

Nouns, verbs and flexibles: implications for typologies of word classes

Erkki Luuk

Institute of Estonian and General Linguistics, University of Tartu, Estonia

Received 17 November 2008; received in revised form 18 January 2009; accepted 4 February 2009

Abstract

Following Farrell's analysis [Farrell, P., 2001. Functional shift as category underspecification. *English Language and Linguistics* 5 (1), 109–130], I submit that the most parsimonious hypothesis for stems that are ambiguous with respect to the noun/verb distinction (such as English *walk, love, kill*, etc.) is that they are neither nouns nor verbs but flexibles. It is generally agreed that the main functions of nouns and verbs are linguistic argument and linguistic predicate, respectively, and that the linguistic predicate/argument distinction is universal in the world's languages. I claim that if all languages have linguistic predicates and arguments, all languages must have at least one lexical class that maps to linguistic predicate and at least one lexical class that maps to linguistic argument. There are only three possibilities as to how a lexical class can map to linguistic predicate or argument: it can map to predicate, it can map to argument, or it can map to both. These three ways correspond to verbs, nouns and flexibles, respectively. Given this inventory of lexical classes, together with the premise that all languages have at least one lexical class that maps to argument and at least one that maps to predicate, the following five logically possible language types emerge: *noun/verb/flexible*, *noun/flexible*, *verb/flexible*, *noun/verb*, and *flexible*. After analyzing typological evidence for each of these types, I conclude that, if the criterion of pervasiveness of the typological trait is applied, type *noun/verb/flexible* is by far the most common, if not the only one present among the world's languages, with type *flexible* ranking next in probability. In addition, as 'word' has traditionally been found difficult to define [Broschart, J., 1997. Why Tongan does it differently: categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1 (2), 123–166; Di Sciullo, A.M., Williams, E., 1987. On the definition of word. *Linguistic Inquiry Monographs*, vol. 14. MIT Press, Cambridge, MA; Greenberg, J.H., 1963. *Essays in Linguistics*, Phoenix Books. Chicago, London], I propose the following definition for elementary word: a minimal unit of speech understood (though not necessarily used) outside context.

© 2009 Elsevier Ltd. All rights reserved.

Abbreviations: 1, 2, 3, first, second, third person; A, argument; ABS, absolutive case = allomorph of alienable genitive; Adj, adjective; Adv, adverb; AG, agent; ART, article; ARV, actor voice; AVV, active voice; COP, copula; DEF, definite; DET, determiner; DP, determiner phrase; F, flexible; FOC, focus; GEN, genitive; HUM, human; IMPF, imperfective; INDIC, indicative; LA, linguistic argument; LOC, locative; LP, linguistic predicate; LP/A, linguistic predicate/argument; N, noun; NL, natural language; NP, noun phrase; OBL, oblique; P, predicate; P/A, predicate/argument; PAST, past tense; PL, plural; POSS, possessive; REAL, realis mood; REFL, reflexive; S, singular; TAM, tense-aspect-mood; V, verb; VC, voice; XP, x phrase (see footnote 7).

E-mail address: erkkil@gmail.com

Keywords: Argument; Noun; Precategoriality; Predicate; Typology; Verb

1. Introduction

Nouns and verbs appear to be the most common lexical categories (Hockett, 1966; Sapir, 1978). Nevertheless, it is a matter of considerable controversy whether the noun/verb (N/V) distinction is in fact universal in the world's languages (Anderson, 2004; Bach, 2004; Laudanna and Voghera, 2002). On one hand, we have claims that a number of Malayo-Polynesian, Austro-Asiatic and native North American languages lack the N/V distinction (Broschart, 1997; Peterson, 2007; Whorf, 1945). On the other hand, we have strong evidence that at least some of these claims do not stand up to scrutiny (e.g., Baker, 2003; Hopper and Thompson, 1984; Jacobsen, 1979).

Obviously, before there is any hope of deciding on the universality of the N/V distinction, N and V must be defined (Rijkhoff, 2002). Although definitions of N and V are many, the majority (though not all – cf. Borer, 2003, 2005a,b; Marantz, 1984, 1997) of them agree on the following¹:

1. N and V are lexical categories.
2. Stems rather than their functional heads (determiner, aspect, etc.) bear lexical categories.
3. N and V are functionally motivated, i.e. have distinct (and possibly complementary) semantic and/or grammatical functions.

To sum up, N and V are held to be lexical categories that encode certain functions. The next question is, what are these functions? A number of different answers are proposed to this question. Below is a fairly representative, albeit not exhaustive list of functions that N/V is thought to encode:

- argument/predicate (Anward, 2001; Helmbrecht, 2001; Jacobsen, 1979).
- argument/predication (Broschart, 1997).
- argument/predicator (Anderson, 2004).
- nomination/predication (Ramat, 1999).
- referent/predication (Alfieri, 2007).
- reference/predication (Bhat, 2000; Croft, 2005; Peterson, 2007; Sasse, 1993a).
- discourse-manipulable participant/reported event (Hopper and Thompson, 1984).
- time-stable/non-time-stable concepts (Givón, 2001; Heine and Kuteva, 2002, 2007).
- designating a thing/designating a process (Langacker, 2004).

One can easily observe that all the above functions reduce to a basic predicate/argument (P/A) distinction. (With the latter three, this is not so straightforward but nevertheless a possibility.) The main difference is that Hopper and Thompson describe pragmatic functions, Langacker, Givón, Heine and Kuteva describe denotational functions, and the rest describe syntactic/propositional functions. Correspondingly, I take P and A to be the syntactic/propositional functions of V and N, respectively. It is generally agreed upon that P/A distinction is universal in the world's languages. See Hurford (2003a) and Hurford (2003b) for some claims that P/A distinction is even more broadly applicable. I have found only one source where the universality of P/A in natural language is contested. On Riau Indonesian, Gil (1994, p. 194) has written:

Moreover, there is no evidence for any kind of predicate–argument relationship: that is to say, no reason to characterize the meaning of *masok putih* [*enter white*, in playing billiards – L] as either **masok (putih)** “the white one is going in” or **putih (masok)** “the going is of the white one”.

¹ Technically, Marantz (1997) argues that lexical classes are not defined on the word level. The possibility that they are defined on the root (or even stem) level is not precluded by Marantz (1997).

Yet, on the same page, *masok putih* is analyzed as [EVENT[EVENTmasok] [THINGputih]] with respect to its “ontological category structure”. As P/A structure does not have to be grammatically marked – cf. **he run* and **run he* that are both ungrammatical but are nevertheless perceived as having P/A structure – this analysis is sufficient to establish *masok* as predicate and *putih* as argument. The fact that *masok putih* is grammatical without any overt P or A marking might seem puzzling. However, it is possible that P and A are signalled solely by word order in *masok putih*. According to Gil (2000), for example, sentence-initial position is characteristic of predicates in Tagalog, another Malayo-Polynesian language with an especially weak N/V distinction.

