

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS

12

ERKKI LUUK

The noun/verb and
predicate/argument structures



TARTU UNIVERSITY
PRESS

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

Dissertation is accepted for the commencement of the Degree of Doctor of Philosophy in General Linguistics on July 10, 2009 by the Council of the Institute of Estonian and General Linguistics, University of Tartu, Tartu, Estonia.

Supervisor: Professor emeritus Haldur Õim, PhD,
University of Tartu, Estonia

Opponent: Professor Frederick J. Newmeyer,
University of Washington, USA

Commencement: Room 139 in University main building, Ülikooli 18, Tartu, on September 22, 2009 at 14:15.

ISSN 1406–5657

ISBN 978–9949–19–187–1 (trükis)

ISBN 978–9949–19–188–8 (PDF)

Autoriõigus Erkki Luuk, 2009

Tartu Ülikooli Kirjastus

www.tyk.ee

Tellimus nr. 290

PREFACE

This thesis is an unexpected result of three years of hard work that took me from Estonian linguistics via evolutionary linguistics to some of the most obscure problems of general linguistics, viz., the predicate/argument structure of natural language and typology of major word classes. The result is unexpected in the sense that, half a year ago, I had no idea that it will become the subject of my PhD thesis. I feel very lucky that it did, and would like to thank a number of people without whose help all this would not have been possible – at least not in such a reassuring way. Thus, without further ado, I thank the following people:

- First and foremost, my supervisor Haldur Õim, one of the main motivators of my work, and the person whose role was decisive in me turning to linguistics. All this would not have been possible without his support and encouragement over the years that I have known him.
- Frederick J. Newmeyer, Renate Pajusalu, Urho Määttä and Ray Jackendoff for their many useful suggestions and comments on the preliminary version of this thesis.
- John Peterson for his endless positivity and help, on Kharia and far beyond.
- Marian Klamer and Lukas Neukom for their useful information on Kambera and Santali, respectively.
- Tania Kuteva for her kind help with a pre-copy of one of her books, co-authored with Bernd Heine.
- James Hurford and Kees Hengeveld, for their comments on earlier drafts of one of the papers this thesis is based on.
- Noam Chomsky for a lengthy and extremely useful email correspondence. Although our discussions hardly ever touched the topics in this thesis, it was educating as well as motivating to dispute over language with him, not to mention the privilege of witnessing the parsing power of one of the most brilliant minds.
- I also owe much to the inspiring events of the Graduate of Linguistics and Language Technology, and to the discussions with my fellow students and professors there.
- I am extremely grateful to all the anonymous reviewers who have helped me to elaborate and elucidate the two papers that the thesis is based on. Without your invaluable help my case would have been much weaker.
- Most importantly, I thank my parents for their understanding and support. I owe you more than I can ever express. And I thank my friends and brothers, especially Hendrik for dragging me into evolutionary linguistics, which led (besides, hopefully, some other things) to this very thesis.

CONTENTS

LIST OF PUBLICATIONS	9
ABBREVIATIONS	10
Conventions	11
INTRODUCTION	12
1. NOUNS, VERBS AND FLEXIBLES: TYPOLOGICAL IMPLICATIONS	16
1.1. Introduction	16
1.2. Nouns, verbs and flexibles	17
1.3. (Apparently) conflicting argument and predicate markers	24
1.4. The pervasiveness of a typological trait	27
1.5. The five logically possible language types	28
1.5.1. Type <i>N/V/F</i>	29
1.5.2. Type <i>N/F</i>	30
1.5.3. Type <i>V/F</i>	31
1.5.4. Type <i>N/V</i>	32
1.5.5. Type <i>F</i>	32
1.5.6. The five logically possible language types: Summary	34
1.6. Typology of the noun/verb distinction: Questions for future research	34
1.7. Summary	35
2. THE NOUN/VERB AND PREDICATE/ARGUMENT STRUCTURES	37
2.1. Introduction: Language fossils	37
2.2. The linguistic predicate/argument distinction and its relations to other similar distinctions	37
2.3. Noun-argument and verb-predicate correspondences in natural language	39
2.3.1. Adjectives, adverbs and adpositions in the linguistic predicate/argument structure	45
2.3.2. The linguistic predicate/argument conversion system. Noun-argument and verb-predicate correspondences	47
2.4. Summary	48
3. THE EVOLUTION OF THE NOUN/VERB AND LINGUISTIC PREDICATE/ARGUMENT STRUCTURES	50
3.1. Introduction	50
3.2. The evolution of the noun/verb and linguistic predicate/argument structures	50
3.3. Eleven arguments for the evolutionary primacy of LA over LP	53
3.4. Summary	59

