structure terms, and as such they are a problem for approaches that require all phrases to be endocentric, i.e. headed.

—They are no problem for the notion of NUCLEUS, which is not restricted to lexical heads of certain categories. Embedding a clause in a NUCLEUS is highly marked in terms of the strong tendency towards symmetrical embedding, but it is not excluded by the theory.

—Everett (2008) shows that reformulating the description of Wari' clause structure in terms of NUCLEUS in place of V or VP leads to a comprehensive account of all of the ISC's properties.

—Thus the notions of RP and NUCLEUS are lexical-category-neutral syntactic categories, which make the analysis of (1b)-(3b) straightforward, and the fact that they are not projected from a lexical head but can accommodate phrases leads to an explanatory account of Wari' ISCs.

## Do all languages distinguish nouns from verbs?

- One of the ways the data in (1)-(3) have been interpreted is as evidence that these languages do not distinguish nouns from verbs.
- Facts like these have led some linguists to argue that Tagalog roots have no inherent lexical category and get assigned one on the basis of how they are used in an utterance (Foley 1998).
- This is questionable, however, because it confuses lexical category with grammatical function: *nagtrabaho* ‘worked’ is a predicate in (1a) and an argument in (1b), while *babae* ‘woman’ is an argument in (1a) and a predicate in (1b), but the category of each word does not change.
- The point is not that Tagalog lacks lexical categories but rather that there are few restrictions on what can function as the predicate in the NUCLEUS or the head of an RP.
- The same would hold true for Nootka and Lakhota.

## **Do all languages distinguish nouns from verbs?**

—In fact, detailed investigation of these languages revealed that there are morphophonological and/or morphosyntactic asymmetries which justify positing (at least) two distinct lexical classes.

—Himmelman (2008) distinguishes nouns from verbs in Tagalog based on a morphological criterion: verbs can take voice morphology, while nouns cannot.

—Jacobsen (1979) shows that in Nootkan languages there are subtle asymmetries between two classes of words in terms of ease of serving as the predicate in the nucleus: the favored class contains words denoting actions, events and states of affairs, while the other class contains words referring to objects, people, animals, etc.

—Boas & Deloria (1942:23) note there are morphophonological differences between nouns and verbs in Lakota, e.g. vowel contraction (nouns tend to contract adjacent vowels while verbs do not), reduplication (common with verbs, very rare with nouns).

## Do all languages distinguish nouns from verbs?

—A further difference can be observed in (3), repeated below.

(3) a. [CORE [RP Wíŋyaŋ kiŋ hená] [NUC wačhí-pi-kte-šni]]. Lakhota  
          woman the those           dance-PL-FUT-NEG

‘Those women will not dance.’

b. [CORE [RP Wačhí-pi kiŋ hená] [NUC wíŋyaŋ-pi-kte-šni]].  
          dance-PL the those           woman-PL-FUT-NEG

‘Those dancers will not be women.’

—It is not possible to mark number on a noun functioning in an RP as an argument or adjunct; plurality can be marked in the RP by a quantifier or plural determiner, or it can be signaled by plural cross-reference on the verb. Both of these are illustrated in (3a).

—When the head of the RP is a verb and is plural, however, it must take plural marking in addition to any quantifiers or plural determiners and plural cross-reference on the predicate, as in (3b).

—This is a morphosyntactic difference between nouns and verbs.

## **Are nouns and verbs universal?**

—Given the discussion of the previous question, the answer would seem to be ‘yes’.

—This conclusion is supported by the nature of language as a communication system. One of the major functions of language is communication with other humans, and what is communicated in many instances is information, which can be represented as propositions. Propositions involve reference and predication, and therefore languages have elements that are specialized for reference and other elements specialized for predication. The former are nouns and the latter are verbs.

—Hence the universality of nouns and verbs follows from the very nature of language as a system of communication, just like the universal parts of the layered structure of the clause.