2. Nouns, verbs and flexibles

As pointed out in the previous section, the N/V distinction, let alone its (non-)universality, cannot be discussed before N and V are defined. The definitions of N and V that aim at both cross-linguistic universality and scientific rigour are scarce. Let us consider two well known but very different examples. Croft defines N and V as prototypical correlations of propositional act functions and semantic classes. For V, the respective values are ‘predication’ for function and ‘actions (relational, dynamic, transitory, nongradable)’ for semantic class; for N, the respective values are ‘reference’ and ‘objects (non-relational, static, permanent, nongradable)’ (Croft, 2001, pp. 87–88; Croft, 2000, pp. 88–89; Croft, 2005, p. 438). While cross-linguistically applicable, the definitions are vague as they rely on prototypes (e.g., *exist*, whilst being a verb, is not a prototypical action, as it is neither relational nor dynamic nor transitory; *embarrassment*, whilst being a noun, is not a prototypical object, as it is neither non-relational nor static nor permanent nor nongradable). The other problem is that the definitions have no obvious implications for syntax or morphology, i.e. they are detached from the level of description that linguists are most accustomed to. It is possible, of course, that cross-linguistically valid definitions of N and V with clear syntactic or morphological implications are simply unattainable (Croft, 2001).² Baker gives V the following definition: “X is a verb if and only if X is a lexical category and X has a specifier” (Baker, 2003, p. 23). However, NPs and adjectives seemingly also take specifiers in predicate constructions like *John is hungry* or *John is a skier*. In order to protect the definition, Baker introduces a hidden functional category he calls Pred. It is this category, he argues, not noun or adjective, that takes specifiers in constructions like the above-mentioned. There are three caveats to Baker’s definition of V. First, hidden structures should be avoided unless they explicitly simplify the explanatory framework, which does not seem to be the case here. Second, degree modifiers are sometimes regarded as specifiers of adjectives, while determiners and quantifiers are regarded as specifiers of nouns (Bennett, 1995; Putseys, 1989; Wehrli, 1988). Third, it is not clear what counts as a specifier. Consider, for example, the following Estonian sentence:

- (a) *Hämar-du-b.*
 dark-REFL-3S
 ‘It gets dark’

(a) is a full sentence but also a verb. The only way to preserve Baker’s definition with (a) is to assume that person-number markers can be specifiers for verbs. However, this assumption is unconventional, and might therefore be rejected.

Before we proceed with defining N and V, let us have a look at a language that putatively lacks these categories. The most striking example usually given is the one below, from the Wakashan language Nootka (Swadesh, 1939, pp. 78–79):

² Haspelmath (2007), for example, claims that universal pre-established (a priori, innate) formal categories do not exist. Observe that, even if he is correct, this does not preclude universal a posteriori formal categories, as (sub)optimal descriptions that balance exactness and parsimony. In addition, depending on their definitions, many categories (e.g. ‘word’) can be formal as well as based on semantic-pragmatic or phonetic substance (cf. below).

- (b)
- | | |
|----------------------------|------------------|
| <i>mamu:k-ma</i> | <i>qu:ʔas-ʔi</i> |
| work-INDIC | man-the |
| ‘The is man working’ | |
| <i>qu:ʔas-ma</i> | <i>mamu:k-ʔi</i> |
| man-INDIC | work-the |
| ‘The working one is a man’ | |

The functional heads in these two-word sentences, the mood marker and the article, are interchangeable, and both *mamu:k* and *qu:ʔ as* can function as argument and predicate. As the stems are symmetric with respect to P/A structure, they cannot be nouns or verbs, as nouns and verbs encode arguments and predicates, respectively.³ Arguments and predicates are not missing, though – they are encoded by the functional heads -*ma* and -*ʔ i*. Cf. the English *he worked* [work-PAST] and *he did the work* [DET work]. In English, like in Nootka or any other language (except a few select cases we will discuss below), tense-aspect-mood marking encodes predicate and determiners encode argument. The English *work* is no different from its Nootka counterpart in that it can be marked by both determiners and tense-aspect-mood (TAM) markers. The observation that English has a class of stems that are flexible with respect to N/V is not new (Jespersen, 1924). Farrell (2001) has argued for category underspecification and against zero derivation for such English stems. There are three possibilities with stems like *work*, *love*, *kill*, *walk*, etc. that can function both nominally and verbally:

- (1) Zero derivation: the noun is zero-derived from the verb and/or vice versa, or both are zero-derived from the uncategorized stem.
- (2) Homophony: there are two identical stems, the noun stem and the verb stem, no derivation.
- (3) Underspecification: the stem is underspecified (flexible, symmetric, universal) with respect to N/V, no derivation.

All these hypotheses are untestable but (3) is the most parsimonious.⁴ The second option under (1) differs from (3) in that in (3), there is no N/V derivation (only F), but by stipulating that “both are zero-derived from the uncategorized stem”, N/V is derived from F. Observe that the derivation is redundant, as the P/A functions can be assigned contextually by markers (cf. (b) and (4)–(8) below). By ‘untestable’ I mean that they are untestable by the present day methods. If lexical entries could be evidenced and identified in the brain, this issue could be resolved. In the meantime, I adopt hypothesis (3). A corollary of (3) is that there is a class of stems that are underspecified with respect to the N/V distinction. Borrowing a term from a related notion, ‘flexible parts-of-speech systems’ (e.g., Don and Van Lier, 2007; Hengeveld, 1992; Hengeveld and Rijkhoff, 2005; Rijkhoff, 2002), I call this class flexible (F).⁵ In encoding predicates and arguments, then, English has the 3-way distinction of N/V/F instead of just N/V (see Section 5.1). F is defined as necessarily flexible with respect to the N/V distinction and possibly flexible with respect to other parts-of-speech distinctions (e.g., V/

³ Observe that arguments are also encoded by NPs, and predicates are also encoded by COP + NP and COP + Adj constructions (see Section 3 for more details).

⁴ An anonymous reviewer suggests that semantics (e.g. polysemy) may provide crucial evidence in deciding between (1) and (3). Essentially, the choice between (1) and (3) boils down to the question of what is stored in the brain. Semantic criteria are not particularly ‘helpful in deciding on this. For example, the word *lock* can be argued to be polysemic, as it has related predicative and argumental meanings (plus an unrelated homonymous one). However, polysemy is consistent with all three hypotheses: e.g., the meaning of the lexical entry ‘lock’ can be the set-theoretic union of the predicative and argumental meanings (3), the predicative meaning can be synchronically derived from the argumental one (1), or there may be two lexical entries associated with the form *lock*, the predicative and the argumental (2) (all, of course, in addition to the homonymous entry). If the meaning of the lexical entry ‘lock’ is the set-theoretic union of the predicative and argumental meanings, one of the two can be switched off contextually by LP/A markers. The exact content and number of lexical entries cannot be decided by their use alone. Use, on the other hand, is our only clue to semantics. Accordingly, semantic predictions and tests are principally detached from exact descriptions of what is stored in the brain.

⁵ Flexible parts-of-speech systems contain at least one flexible lexeme class. Sometimes, the term ‘universal’ has been used instead of ‘flexible’ (Biggs, 1971; Pawley, 1966). The languages with flexible parts-of-speech systems are sometimes also labeled ‘pre-categorial’ (Evans and Osada, 2005). Apart from focusing on different aspects of the same phenomenon, all these differences are largely terminological.

Adj, N/Adj/Adv, etc.). For example, if a part of speech in a language conflates N/V/Adj or N/V/Adj/Adv, etc., it is F. On the other hand, if it conflates only V/Adj or N/Adj/Adv, it is not F.⁶

Observe that, in many cases, the noun use of the flexible is much more frequent than the verb use (or vice versa). Take, for example, the English word *soldier*. It is rather obvious that the noun use of *soldier* predates the verb use. According to *The Oxford English Dictionary* (1991, p. 956), the noun use of *soldier* is attested continuously from 1300 on, while the verb use is attested continuously from the 19th century (plus four citations from 1647–1800). Thus, one might be tempted to take it as an evidence for hypothesis (1). However, this reasoning is erroneous, as it confuses two entirely different things: diachronic and synchronic word derivation. Diachronically, the verb use of *soldier* is derived from the noun use. Hypothesis (1), however, is about synchronic word derivation. The fact that a word is diachronically derived from another does not entail a corresponding synchronic derivation. According to hypothesis (3), *soldier* is synchronically a flexible. Diachronically, it is a case of N becoming F. Similarly, all asymmetries between the noun and the verb use of a flexible can be accounted for by hypothesis (3).

It is now time to define nouns, verbs and flexibles. In the present paper, I adhere to the following definitions:

- (4) N = the propensity of stems to receive LA markers but not LP markers.
- (5) V = the propensity of stems to receive LP markers but not LA markers.
- (6) F = the propensity of stems to receive both LA and LP markers.

LA markers mark LAs, and LP markers mark LPs, by definition. Thus, N is aligned with LA, V is aligned with LP, and F is aligned with both LA and LP. LA and LP markers are defined in (7)–(8). Definitions (4)–(6) comply with the common assumption that stems rather than functional heads bear lexical categories and with several specific observations (e.g. that tense-aspect-mood markers attach to stems rather than to full words in most languages – see below). If we were to assume that functional heads rather than stems bear lexical categories (Borer, 2003, 2005a,b; Marantz, 1984), we could have definitions like “N = the propensity of LP but not LA markers to mark a particular stem” instead of (4)–(6). The definitions would work either way.