SUMMARY AND CONCLUSIONS	61
SUMMARY IN ESTONIAN	65
REFERENCES	88
CURRICULUM VITAE	97

LIST OF PUBLICATIONS

- I. Luuk, Erkki. 1999. Tähenduse lineaarne ja hierarhiline struktuur [The linear and hierarchic structure of meaning]. *Akadeemia*, 8/1999, 1630–1660.
- II. Luuk, Erkki. 2006. Kõnetegude teooria, illokutiivse loogika ja pragmapoeetika analüüs [An analysis of speech act theory, illocutionary logic, and pragmapoetics]. *Akadeemia*, 2/2006, 399–438.
- III. Luuk, Erkki; Luuk, Hendrik. 2008. Evolutionary framework for the language faculty. In: Andrew D. M. Smith, Kenny Smith, Ramon Ferrer i Cancho (eds.). *The Evolution of Language. Proceedings of the 7th International Conference (EVOLANG7)*. New Jersey etc: World Scientific.
- IV. Luuk, Erkki; Luuk, Hendrik. 2008. On certain core properties of the language faculty: an analysis of embedding, recursion and conceptual embedding. In: *Language, Communication & Cognition: Book of Synopses*. Language, Communication & Cognition conference, Brighton, 04–08.08.08.
- V. Luuk, Erkki. 2008. Semantilised tasandid ja semantilised primitiivid [Semantic levels and semantic primitives]. *Keel ja Kirjandus*, 12/2008, 949–967.
- VI. Luuk, Erkki. in press. Nouns, verbs and flexibles: implications for typologies of word classes. *Language Sciences*, doi:10.1016/j.langsci.2009.02.001
- VII. Luuk, Erkki. in press. The noun/verb and predicate/argument structures. *Lingua*, doi:10.1016/j.lingua.2009.02.009

ABBREVIATIONS

1	– first person
2	– second person
3	– third person
A	– argument
ABS	– absolutive case = allomorph of alienable genitive
AC	– argument clause
ADJ	– adjective
ADP	– adposition
ADV	– adverb
AG	– agent
ART	– article
ARV	– actor voice
AUX	– auxiliary verb
AVV	– active voice
COP	– copula
DEF	– definite
DET	– determiner
DP	– determiner phrase
F	– flexible
FOC	– focus
FOPL	– first order predicate logic
GEN	– genitive
GER	– gerund
HUM	– human
IMPF	– imperfective
INDIC	– indicative
INF	– infinitive
INFL	– verbal inflection
IP	– inflection phrase (see INFL)
LA	– linguistic argument
LAx	– linguistic argument marker
LOC	– locative
LP	– linguistic predicate
LPx	– linguistic predicate marker
LP/A	– linguistic predicate/argument
M	– masculine
N	– noun
N/V	– noun/verb
NL	– natural language
NP	– noun phrase
O	– object
OBL	– oblique
P	– predicate

P/A	– predicate/argument
PAST	– past tense
PC	– predicate clause
PJ	– projection
PL	– plural
POSS	– possessive
REAL	– realis mood
REFL	– reflexive
S	– sentence
S/NP	– sentence/noun phrase
S/XP	– sentence/x phrase (see fn. 10)
SG	– singular
SOPL	– second order predicate logic
SOV	– subject object verb (word order)
TAM	– tense-aspect-mood
V	– verb
VC	– voice
XP	– x phrase (see fn. 10)

Conventions

Language types are set in capital letters and italics, e.g., *N/V*, *N/V/F* etc.

INTRODUCTION

The objects of this thesis are the noun/verb and predicate/argument structures. Three main issues that the thesis investigates are as follows:

1. Is the noun/verb distinction universal in the world's languages?
2. Is there a correspondence between the noun/verb and predicate/argument structures?
3. What (if anything) could be conjectured about the evolution of the noun/verb and predicate/argument structures besides the considerations put forward in Heine and Kuteva (2002, 2007)?

All these issues are complex and currently unresolved. Although 1–2 are formulated as yes/no questions, one can scarcely hope for simple yes/no answers to them. As we will see in the following chapters, 1–3 expand to a series of subproblems that must be solved before the answers to the original questions can be sought. Much of the thesis details and addresses these subproblems. The structure of the thesis is straightforward: the first chapter deals with issue 1, the second chapter with issue 2, and the third one with issue 3. However, depending on one's perspective, the treatment of subproblems (e.g., the (im)possibility of formal and cross-linguistically universal definitions of N and V) may be even more important than that of the original ones.