—However, languages differ as to the semantic domains encompassed by the nouns and verbs in them.

## Are nouns and verbs universal?

—While words referring to objects, entities, people, animals, etc., will typically be realized as nouns, and those referring to actions, events, states of affairs, etc., will typically be instantiated as verbs, others may be realized in different ways in different languages.

—Consider property expressions, e.g. *tall, little, red, black, rich, happy*, etc. In some languages, e.g. English, they constitute a word class all their own, namely, adjectives.

—In other languages, however, they do not constitute an independent word class but rather are treated as a subtype of noun or verb. In Dyirbal (Dixon 1972) property words are assimilated to the category of noun and are treated morphosyntactically like nouns. In Lakhota, on the other hand, they are construed as a subclass of verbs and take the same inflectional morphology as verbs.

(5) Wínyañ kiñ hená **háŋska**-pi-kte-šni.  
woman the those tall-PL-FUT-NEG  
'Those women will not be tall.'

Lakhota

## Are nouns and verbs universal?

—Thus, even though all languages have nouns and verbs in their lexical inventory, the substantive content of these categories may vary from language to language.

—Another kind of cross-linguistic variation is how ‘deep’, so to speak, the distinction between the two lexical classes is. In the languages we have looked at so far, Tagalog, Nootka and Lakhota, the contrast is not very deep at all, with nouns and verbs sharing a lot of grammatical properties.

—This contrasts sharply with languages like Latin, Russian and Dyirbal, which make strict divisions between the two main categories, with English somewhere in the middle.

—All of this variation suggests that a firmer foundation is needed for a theory of lexical categories in RRG.

# Laying the foundation for an RRG approach to lexical categories

- At the most basic level, lexical items fall into one of two classes: they are either REFERRING EXPRESSIONS [REs] or PREDICATES.
- This is related to the fundamental opposition motivating the layered structure of the clause, namely, the opposition between predicating and non-predicating elements, which, as mentioned earlier, follows from the nature of language as a system of human communication.
- REs and predicates can have one of three grammatical functions: ARGUMENT, PREDICATOR, or MODIFIER.
- In (1)-(3), a predicate is functioning as the predicator in the nucleus and an RE as the core argument in the (a) examples, whereas in the (b) examples an RE is the predicator in the nucleus and a predicate serves as the core argument.

# Laying the foundation for an RRG approach to lexical categories

- The traditional categories of verb, adjective, adverb and adposition are semantically predicates and grammatically predicators or modifiers, as defaults.
- In RRG semantic representations it has always been the case that lexical modifiers and meaning-bearing adpositions are represented as predicates.
- REs would form the core of the language-specific category of noun, and predicates expressing actions and events would form the core for the category of verb in a language.
- Languages vary as to how they treat predicates expressing states, predicates expressing properties/attributes, predicates expressing manner, predicates expressing location, etc. In Dyirbal, predicates expressing states and manner are assimilated to the verb category, while those expressing properties/attributes are part of the noun category. In Lakota, on the other hand, all of those types of predicates fall into the verb category.

# Laying the foundation for an RRG approach to lexical categories

- Defining lexical categories in terms of their morphophonological and morphosyntactic properties leads to the observation that Tagalog, Nootka, Lakhota and other such languages show that languages need not make ‘deep’ lexical category distinctions, and this seems to follow from ability of lexical items to function as an argument, predicator or modifier rather freely.
- This ability is an important morphosyntactic property which many or all lexical items share, and it is precisely the morphosyntactic properties that different words do not share that is the basis for assigning them to different lexical categories.
- In languages with well-defined lexical categories there are strict constraints on the grammatical function that a given lexical category may have, e.g. nouns (REs) can be arguments, verbs (predicates of certain semantic types) can be predicators, and adjectives and adverbs (predicates of certain semantic types) can be modifiers.