There are two reasons why I assume that stems rather than functional heads bear lexical categories. First, this assumption is intuitively more plausible and thus much more common (cf. Anderson, 2004; Anward, 2001; Croft, 2000; Crystal, 2004; Hopper and Thompson, 1984; Jespersen, 1924; Langacker, 2004; Lyons, 2004; Sasse, 1993a; Wierzbicka, 2000). Second, the hypothesis that functional heads rather than stems bear lexical categories cannot successfully account for the failure of functional heads to bear lexical categories, as in **the agree*, **potatoed*, **prairied*, etc., where the functional heads TAM and DET cannot bear lexical categories (presumably V and N, respectively). The hypothesis can account for this situation (e.g., “N = the propensity of LP but not LA markers to mark a particular stem”) only with self-subversive implications, by tacitly assuming that stems have a hidden property, a ‘functional value’ (Borer, 2005a, pp. 354–355), alluded to by “a particular stem...” in the definition above, which drives the marking. Thus, the hypothesis is not only counterintuitive but also non-parsimonious and self-subversive. In addition, Borer’s and Marantz’s hypothesis contradicts Heine and Kuteva (2002) and Heine and Kuteva (2007) theory of grammatical evolution, according to which nouns, verbs, adjectives and adverbs predate functional heads (demonstratives, case and tense markers, etc.). Definitions (4)–(6), on the other hand, provide a straightforward and uniform account of lexical categories.

The word has traditionally been a difficult concept to define (Broschart, 1997; Di Sciullo and Williams, 1987; Greenberg, 1963). Although Dixon and Aikhenvald (2002) propose separate definitions for ‘phonological word’ and ‘grammatical word’, I do not find the distinction particularly useful. First, their definition of ‘phonological word’ has no bearing on semantics or grammar. Clearly, this is not what is meant by ‘word’

⁶ Different classes of flexibles can be defined with respect to different parts-of-speech distinctions by using, for example, the following notation: $F_{N/V}$, $F_{V/Adj}$, etc. – flexible with respect to the N/V distinction, flexible with respect to the V/Adj distinction, etc. As the present paper is primarily concerned with only one class of flexibles (flexible with respect to the N/V distinction), I label this class F. Thus, henceforth, F or flexible refers strictly to ‘flexible with respect to the N/V distinction’.

in any other (i.e. non-phonological) sense. In order to have any non-phonological significance, ‘phonological word’ must refer to phonological criteria for ‘word’. As this would, obviously, require ‘word’ to be defined beforehand, ‘phonological word’ is either (a) of no consequence outside phonology, or (b) by definition secondary, and of secondary importance, as compared to a definition of ‘word’. In addition, ‘grammatical word’ is already used as a cover term for words that fill grammatical functions, e.g. determiners, conjunctions and adpositions.

I propose the following definition for ‘word’: a minimal unit of speech understood (though not necessarily used) outside context. This defines elementary word and does not preclude compound words (*doormat*) or word sentences like Yup’ik *kaipiallrulliniuk* (‘the two of them were apparently really hungry’ – Mithun, 1999, p. 38). Homonymy aside, the meaning of isolated articles and adpositions (*a, the, in, to* etc.) is transparent though they are never used outside context. The meaning of isolated suffixes (*-y* in *discovery, -ed* in *walked* etc.), however, is transparent only in the context of words they appear in.

Morphologically, ‘root’, ‘stem’ and ‘word’ are designated as follows: root = the base form; stem = root plus any number of derivational affixes (if applicable); word = stem plus any number of inflectional affixes (if applicable). If a language has N and/or V, it may have N and/or V derivation. When nouns, verbs or flexibles are nominalized or verbalized, the roots and lower order stems do not change their propensity. Instead, the derivational affix has changed the propensity of the higher order stem. Correspondingly, we have to assume that the higher order stem’s propensity overrides the propensities of lower order stems. In effect, we can have derivational chains like *agree* (V) → *agreement* (N) and *revolve* (V) → *revolution* (N) → *revolutionize* (V). Observe that, by the definition (4), pronouns and proper names constitute a subclass of N, as they, too, receive LA but not LP markers.

Thus, the universality of the LP/A distinction is independent of the universality of the N/V distinction, but not vice versa, if N and V are defined by their syntactic/propositional functions LA and LP, as they usually are (cf. (1)–(3) and Bhat, 2000; Croft, 2005; Peterson, 2007; Sasse, 1993a). I argue that, in addition to the universality of the LP/A distinction, linguistic arguments are universally marked by (a subset of) determiners, possessives and LP/A word order constraints, and linguistic predicates are universally marked by (a subset of) tense-aspect-mood, voice markers and LP/A word order constraints. Thus:

(7) LA markers = determiners, possessives and LP/A word order constraints.

(8) LP markers = tense-aspect-mood, voice markers and LP/A word order constraints.

As suggested with *masok putih* in Section 1, LA and LP can be marked by word order. Observe that many languages do not have some of these grammatical categories (e.g. determiners). However, I know of no language that fails to have at least one category of LA and one category of LP markers. According to Heine (1997), possessives are universal. Observe also that a single word order constraint could mark both LA and LP. Several approaches have proposed (a subset of) these categories, as well as gender, number, case and person, to be indicative of nouns and verbs or arguments and predicates (Broschart, 1997; Croft, 1990; Hopper and Thompson, 1984; Peterson, 2005, 2007; Sasse, 1993a). For particular languages, the selection of markers in (7)–(8) may seem too narrow or conservative but the aim is to establish sufficient and universal sets. The sets of markers in (7)–(8) appear to mark LAs and LPs universally, i.e. in all languages. The issue is not simple, however. For example, N and even DET accept TAM marking in some languages, e.g. Wakashan, Chamicuro and Lardil (Jacobsen, 1979; Nordlinger and Sadler, 2004). Importantly, however, TAM marking on N or DET appears to be restricted to a proper subset of the TAM paradigm in these languages. Thus, we should consider a distributional criterion which incorporates this proper subset condition in establishing LAs and LPs and, consequently, the word classes N, V and F in the world’s languages. Observe that this subset condition is covered by the formulation “...universally marked by (a subset of)...” above. In the next section, I analyze the problem of conflicting LA and LP markers in more detail.

The reason why number, gender and case are not included among the categories in (7)–(8) is that in many languages they mark other lexical classes besides nouns (Blake, 1994; Polinsky and Van Everbroeck, 2003). In most of these cases, the marking on non-noun appears to be an instance of agreement. In some languages it is rather obvious that the words that agree with nouns or flexibles in case and/or gender bear a variety of LA

marking that marks the whole XP instead of just N or F.⁷ This makes the particular LA marking, of course, useless for defining N and F. It might be argued that nouns are distinct from adjectives in that the former have fixed gender whereas for the latter gender can vary. This is certainly true but the question is the applicability of gender marking as such, not its particular qualities. In Russian, for example, verbs can be also marked for gender. Cross-linguistically, number or person-number marking on verbs is common (found in, e.g., English, Estonian and Russian). Moreover, nouns and verbs can take identical person-number affixes, e.g., in Estonian and Classical Nahuatl (Evans and Osada, 2005).