Whereas predicate/argument structures are parts of natural language and mathematical logic (as well as, perhaps, parts of cognition and perception in some species – Hurford 2001, 2003b, 2003c), the noun/verb structure pertains to natural language only. However, it is not clear whether all languages have nouns and verbs (Anderson 2004; Bach 2004; Laudanna & Voghera 2002). Although the relationships between the linguistic, the logical, and the cognitive/perceptual predicate/argument structures are discussed (in sections 2.2, 2.3 and 3.2), the thesis focuses on the linguistic predicate/argument distinction. A claim is that the linguistic predicate/argument structure is a universal characteristic of all sufficiently developed human languages, both possible and actual. Insofar as possible, the claim is tested in sections 1.1 and 3.2. I do not maintain that the claim itself is new (although I have never seen it formulated this way before). However, a number of implications of the claim that are pursued in chapters 1–3 are novel, as are the analyses of the noun/verb and linguistic predicate/argument structures in chapters 1–2.

As mentioned above, one of the most controversial subjects in the typology of word classes is whether all languages have nouns and verbs. There is a long tradition of conflicting opinions and lack of consensus on this matter (cf. Baker 2003; Broschart 1997; Evans & Osada 2005; Hopper & Thompson 1984; Jacobsen 1979; Peterson 2007; Whorf 1945). The research that forms the foundation of chapter 1 began as an attempt to investigate the roots of this problem (Luuk in press-b). Soon it became obvious that one of the major obstacles for consensus in this area is the lack of rigorous and cross-

linguistically valid definitions of ‘noun’ and ‘verb’. While it is possible that formal and universal pre-established categories do not exist (Haspelmath 2007), ‘noun’, ‘verb’ and ‘word’ are, at the very least, useful approximations thereof, the definitions of which must be as precise and universal as possible. After analyzing a few most well-known definitions of ‘noun’ and ‘verb’ that aim at both rigor and cross-linguistic applicability (Baker 2003; Croft 2000, 2001), and being somewhat dissatisfied with them, I started work to narrow down my own definitions, finally arriving at formulations (4)–(5) in section 1.2. A feature of definitions (4)–(5) is that they are both formal and substance-based, as they are tied to specific grammatical markers (the linguistic predicate/argument markers as defined in (7)–(8)) as well as to the semantic substance of the linguistic predicate/argument structure. Chapter 1 makes it clear that linguists are very far from answering the question of the universality of the noun/verb distinction. As the problem of the universality of the noun/verb distinction is inseparable from the problem of universal definitions of ‘noun’ and ‘verb’, a solution to the former depends on the solution to the latter (but not vice versa). As long as the definitions of ‘noun’ and ‘verb’ (and more generally, the criteria for certain linguistic categories as well as for ‘linguistic category’ as such) are not agreed upon, no progress on the problem of the universality of the noun/verb distinction is possible. Currently, there is no agreement on the definitions of ‘noun’, ‘verb’ and ‘word’ (cf. Anderson 2004; Anward 2001; Baker 2003; Broschart 1997; Croft 2000; Di Sciullo & Williams 1987; Dixon & Aikhenvald 2002; Greenberg 1963; Sasse 1993b). I can only hope that chapter 1 has met some success in identifying relevant criteria for these linguistic categories.

Further, and rather to my surprise, the research behind chapter 1 led to several specific predictions for the world's languages, and for typologies of word classes in particular. An intuitive appeal of the logico-typological framework presented in sections 1.5 and 1.7 is in its parsimony and simplicity. Essentially, the whole framework is built on three foundational cross-linguistic universals: 1) the linguistic predicate/argument structure, 2) lexical class, and 3) a correlation between the two. Given these three premises, one arrives at this framework by logical inference alone. Thus, a feature of the framework is logical necessity. Due to its very nature, the framework makes exact predictions as to the logically (im)possible language types (see section 1.5). However, it cannot predict which of the five logically possible language types are realized in the world's languages (other than that at least one of them¹ must, by definition, be realized). A more exact answer to the question of which language types actually exist(ed) in the world requires a colossal work of descriptive and typological linguistics, far beyond the scope of chapter 1 (and of this thesis as a whole).

