

On the status of lexical categories in RRG

Robert D. Van Valin, Jr.

University at Buffalo & Heinrich Heine University Düsseldorf

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 predicate in the clause, 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. The appropriateness of these notions is illustrated in the following well-known example from Tagalog.

- (1) a. [CORE [NUC Nagtrabaho] [RP ang babae]].
 worked NOM woman
 ‘The woman worked.’
 b. [CORE [NUC Babae] [RP ang nagtrabaho]].
 woman NOM worked
 ‘The one who worked is/was a woman.’

In (1a) the predicate in the nucleus is a verb, and the head of the RP is a noun, following Himmelmann (2008); in (1b), 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. 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 possibly 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.

How can RRG approach the issue of lexical categories, given the range of cross-linguistic variation, from languages like Tagalog (Himmelmann 2008), Nootka (Swadesh 1939, Jacobsen 1979) and Lakhota (Boas & Deloria 1942), which seem to make only weak distinctions between nouns and verbs, to languages like Latin, Russian and Dyirbal, which make strict divisions between the two main categories, with English somewhere in the middle? One possibility is to claim that 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. This distinction derives from the nature of language as a system of communication: communication involves conveying information, which involves propositions, which involve reference and predication, hence REs and predicates. REs and predicates can have one of three grammatical functions: ARGUMENT, PREDICATOR, or MODIFIER. In (1a) a

predicate is functioning as the predicator in the nucleus and an RE as the core argument, whereas in (1b) an RE is the predicator in the nucleus and a predicate serves as the core argument. 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.

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. These restrictions are among the morphosyntactic properties which define each 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 linguists, when confronted with Tagalog-type languages, to posit derivational morphology for verbalization and nominalization which is all 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'. The RRG analysis of (1b) given above does not involve any zero-marked derivational morphology.

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. Adpositional predicates can be arguments, as with verbs like *put*, or adjunct modifiers. Thus, from an RRG perspective, lexical categories are like grammatical relations: language-specific but with a universal semantic foundation. Moreover, they play a rather different role in the non-endocentric syntax of RRG than they play in the endocentric syntax of other theories.

References

- Baker, M. 2003. *Lexical categories*. CUP.
Boas, F. & E. Deloria. 1942. *Dakota Grammar*. Washington, DC: US Govt. Printing Office.
Foley, W. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane.

- Himmelmann, 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. The noun and verb in Nootkan. *The Victoria Conference on Northwestern Languages*, 83-155. Victoria: British Columbia Provincial Museum.
- Swadesh, Morris. 1939. Nootka internal syntax. *IJAL* 9:77-102.