3. (Apparently) conflicting argument and predicate markers

Normally, F is marked by either LA or LP marking. One variant of LA and LP marker conflict refers to the situation when a stem is **simultaneously** marked by both LA and LP markers. Here are two examples from Tagalog and Tongan:

- (c) Tagalog
ang b-um-ib-ilí
 DET buy-ARV-IMPF.REAL-buy
 ‘the one buying’
- (d) Tongan (Broschart, 1997, p. 136)
na’e kau faiakó (‘a) e Siasí
 PAST PL.HUM teacher.DEF ABS ART Church
 ‘The Church provided the teachers’

I suggest that the apparent conflict can be resolved by the simple rule that the element’s identity is determined by its outmost, i.e. syntactically most distant marker. Thus we get [DET *ang* [VC/TAM *bumibilí*]] and [TAM *na’e*[*kau* [*faiakó* DET]]] (‘provided the teachers’). DET and VC/TAM mark elements as arguments and predicates, respectively (VC is voice). Cf. the following Tagalog example:

- (e) Tagalog (Himmelman, 2007)
ang langgam rin ang t-um-ulong sa mga bata
 DET ant also DET help-ARV-help LOC PL child
 Lit: the ones who helped the children were also the ants
 ‘The ants helped the children’

Here we have [DET *ang* [VC *t-um-ulong*]], ‘the ones who helped’. Again, DET and VC mark elements as arguments and predicates, respectively, with the outmost marker DET specifying *ang tumulong* as an argument. The rule that the outmost marker determines the element’s identity also accounts for LA and LP marker conflicts in the Munda language Kharia. According to Peterson (2007), what appear to be nouns marked for genitive case can also function as predicates:

- (f) Kharia (Peterson, 2007, p. 280)
iñ ho-kaR-te iñ-a?-y-o?j.
 1S 3-S.HUM-OBL 1S-GEN-y-AVV.PAST.1S
 I him/her mine made
 ‘I adopted her’

⁷ Hurford (2007) avoids the DP vs. NP debate (e.g., Bošković, 2007; Bruening, 2008; Lulu and Haitao, 2007; Progovac, 1998) by labeling the phrase XP. I suggest that, besides being a sensible precaution until the possible verdict is reached on the universality of either NP or DP, XP has the additional benefit of accommodating projections of proper names as well as argumental projections of F (e.g., *the handsome Mr. Smith, the run, a long walk*). Observe that this does not necessitate the NP-style analysis for XP but merely includes it as an option alongside with the DP analysis. XP is a theory-neutral term for all LA projections regardless of their licensing heads.

Here we have $[[i\ddot{n}-a\dot{?} \text{POSS}]-y-o\dot{?}j \text{VC/TAM}]$, 'mine made'. POSS marks argument; VC and TAM mark predicate. As the outmost markers are predicate markers, $i\ddot{n}-a\dot{?}-y-o\dot{?}j$ is a predicate.⁸ Observe also that the rule that the outmost marker determines element's identity is in concordance with the rule, specified in Section 2, that the higher order stem's propensity overrides the propensities of lower order stems.

Arguments can also be incorporated within (or converted to) predicates, as in the following example:

(g) *this is my book*

Here, the copula *is* specifies *my book*, which is an argument, as a part of the predicate *is my book*. If one analyzes predicates and arguments in this sentence, one gets $[A \text{ this}][P \text{ is } [A \text{ my book}]]$. There is no A and P marker conflict here, as the POSS *my* marks the argument *book*, which is specified as a part of the predicate by the copula *is*. Alternatively, we may say that the copula converts the argument that follows it into a predicate. In some other languages, e.g., Russian, Maltese and Moroccan Arabic, a predicate is specified by juxtaposing two arguments, as in

(h) Russian
eta moya kniga
 this my book

analyzable as $[A \text{ eta}][P [A \text{ moya kniga}]]$. In these languages, a juxtaposition of two arguments converts the second one to a predicate. In some languages at least, this holds for present tense indicative mood constructions only. As soon as, e.g., past or different mood is specified, the copula is required:

(i) Russian
eta byla moya kniga
 this was my book

(j) *eta dolzhen bytj moya kniga*
 this must be my book

My alignment of predicates and arguments is different from that of, e.g. Hengeveld (1992), according to which *my book* ~ *moya kniga* would simply be predicates. The present alignment follows from the definitions of LA and LP markers (7)–(8). As *book* ~ *kniga* are marked by LA markers (the possessives *my* ~ *moya*), they are LAs by definition. All stems, words and phrases that can be marked by LA or LP markers are LAs or LPs by definition.

Besides concurrent LA and LP marking, as in the examples (c)–(f), LA and LP markers may also conflict **distributionally**. If a lexeme accepts all (or at least a roughly equal proportion of) LA and LP markers in a language, that lexeme is F. Significant differences in the proportion of accepted LA and LP markers, however, pose a typological problem. As a solution, at least three different distributional criteria might be considered:

- (9) If a stem accepts at least one LA and at least one LP marker, it is F.
- (10) If a stem accepts all LA but only some LP markers in a given language, it is N (and if vice versa, it is V).
- (11) If a stem accepts a majority of LA and a minority of LP markers in a given language, it is N (and if vice versa, it is V).

⁸ An anonymous reviewer notes that the combination of case markers with predicative markers is possible only with genitive, not with oblique case. This makes perfect sense, as genitive is POSS, i.e. an argument marker that can (theoretically) be overridden by an outer predicate marker. This suggests that other case markers are not an integral part of the Kharia P/A marking paradigm – a result predicted by definitions (7)–(8).

For the sake of simplicity, let us name the criterion (9) exclusive and the criteria (10) and (11) inclusive. In the face of it, (10) and (11) may seem like hairsplitting. This intuition is supported by the fact that in many languages the conditions specified in (10) or (11) do not arise at all. In some other languages, however, the stems' propensity is sufficiently idiosyncratic for (10) and (11) to be applicable. According to (9), the lexemes that accept (a subset of) both LA and LP marking in a language are F. According to (10) and (11), however, (some of) these lexemes can be N and V instead. To give an example, Jacobsen (1979) analysis of Wakashan languages is based on (10) or (11). As a result, the lexemes that accept (a subset of) both LA and LP markers in Wakashan are classified not as F but N and V. His classification is based on the fact that – though both Jacobsen's N and V take TAM marking – TAM marking on N is limited to durative aspect and does not occur in future tense, whereas TAM marking on V covers all the TAM paradigm. Similarly, though both his N and V accept determiners, possessives (or at least some of them) attach to N only.

The fact that only one criterion at a time can be chosen for a (typo)logically coherent framework, poses a difficult trilemma.⁹ Needless to say, failure to distinguish these criteria from one another and to stick to only one of them at a time are major sources of confusion in N/V typology. It is a common practice (and a logical error) for scholars who have used the inclusive criteria to criticize the **results** of those who adhere to the exclusive criterion and vice versa – whereas in fact their critique could be relevant to the criterion only. Unfortunately, there is no a priori reason to prefer one criterion to the other.¹⁰ A necessary result of the situation is that the scholars applying inclusive criteria find N/V in all languages they investigate (Baker, 2003; Evans and Osada, 2005; Hopper and Thompson, 1984; Jacobsen, 1979), whereas those that apply the exclusive criterion find N/V absent in a number of languages (Gil, 1994; Itkonen, 2000). For a more general critique of the distributional method of establishing word classes, see Croft (2005).

4. The pervasiveness of a typological trait

A typological trait's pervasiveness in a particular language is another possible criterion that might or might not be considered. For instance, if a language has only one noun, does it have the category N or not? If the pervasiveness criterion is applied, it does; if it is not applied, it might not. The pervasiveness criterion is related to Evans and Osada (2005) 'exhaustiveness through the lexicon', which stipulates that a language must lack nouns and verbs altogether to be of type *F*. The difference is that Evans and Osada's principle is a criterion for languages of type *F*, whereas the pervasiveness criterion stipulates that a language has a particular category (e.g., the lexical category F) if it has at least one member of this category. Thus, Evans and Osada's exhaustiveness principle is about language type *F*, but the pervasiveness criterion is about linguistic categories (in general). Hengeveld (1992), for example, dismisses the pervasiveness criterion for his typology. As a result, the language (Tuscarora) that is claimed to have "a reduced number of true nouns" lacks the category noun on the next page (Hengeveld, 1992, pp. 67–68). Most researchers would reject this kind of reasoning. Indeed, it has been stressed that the pervasiveness criterion must be applied (Baker, 2003; Evans and Osada, 2005; Itkonen, 2000), and I am not aware that anyone would have insisted on the opposite, even in the cases where the criterion has been dismissed (Hengeveld, 1992; Rijkhoff, 2002). Moreover, in their recent papers, Hengeveld and Rijkhoff adhere to the pervasiveness criterion (Hengeveld et al., 2004; Hengeveld and Rijkhoff, 2005).