An important result of chapter 1 is the substantiation of the word type ‘flexible’ as the parsimonious alternative to zero derivation and homophony (see section 1.2). Flexibles are the stems that are used nominally as well as

¹ Namely, type *N/V/F* (see section 1.5.1).

verbally, such as the English *walk*, *run* and *lock*. Other noteworthy results in chapter 1 include the definitions of ‘noun’, ‘verb’, ‘flexible’ and ‘word’, and the framework of the five logically possible language types (*N/V/F*, *N/F*, *V/F*, *N/V* and *F* – see section 1.5).

Jackendoff (1999) has used the notion ‘language fossil’ for evolutionarily basic structural-functional types in language. His list of language fossils includes principles such as grouping (modifiers appear adjacent to what they modify), Agent First, Focus Last, et al. I suggest that the linguistic predicate/argument structure qualifies as an evolutionarily basic structural-functional type, and should thus be included among language fossils. According to Heine and Kuteva (2002, 2007), nouns and verbs emerge at the earliest stages of language evolution. In chapter 2, I give a detailed overview of the linguistic predicate/argument structure, and show how noun – linguistic argument and verb – linguistic predicate correspondences are established in natural language (Luuk in press-a). The latter is important, as it has been standardly assumed that (1) the predicate/argument structure of natural language corresponds closely or is identical to that of first order predicate logic, and (2) there is no noun-argument and verb-predicate correspondence in natural language, as the same kind of term can appear as both argument and predicate (as, e.g., *a man* in *A man dies* and *Plato is a man*). Hurford (2003b, 2003c) refers to (2) as the ‘Aristotle problem’. In chapter 2, I show that the Aristotle problem arises from assumption (1). The solution to the problem lies in acknowledging that the linguistic predicate/argument structure is more complex than that of first and second order predicate logics, with up to three levels of linguistic arguments and linguistic predicates and conversion rules from linguistic argument to linguistic predicate and vice versa. A detailed account of the functioning of the linguistic predicate/argument structure is given, complete with sets of rules which generate higher order linguistic predicates and arguments, and allow linguistic predicate/argument conversion (see sections 2.3–2.4). This is the most important result in chapter 2.

Certain asymmetries in these rules, in the linguistic predicate/argument structure, and in predicate/argument structure as such, suggest that linguistic arguments may be evolutionarily more fundamental than linguistic predicates. Previously, a similar argument has been made about nouns and verbs within the framework of grammaticalization (Heine & Kuteva 2002, 2007). Chapter 3 analyzes this claim. First, some general assumptions and observations are made about the evolution of the linguistic predicate/argument structure. Quite obviously, the linguistic predicate/argument structure must reflect on the conceptual level, as it requires the ability to conceptualize the corresponding functions. As the linguistic predicate/argument structure is functionally motivated by, e.g., the ability to talk about events (i.e. objects/properties caught in actions/changes²), it is plausible that the underlying ability to conceptualize

² Prototypically, linguistic arguments refer to objects/properties and linguistic predicates refer to actions/changes.

events in the above defined sense (which relies on a conceptual predicate/argument structure) predates the linguistic predicate/argument structure, i.e. predicate/argument marking in language. This conjecture is supported by the fact that the linguistic predicate/argument structure is useless without the ability to conceptualize the functions of linguistic argument and linguistic predicate but the ability to conceptualize events is useful even in the absence of language (e.g., in behavioral planning, which increases the individual's fitness). Then, if Hurford (2001, 2003b, 2003c) is correct about the cognitive/ perceptual predicate/argument structure being common among mammals, it is plausible that the cognitive/perceptual predicate/argument structure was mapped onto conceptualization in some species³, resulting in a conceptual predicate/argument structure, which in turn gave rise to the linguistic predicate/argument structure in humans (after the emergence of protolanguage – cf. Bickerton 1998, 2000, 2002, 2007; Jackendoff 1999; Nowak & Krakauer 1999; Wray 2000). Section 3.3 analyzes the question of a possible evolutionary priority of linguistic argument over linguistic predicate or vice versa, concluding, with a number of new (as compared to Heine & Kuteva 2002, 2007) arguments from a variety of domains that the evidence for linguistic arguments predating linguistic predicates is overwhelming (Luuk in press-a).

³ Conceptualization requires a capacity for mental imagery (observe that mental imagery does not have to be visual or visuomotor but can be auditory, olphactory, etc. as well – Belardinelli et al. 2004; Bensafi et al. 2003; Cooper 1995; Klatzky, Lederman & Matula 1991; Pylyshyn 2003; Reisberg, Wilson & Smith 1991). So far, only certain species have been shown to be capable of mental imagery (cf. Herman, Richards & Wolz 1984; Mauck & Dehnhardt 1997; Savage-Rumbaugh, Shanker & Taylor 1998; Vauclair, Fagot & Hopkins 1993). However, there is every reason to believe that the list of species possessing mental imagery is substantially longer than could be inferred from these four citations.