# Laying the foundation for an RRG approach to lexical categories

—These restrictions are among the morphosyntactic properties which define the category. If a category is to have a different grammatical function, e.g. noun or adjective as predicator or verb as argument, special morphosyntactic treatment is necessary, e.g. the use of an auxiliary verb or verbalizing derivational morphology for non-verbal predicators, or nominalizing derivational morphology for non-nominal arguments. These special treatments are also part of the properties defining the different categories.

— This pattern is found in many European languages, and it has been taken as the norm for human language.

— This view has led some linguists, when confronted with Tagalog/Nootka/Lakhota-type languages with ‘shallow’ category distinctions, to posit derivational morphology for verbalization and nominalization, all of which is marked by zero-morphemes. Thus, on this type of analysis, in (1b) *babae* ‘woman’ has undergone zero-marked verbalization, yielding ‘to be a woman’, and *nagtrabaho* ‘worked’ has undergone zero-marked nominalization, yielding ‘the one who worked’.

## Laying the foundation for an RRG approach to lexical categories

- The RRG analysis of (1b)-(3b) given above does not involve any zero-marked derivational morphology. Rather, the category-neutral nature of RP and NUCLEUS make possible a straightforward description of these phenomena.
- For languages with ‘deep’ category distinctions, they are readily formulable in terms of RP and NUCLEUS, e.g., in English a non-verbal predicator requires an auxiliary verb, normally *be*, in order to have a well-formed nucleus.
- Note that in terms of this scheme, the ‘PRED’ node under the NUCLEUS node in the LSC should be interpreted as ‘PREDicator’ rather than as ‘PREDicate’.
- Adpositional predicates can be arguments, as with verbs like *put*, or adjunct modifiers. They may also function as predicators in locational predications like *John is in the house*.

# Laying the foundation for an RRG approach to lexical categories

— What, then, is the status of lexical categories in RRG?

—The universal semantic distinction is between REs and predicates, which underlies the noun-verb dichotomy. It is analogous to the actor vs. undergoer distinction, which is semantic and universal and which underlies the traditional subject-object dichotomy.

—The lexical categories beyond noun and verb in a language are differentiations of the functions of predicates as modifiers (i.e. as adjectives and adverbs), and all of the distinctions must be justified morphophonologically and/or morphosyntactically.

—Thus, from an RRG perspective, lexical categories are like grammatical relations: language-specific but with a universal semantic foundation.

## References

- Baker, M. 2003. *Lexical categories*. Cambridge: Cambridge University Press.
- Boas, F. & E. Deloria. 1942. *Dakota Grammar*. Washington, DC: US Govt. Printing Office.
- Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
- Everett, Dan. 2008. Wari' Intentional State Constructions. In Van Valin, ed., 381-410.
- Foley, W. & R. Van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: CUP.
- Foley, W. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane.
- Himmelman, N. 2008. Lexical categories and voice in Tagalog. In *Voice and Grammatical Relations in Austronesian Languages*, 247–293. CSLI Publications, Stanford, CA.
- Jacobsen, W. 1979. Noun and verb in Nootkan. *The Victoria Conference on Northwestern Languages*, 83-155. Victoria: British Columbia Provincial Museum.
- Schachter, Paul. 1985. Parts-of-speech systems. In T. Shopen, ed., *Language, typology and syntactic description*, vol. 1. Cambridge: Cambridge University Press.
- Swadesh, Morris. 1939. Nootka internal syntax. *IJAL* 9:77-102.
- Van Valin, Robert. 2008. RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In R. Van Valin, ed., 161-78.
- . 2013. Head-marking languages and linguistic theory. B. Bickel, L. A. Grenoble, D. A. Peterson, & A. Timberlake (Eds.), *Language typology and historical contingency. In honor of Johanna Nichols*, 91-124. Amsterdam: Benjamins.
- Van Valin, Robert, ed. 2008. *Investigations of the syntax-semantics-pragmatics interface*. Amsterdam: Benjamins.