I agree that the idea that only one N, V or F can make a language fall into a different typological class may seem daunting. However, the absolute number of 1 is the simplest cut-off point to observe while being, logically, the most important one, as it indicates whether the language logically has a particular category or not. Part of the problem is that, typologically, any sensible cut-off point could be used as long as it is identified – and any absolute limit of, say, 1...40 and any percentage of 0.01...3 seems to be within the range of typologically sensible. It is very hard to motivate an exact limit in such conditions. In addition, a percentage limit would be impractical, as it would require analyzing representative random samples. However, if we could ascertain that, e.g., all languages are of type *N/V/F* with the absolute cut-off point of 1, it would be instructive

⁹ Assuming that (10) is a subcase of (11), the two are not necessarily mutually exclusive.

¹⁰ Typologically, it would be perhaps wiser to apply inclusive criteria (cf. the penultimate paragraph of Section 2).

to apply other limits instead. At this time, however, we lack certainty even with the simplest, the absolute limit of 1.

5. The five logically possible language types

Let us repeat the most important point made so far and consider some implications:

- (12) Every language has linguistic predicates (LP) and linguistic arguments (LA).
- (13) A corollary of (12): Every language has at least one lexical class that maps to LP and at least one lexical class that maps to LA.
- (14) There are only three possibilities how a lexical class can map to LP or LA: it can map to LP, it can map to LA, it can map to both. According to the definitions (4)–(6), these three ways correspond to V, N and F, respectively.
- (15) From (13) and (14) it follows that, with respect to lexical classes that map to LP and/or LA, there are five logically possible language types: $N/V/F$, N/F , V/F , N/V and F .

Observe that this is not a typology of real language types, but of logically possible ones that real types must map on to. I make no claim as to the reality of any of these types except $N/V/F$ (see Section 5.1). The present typology predicts, however, that types $N\cup$ and V (lacking both F and V or N, respectively) are not expected to be found among the languages of the world. The prediction is not trivial, as there are some claims of languages of type N (Itkonen, 2000; cf. Sasse, 1993) and type V (Hengeveld, 1992; Itkonen, 2000; Rijkhoff, 2003). In Sections 5.2 and 5.3, I show that these claims do not hold up to scrutiny. The typology is purely logical, i.e. it does not follow that any particular of the five proposed types exists. However, at least one of them must, by definition, be realized in the languages of the world. In what follows, I take a closer look at all these types individually. I hypothesize that, if the pervasiveness criterion is applied, the most plausible situation is either (16) or (17):

- (16) All the world's languages belong to type $N/V/F$.
- (17) All the world's languages belong to types $N/V/F$ and F .

Observe that both cases involve F as a separate, cross-linguistically universal part of speech. In addition, the possibility of a language of type $N/V/F$ has not received much attention in the literature so far. The question whether (16) or (17) is more plausible of the two lies outside the scope of the present study; in fact, it is possible that both are wrong, as the existence of any of the five types cannot be precluded at this point.

5.1. Type $N/V/F$

In all figures except 4, F is represented by the two ovals minus the circles of N and V (if any). Thus, in Fig. 1, $F = (LA - N) \cup (LP - V)$. N, V and F are lexical classes, LA and LP are their functions. N maps to LA, V maps to LP, and F maps to both LA and LP. Theoretically, N, V and F may also map to adjectival and other functions, not included in this schema. The only restrictions are that N cannot map to LP and V cannot map to LA. The vast majority (if indeed not all) the world's languages are of this type. English, for one, has a rich assortment of nouns (e.g. *society*, *life*, *prairie*, *child*, *lizard*), verbs (*agree*, *write*, *ask*, *comprehend*,

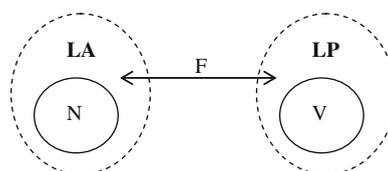
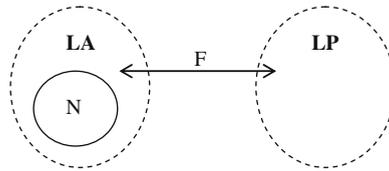


Fig. 1. Type $N/V/F$.

Fig. 2. Type *N/F*.

engage) and flexibles (*round, love, kill, walk, run*) (Bierwisch, 2001; Don and Van Lier, 2007; Farrell, 2001; Jespersen, 1924; Vogel, 2000).

5.2. Type *N/F*

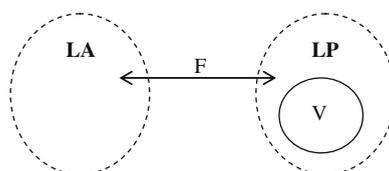
Type *N/F* (Fig. 2) has been proposed for Maori (Biggs, 1971) and Niuean (Massam, 2005). However, Bauer (1993, p. 259), while admitting that “there do not appear to be clear-cut structural properties which could help to provide an operational definition for the class verb”, suggests the capacity to take the nominalizing suffix -*Canga* and tense-aspect markers to be among the criteria for verbs in Maori. In any case, the evidence that Maori and Niuean are *N/F* languages is inconclusive, as the finding of only one *V* stem in them would be sufficient for these languages to fall into type *N/V/F* instead (if the pervasiveness criterion is applied, as it should be). It is not clear whether Biggs (1971) and Massam (2005) pay any attention to the pervasiveness criterion. It is also worth mentioning that, for Ancient Tamil, the logically impossible type *N*, lacking both *V* and *F*, has been tentatively proposed by Itkonen (2000). However, Itkonen admits that the (seemingly SOV) word order still maintains the *P/A* distinction in Ancient Tamil. According to (12)–(15), this is sufficient to rule out type *N*. Cf. Sasse (1993a) for a short overview on languages that have been claimed to belong to the logically impossible types *N* and *V*.

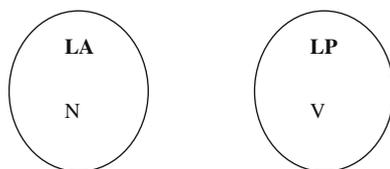
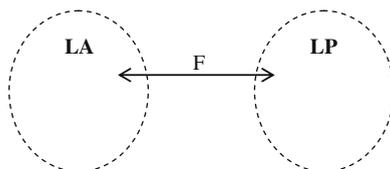
5.3. Type *V/F*

There is no conclusive evidence that a language of type *V/F* exists (Fig. 3). According to Hengeveld (1992), Quechua has two parts of speech, *V* and *N/Adj/Adv* (i.e. the language does not differentiate between nouns, adjectives and adverbs). *N/Adj/Adv*, however, is not flexible with respect to the *N/V* distinction – hence, it is not *F*. The same holds for Ngiyambaa, proposed to have the same parts of speech as Quechua (Rijkhoff, 2002).

Interestingly, there have also been a few alleged claims of languages of a logically impossible type *V*, lacking both *N* and *F* (Hengeveld, 1992; Itkonen, 2000; Rijkhoff, 2003). The claims are about two Iroquoian languages, Cayuga and Tuscarora. However, it is not the issue that Tuscarora has no nouns – there is just a reduced number of them (Hengeveld, 1992; Hengeveld et al., 2004). To repeat, even one noun would be sufficient for a language to logically have that category. Similarly, one noun and one verb would be sufficient for a language to have the *N/V* distinction. Cf. Baker (2003, p. 177): “An important typological difference exists only if categorial ambiguity extends to an entire open class of inflectionally similar words, thereby affecting the overall grammar of the language”. Tuscarora falls into type *N/V* or *N/V/F*, then.

For Cayuga, Sasse (1993a) and Sasse (2001) has posited a class of simplex forms such as *só:wa:s* ‘dog’ that reject verbal inflection. Mithun (2000) argues for a clear, even robust *N/V* distinction in all Iroquoian lan-

Fig. 3. Type *V/F*.

Fig. 4. Type *N/V*.Fig. 5. Type *F*.

languages already at the stem level. In Cayuga, only nouns take possessive prefixes and the noun suffix *-aʔ*, and only verbs are TAM-marked (Mithun, 2000). Presently, the dominant opinion is that both Cayuga and Tuscarora have nouns (Anward, 2001). Sasse (1993a) and Sasse (2001) suggest that Cayuga has nouns (that may, however, constitute a closed class), while giving a more hesitating estimate elsewhere (Sasse, 1993b). If it turned out that Cayuga has no nouns (which is unlikely), it would be a candidate for type *V/F* – assuming that it has at least one *F* stem. A more plausible candidate is Samoan, which has flexibles as well as (perhaps) the possibility to derive verbs (anonymous reviewer, p.c.). In addition, Neukom (2001) has hypothesized that a Munda language, Santali is of type *V/F*.

5.4. Type *N/V*

There is no conclusive evidence of a natural language of type *N/V* (Fig. 4). A language of this type would be perfectly conceivable, however, with the artificial languages like Ido or Esperanto being the closest examples (Jespersen, 1924). It is uncertain whether any natural language belongs to this type, as only one *F* stem would suffice for it to fall to type *N/V/F* instead, but the Iroquoian and Bantu languages which may lack *F* altogether are the prime possible candidates (Baker, 2003; Hengeveld, 1992; anonymous reviewer, p.c.).

5.5. Type *F*

A number of languages have been proposed to belong to this type.¹¹ The condition of type *F* (Fig. 5) is that the language has no stems that map to either *LA* but not *LP* or *LP* but not *LA* – in other words, the language has neither nouns nor verbs. As only one *N* or *V* would suffice for the language to belong to a different type, the condition should not be too hard to test. Nevertheless, *F* type has been posited many times and for a number of languages. Here is a typical testimony: “(1) all full words, including names, may serve as predicates and may be inflected using person markers/–/, and (2) any lexical item can become a referring expression by positioning a determiner in front of it” (Czaykowska-Higgins and Kinkade, 1998, p. 36, on Salishan). At least the following languages have, sometimes tentatively, been proposed to belong to type *F*:

1. Wakashan languages, once spoken in American Northwest, e.g. Nootka, Makah, Nitinat and Kwakiutl (Boas, 1947; Swadesh, 1939; Whorf, 1945; cf. Jacobsen, 1979).
2. Salishan languages, once spoken in American Northwest, esp. Straits Salish, Squamish and Upper Chehalis (Czaykowska-Higgins and Kinkade, 1998; Jelinek and Demers, 1994; Kinkade, 1963; cf. Sasse, 1993a).

¹¹ Languages of this type are alternatively labeled as ‘precategorial’ or having ‘flexible parts of speech’ (Don and Van Lier, 2007; Evans and Osada, 2005; Hengeveld, 1992; Hengeveld and Rijkhoff, 2005; Rijkhoff, 2002). The differences between these notions, if applicable, and type *F* are largely terminological.

3. Munda languages, spoken in India, esp. Mundari and Kharia (Bhat, 2000; Hengeveld and Rijkhoff, 2005; Peterson, 2005, 2007; Rijkhoff, 2003).
4. Several Malayo-Polynesian languages (Tongan, Samoan, Tagalog, Kambera,¹² *Tukang Besi*, Malay/Indonesian, Riau Indonesian, Ilokano – Bloomfield, 1942; Broschart, 1997; Gil, 1994, 2000, 2007; Hengeveld, 1992; Itkonen, 2000; Lambert, 1998; cf. Baker, 2003).
5. Vietnamese (Gil, 1994).

It must be pointed out that, for different languages and by different authors, *F* type has been posited at different levels of analysis (cf. a review in Jacobsen, 1979). Kwakiutl and Upper Chehalis, for example, have usually been claimed to be of type *F* at stem and root levels. For Nootka, the claim has been about stem or full word level, depending on the author. Finally, the Malayo-Polynesian languages have usually been claimed to be of type *F* at full word, stem and root levels. Another possible problem that a claim of an alleged *F* type language has to face is exemplified by the following case: It has been noted that in Samoan many roots can be found in the function of both verb phrase and noun phrase. However, the researchers (Mosel and Hovdhaugen, 1992, as reviewed in Baker, 2003, p. 177) have never observed *alu* ‘go’ as a noun or *mea* ‘thing’ as a verb. Does this mean that the lack of the N/V distinction in Samoan is not pervasive? It certainly might – but the absence of data is in itself inconclusive evidence. Without any positive evidence, only a native speaker’s competence could help to resolve the issue.

5.6. Summary

To sum up, the evidence for the existence of types *N/F*, *V/F*, *N/V* and *F* is inconclusive at best, as it depends, besides the data available to the researcher, on the following objective criteria:

- (18) The level of analysis (root, stem, full word or phrase).
- (19) The definitions of N, V and F.
- (20) Whether the same distributional criterion is used in interpreting the definitions of N, V and F (see (9)–(11)).
- (21) Whether the typological trait’s pervasiveness is taken to be a part of the criteria of its existence in the language (see Section 4).

Only when all the four criteria are matched is there hope for consensus on the actual existence of any one of the five logically possible language types and, by implication, on the universality of the N/V distinction. Definitions (4)–(6) fix the level of analysis to stem. The distributional criteria do not affect the logical impossibility of types *N* and *V* (they affect only whether a language is categorized as having the N/V distinction or being of type *F* instead). As suggested in footnote 10, the inclusive criteria are preferable. As argued in Section 4, the pervasiveness criterion should be applied. With these definitions and criteria, the most plausible conclusion is perhaps that all the world’s languages belong to either type *N/V/F* or types *N/V/F* and *F*. This is a preliminary estimate, and none of the other types is ruled out. However, I suggest the following frequency hierarchies for the five types (from most to least frequent in the world’s languages):

- (22) $N/V/F > F; N/V; V/F; N/F.$
- (23) $N/V/F > F > N/V; V/F; N/F.$
- (24) $N/V/F > F > N/V > V/F; N/F.$

Current data is insufficient to decide whether (22), (23), or (24) is the most plausible one.

6. Typology of the noun/verb distinction: questions for future research

Drawing from what has been said, it is possible to formulate several important questions for future research:

¹² However, Klamer (1998) and Klamer (2005) have offered convincing evidence that Kambera is of type *N/V/F*.

1. Is there a language that does not have at least two stems, one that maps to LA but not LP, and the other that maps to LP but not LA (type *F*)? (See Section 5.5 for the list of languages that have been suggested to belong to this type.)
2. Is there a language that does not have a stem that maps to both LA and LP (type *N/V*)? In particular, are any of the Iroquoian and Bantu languages of type *N/V*?
3. Is there a language that has at least two stems, one that maps to LP but not LA and the other that maps to both, but does not have a stem that maps to LA but not LP (type *V/F*)? In particular, is Santali, Samoan or Cayuga of type *V/F*?
4. Is there a language that has at least two stems, one that maps to LA but not LP and the other that maps to both, but does not have a stem that maps to LP but not LA (type *N/F*)? In particular, is Maori or Niuean of type *N/F*?

7. Conclusion

I submit that the most parsimonious hypothesis for the stems that are ambiguous with respect to the noun/verb distinction like the English *walk*, *love*, *run*, etc. is that they are neither nouns nor verbs but flexibles. The remaining two alternatives – zero derivation and homophony – are equally untestable but also posit unnecessary hidden structure. It is generally agreed upon that the main functions of nouns and verbs are linguistic argument (LA) and linguistic predicate (LP), respectively, and that the linguistic predicate/argument distinction is universal in the world's languages. I argue that determiners, possessives and LP/A word order constraints universally mark linguistic arguments, and that tense-aspect-mood, voice markers and LP/A word order constraints universally mark linguistic predicates. Importantly, a language needs only one LP marker and one LA marker (or just a single LP/A word order constraint) to make the LP/A distinction. Assuming that the main functions of N and V are LA and LP, respectively, and that LA and LP are universally marked, N, V and flexible (F) can be given the following definitions:

- (25) N = the propensity of stems to receive LA markers but not LP markers.
- (26) V = the propensity of stems to receive LP markers but not LA markers.
- (27) F = the propensity of stems to receive both LA and LP markers.

Premise (28) has certain typological implications:

- (28) Every language has linguistic predicates (LP) and linguistic arguments (LA).
- (29) A corollary of (28): every language has at least one lexical class that maps to LP and at least one lexical class that maps to LA.
- (30) There are only three possibilities how a lexical class can map to LP or LA: it can map to LP, it can map to LA, or it can map to both. According to the definitions (25)–(27), these three ways correspond to V, N and F, respectively.

From (29)–(30) it follows that, with respect to the lexical classes that map to LP and/or LA, there are exactly five logically possible language types: *N/V/F*, *N/F*, *V/F*, *N/V* and *F*. We took a closer look at each of these types, and found *N/V/F* to be by far the most common among the world's languages. The evidence for types *N/V*, *V/F* and *N/F* is modest. Indeed, it is possible that all the world's languages are of type *N/V/F*, if the criterion of pervasiveness of the typological trait is applied, specifying that a language has a category if it has at least one member of this category. In typology, the question has long been whether the N/V distinction is universal in the world's languages. At this time, the numerous claims of languages of type *F* remain the biggest challenge for the universality of the N/V distinction.

Acknowledgements

I am most grateful to John Peterson, Hendrik Luuk, Haldur Õim, Kees Hengeveld, Marian Klamer and the three anonymous reviewers for their comments, suggestions and criticism on this paper. All remaining mistakes are, of course, my own.

References

- Alfieri, L., 2007. The contribution of Proto-Indo-European to a part-of-speech typology. Workshop on Languages with Flexible Parts-of-Speech Systems. University of Amsterdam.
- Anderson, J.M., 2004. Syntactic categories and notional features. In: Aarts, B. (Ed.), *Fuzzy Grammar: A Reader*. Oxford University Press, Oxford; New York, pp. 225–237.
- Anward, J., 2001. Parts of speech. In: Haspelmath, M., König, E., Oesterreicher, W., Raible, W. (Eds.), *Language Typology and Language Universals. An International Handbook*. Walter De Gruyter, Berlin; New York, pp. 726–735.
- Bach, E., 2004. Linguistic universals and particulars. In: Sterkenburg, P.G.J.V. (Ed.), *Linguistics Today: Facing a Greater Challenge*. John Benjamins, Amsterdam, Philadelphia, p. vi, 47–60.
- Baker, M.C., 2003. In: *Lexical Categories: Verbs, Nouns, and Adjectives: Cambridge Studies in Linguistics*, vol. 102. Cambridge University Press, Cambridge, New York.
- Bauer, W., 1993. *Maori*. Routledge, London, New York.
- Bennett, P., 1995. *A Course in Generalized Phrase Structure Grammar*. UCL, London.
- Bhat, D.N.S., 2000. Word classes and sentential functions. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin, New York, pp. 47–63.
- Bierwisch, M., 2001. Repertoires of primitive elements: prerequisite or result of acquisition? In: Weissenborn, J., Höhle, B. (Eds.), *Approaches to Bootstrapping*. John Benjamins, Amsterdam, pp. 281–307.
- Biggs, B., 1971. Languages of Polynesia. *Current Trends in Linguistics* 8 (1), 466–505.
- Blake, B.J., 1994. *Case: Cambridge Textbooks in Linguistics*. Cambridge University Press, Cambridge [England], New York, NY, USA.
- Bloomfield, L., 1942. Outline of Ilocano syntax. *Language* 18, 193–200.
- Boas, F., 1947. Kwakiutl grammar with a glossary of the suffixes. *Transactions of the American Philosophical Society* 37 (3), 201–377.
- Borer, H., 2003. Exo-skeletal vs. endo-skeletal explanations. In: Moore, J., Polinsky, M. (Eds.), *The Nature of Explanation in Linguistic Theory*, CSLI, Chicago, pp. 31–67.
- Borer, H., 2005a. Structuring Sense. In: *The Normal Course of Events*, vol. II. Oxford University Press, Oxford.
- Borer, H., 2005b. Structuring Sense. In: *In Name Only*, vol. I. Oxford University Press, Oxford.
- Bošković, Ž., 2007. What will you have, DP or NP? In: Elfner, E., Walkow, M. (Eds.), *NELS 37*, University of Massachusetts Graduate Linguistic Students' Association, Urbana-Champaign.
- Broschart, J., 1997. Why Tongan does it differently: categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1 (2), 123–166.
- Bruening, B., 2008. Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. *The 32nd Penn Linguistics Colloquium*. University of Pennsylvania.
- Croft, W., 1990. *Typology and universals: Cambridge textbooks in linguistics*. Cambridge University Press, Cambridge [England], New York.
- Croft, W., 2000. Parts of speech as language universals and as language-particular categories. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin; New York, pp. 65–102.
- Croft, W., 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford University Press, Oxford.
- Croft, W., 2005. Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 40 (3), 431–441.
- Crystal, D., 2004. English word classes. In: Aarts, B. (Ed.), *Fuzzy Grammar: A Reader*. Oxford University Press, Oxford; New York, pp. 191–211.
- Czaykowska-Higgins, E., Kinkade, M.D., 1998. Salish languages and linguistics. In: Czaykowska-Higgins, E., Kinkade, M.D. (Eds.), *Salish Languages and Linguistics: Theoretical and Descriptive Perspectives*. Mouton, The Hague, pp. 1–68.
- Di Sciullo, A.M., Williams, E., 1987. On the definition of word. In: *Linguistic Inquiry Monographs*, vol. 14. MIT Press, Cambridge, MA.
- Dixon, R.M.W., Aikhenvald, A.Y., 2002. Word: a typological framework. In: Dixon, R.M.W., Aikhenvald, A.Y. (Eds.), *Word: A Cross-linguistic Typology*. Cambridge University Press, Cambridge, pp. 1–41.
- Don, J., Van Lier, E., 2007. On the lack of word-class distinction: flexibility and conversion. Workshop on Languages with Flexible Parts-of-Speech Systems. University of Amsterdam.
- Evans, N., Osada, T., 2005. Mundari: the myth of a language without word classes. *Linguistic Typology* 40 (3), 351–390.
- Farrell, P., 2001. Functional shift as category underspecification. *English Language and Linguistics* 5 (1), 109–130.
- Gil, D., 1994. The structure of Riau Indonesian. *Nordic Journal of Linguistics* 17, 179–200.
- Gil, D., 2000. Syntactic categories, cross-linguistic variation and universal grammar. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin, New York, pp. 173–216.
- Gil, D., 2007. On the position of Riau Indonesian in a typology of flexible-syntactic-category languages. Workshop on Languages with Flexible Parts-of-Speech systems. University of Amsterdam.
- Givón, T., 2001. *Syntax: An Introduction*. John Benjamins, Amsterdam, Philadelphia.
- Greenberg, J.H., 1963. *Essays in Linguistics*. Phoenix Books, Chicago, London.
- Haspelmath, M., 2007. Pre-established categories don't exist: consequences for language description and typology. *Linguistic Typology* 11, 119–132.
- Heine, B., 1997. *Cognitive Foundations of Grammar*. Oxford University Press, Oxford.
- Heine, B., Kuteva, T., 2002. On the evolution of grammatical forms. In: Wray, A. (Ed.), *The Transition to Language*. Oxford University Press, Oxford, pp. 376–397.
- Heine, B., Kuteva, T., 2007. *The Genesis of Grammar: A Reconstruction*. Oxford University Press, New York.

- Helmbrecht, J., 2001. Head-marking vs. dependent-marking languages. In: Haspelmath, M., König, E., Oesterreicher, W., Raible, W. (Eds.), *Language Typology and Language Universals. An International Handbook*. Walter De Gruyter, Berlin, New York, pp. 1424–1432.
- Hengeveld, K., 1992. In: *Non-verbal predication: theory, typology, diachrony: functional grammar series*, vol. 15. Mouton de Gruyter, Berlin, New York.
- Hengeveld, K., Rijkhoff, J., 2005. Mundari as a flexible language. *Linguistic Typology* 40 (3), 406–431.
- Hengeveld, K., Rijkhoff, J., Siewerska, A., 2004. Part-of-speech systems and word order. *Journal of Linguistics* 40 (3), 527–570.
- Himmelman, N.P., 2007. Lexical categories and voice in Tagalog. In: Austin, P., Musgrave, S. (Eds.), *Voice and Grammatical Functions in Austronesian Languages*, CSLI, Stanford.
- Hockett, C., 1966. The problem of universals in language. In: Greenberg, J.H. (Ed.), *Universals of Language*. MIT Press, Cambridge, Mass, p. xxvii, 1–12.
- Hopper, P.J., Thompson, S.A., 1984. The discourse basis for lexical categories in universal grammar. *Language* 60, 703–752.
- Hurford, J.R., 2003a. Ventral/dorsal, predicate/argument: the transformation from perception to meaning. *Behavioral and Brain Sciences* 26 (3), 301–311.
- Hurford, J.R., 2003b. The neural basis of predicate-argument structure. *Behavioral and Brain Sciences* 26 (3), 261–283, Discussion 283–316.
- Hurford, J.R., 2007. The origin of noun phrases: reference, truth, and communication. *Lingua* 117 (3), 527–542.
- Itkonen, E., 2000. Noun vs. Verb? Parts of Speech in and Across Languages, Symposium Presentation. University of Helsinki.
- Jacobsen, W.H., 1979. Noun and verb in Nootkan. In: Efrat, B. (Ed.), *The Victorian Conference on Northwestern Languages*, British Columbia Provincial Museum, Victoria, pp. 83–153.
- Jelinek, E., Demers, R.A., 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70, 697–736.
- Jespersen, O., 1924. *The Philosophy of Grammar*. H. Holt and Company, New York.
- Kinkade, M.D., 1963. Phonology and morphology of Upper Chehalis. *International Journal of American Linguistics* 29, 345–356.
- Klamer, M., 1998. *A Grammar of Kambera*: Mouton Grammar Library. Mouton de Gruyter, Berlin/New York.
- Klamer, M., 2005. Kambera. In: Adelaar, K.A., Himmelman, N.P. (Eds.), *The Austronesian Languages of Asia and Madagascar*. Routledge, London, pp. 709–734.
- Lambert, S., 1998. Partial overlap of nominal and verbal linkers: the case of Tongan. In: *International Conference of the SFB 'Theory of the Lexicon'*, Wuppertal.
- Langacker, R.W., 2004. Bounded regions. In: Aarts, B. (Ed.), *Fuzzy Grammar: A Reader*. Oxford University Press, Oxford; New York, pp. 239–245.
- Laudanna, A., Voghera, M., 2002. Nouns and verbs as grammatical classes in the lexicon. *Rivista di Linguistica* 14 (1), 9–26.
- Lulu, W., Haitao, L., 2007. A description of Chinese NPs using head-driven phrase structure grammar. In: *14th International Conference on Head-Driven Phrase Structure Grammar*, CSLI, Stanford, pp. 287–305.
- Lyons, J., 2004. A notional approach to the parts of speech. In: Aarts, B. (Ed.), *Fuzzy Grammar: A Reader*. Oxford University Press, Oxford, New York, pp. 213–223.
- Marantz, A., 1984. *On the Nature of Grammatical Relations*. The MIT Press, Cambridge, MA.
- Marantz, A., 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In: A. Dimitriadis, L. Siegel, C. Surek-Clark, A. Williams (Eds.), *21st Annual Penn Linguistics Colloquium*, University of Pennsylvania Department of Linguistics, Philadelphia.
- Massam, D., 2005. Lexical categories in Niuean. In: Carnie, A., Harley, H., Dooley, S.A. (Eds.), *Verb First: On the Syntax of Verb-initial Languages*, John Benjamins, Amsterdam/Philadelphia, pp. 227–241.
- Mithun, M., 1999. *The Languages of Native North America: Cambridge Language Surveys*. Cambridge University Press, Cambridge, UK; New York.
- Mithun, M., 2000. Noun and verb in Iroquoian languages: multicategorisation from multiple criteria. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin; New York, pp. 397–420.
- Mosel, U., Hovdhaugen, E., 1992. *Samoan Reference Grammar*. Scandinavian University Press, Oslo.
- Neukom, L., 2001. Santali. *Languages of the World/Materials*, vol. 323. Lincom Europa, München.
- Nordlinger, R., Sadler, L., 2004. Tense beyond the verb: encoding clausal tense/aspect/mood on nominal dependents. *Natural Language and Linguistic Theory* 22, 597–641.
- Pawley, A., 1966. Samoan phrase structure: morphology-syntax of a Western Polynesian language. *Anthropological Linguistics* 8, 1–63.
- Peterson, J., 2005. There's a grain of truth in every "myth", or, Why the discussion of lexical classes in Mundari isn't quite over yet. *Linguistic Typology* 9 (3), 391–405.
- Peterson, J., 2007. Languages without nouns and verbs? An alternative to lexical classes in Kharia. In: Masica, C.P. (Ed.), *Old and New Perspectives on South Asian Languages. Grammar and Semantics. Papers growing out of the Fifth International Conference on South Asian Linguistics (ICOSAL-5)*, Held at Moscow, Russia in July 2003, Motilal Banarsidass, Delhi, pp. 274–303.
- Polinsky, M., Van Everbroeck, E., 2003. Development of gender classifications: modeling the historical change from Latin to French. *Language* 79 (2), 356–390.
- Progovac, L., 1998. Determiner phrase in a language without determiners. *Journal of Linguistics* 34, 165–179.
- Putseys, Y., 1989. Sentential complementation and the meanings of such. In: Jaspers, D., Klooster, W.G., Putseys, Y., Seuren, P.A.M. (Eds.), *Sentential Complementation and the Lexicon*, Foris, Dordrecht.
- Ramat, P., 1999. Linguistic categories and linguists' categorizations. *Linguistics* 37 (1), 157–180.
- Rijkhoff, J., 2002. Verbs and nouns from a cross-linguistic perspective. *Rivista di Linguistica* 14 (1), 115–147.

- Rijkhoff, J., 2003. When can a language have nouns and verbs? *Acta Linguistica Hafniensa* 35, 7–38.
- Sapir, E., 1978. *Language: An Introduction to the Study of Speech*. Hart-Davis MacGibbon, London.
- Sasse, H.-J., 1993a. Syntactic categories and subcategories. In: Jacobs, J., Von Stechow, A., Sternefeld, W., Vennemann, T. (Eds.), *Syntax: An International Handbook*. Mouton de Gruyter, Berlin, pp. 646–686.
- Sasse, H.-J., 1993b. Das nomen – eine universale kategorie? *Sprachtypologie und Universalienforschung* 46, 187–221.
- Sasse, H.-J., 2001. Scales between nouniness and verbiness. In: Haspelmath, M., König, E., Oesterreicher, W., Raible, W. (Eds.), *Language Typology and Language Universals. An International Handbook*. Walter De Gruyter, Berlin; New York, pp. 495–509.
- Swadesh, M., 1939. Nootka internal syntax. *International Journal of American Linguistics* 9, 77–102.
- The Oxford English Dictionary, 1991. Ser-Soosy, vol. XV. Clarendon Press, Oxford.
- Vogel, P.M., 2000. Grammaticalisation and part-of-speech systems. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin; New York, pp. 259–284.
- Wehrli, E., 1988. Parsing with a GB-grammar. In: Reyle, U., Rohrer, C. (Eds.), *Natural Language Parsing and Linguistic Theories*. D. Reidel, Dordrecht, pp. 177–201.
- Whorf, B.L., 1945. Grammatical categories. *Language* 21 (1), 1–11.
- Wierzbicka, A., 2000. Lexical prototypes as a universal basis. In: Vogel, P.M., Comrie, B. (Eds.), *Approaches to the Typology of Word Classes*. Mouton de Gruyter, Berlin; New York, pp. 285–317.