I. NOUNS, VERBS AND FLEXIBLES: TYPOLOGICAL IMPLICATIONS*

I.1. Introduction

Nouns and verbs appear to be the most common among lexical categories (Hockett 1966; Sapir 1978 [1921]). 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 & 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 hold up to scrutiny (e.g., Baker 2003; Hopper & 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, 2005b; Marantz 1984, 1997) of them agree on the following⁴:

1. N and V are lexical categories.
2. Stems rather than their functional heads (determiner, aspect, etc.) bear lexical categories.
3. N and V 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 1993b).
- * discourse-manipulable participant / reported event (Hopper & Thompson 1984).

* This chapter is based on Luuk (in press-b).

⁴ 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).

- * time-stable / non-time-stable concepts (Givón 2001; Heine & 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 (2003b, 2003c) 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: 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 – Luuk] as either **masok (putih)** “the white one is going in” or **putih (masok)** “the going is of the white one”.

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 signaled 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.

1.2. 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 rigor 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: 87–88; Croft 2000: 88–89; Croft 2005: 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. 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)⁵. 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: 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-3SG
 'It gets dark.'

⁵ Haspelmath (2007), for example, claims that universal pre-established (a priori) 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 language, there are three possible kinds of universal formal categories: innate, acquired, and established only in research. A purportedly universal formal pre-established category is dismissed if we find a language without it. Observe that, even if one could, in this way, succeed in dismissing all the putatively universal formal pre-established categories (a perspective in which I sincerely doubt), there will always remain the possibility of a universal formal category established a posteriori in research. Thus, there is always the possibility that universal formal categories exist. 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). Methodologically, it is extremely difficult to ascertain whether a particular universal formal category is innate, acquired, or established only in research. The problem lies at the level of epistemology already: there are, uncontroversially, linguistic categories apart from linguists’ categorizations, yet linguists have access to the former only by the latter. It is in fact possible for the latter to be **better** (more universal, formal, exact and/or parsimonious) than the former.

(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 thus 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: 78–9):

- (b) *mamu:k-ma* *qu:ʔas-ʔi*
 work-INDIC man-the
 'The man is working.'
- qu:ʔas-ma* *mamu:k-ʔi*
 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:ʔas* 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, respectively⁶. Arguments and predicates are not missing, though – they are encoded by the functional heads *-ma* and *-ʔi*. 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.

⁶ Observe that arguments are also encoded by NPs, and predicates are also encoded by COP+NP and COP+ADJ constructions (see section 2.3 for more details).

All these hypotheses are untestable but (3) is the most parsimonious⁷. 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 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 encode predicates and arguments while being underspecified with respect to the N/V distinction. Borrowing a term from a related notion, ‘flexible parts-of-speech systems’ (e.g., Don & van Lier 2007; Hengeveld 1992; Hengeveld & Rijkhoff 2005; Rijkhoff 2002), I call this class flexible (F)⁸. In encoding predicates and arguments, then, English has the 3-way distinction of N/V/F instead of just N/V (see section 1.5.1). 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/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 F⁹.

⁷ An anonymous reviewer argued that semantics (e.g. polysemy) may provide crucial evidence in deciding between (1)–(3). Essentially, the choice between (1)–(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.

⁸ 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 ‘precategoryal’ (Evans & Osada 2005). Apart from focusing on different aspects of the same phenomenon, all these differences are largely terminological.

⁹ 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 work 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’.

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 (Oxford 1991: 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 property of stems to accept LA markers but not LP markers.
- (5) V = the property of stems to accept LP markers but not LA markers.
- (6) F = the property of stems to accept 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) (additional details on the LP/A distinction are given in section 2.2). 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, 2005b; Marantz 1984), we could have definitions like “N = the property 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, shared by the majority of researchers (e.g., Anderson 2004; Anward 2001; Croft 2000; Crystal 2004; Hopper & Thompson 1984; Jespersen 1924; Langacker 2004; Lyons 2004; Sasse 1993b; 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 embarrass*, **a comprehend*, **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 property 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 2005b: 354–5), 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’s (2002, 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.

Word has traditionally been a difficult concept to define (Broschart 1997; Di Sciullo & 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 for ‘phonological word’ has no bearing on semantics or grammar. Clearly, this is not what is meant by ‘word’ 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 *Kaipiallulliniuk* (‘the two of them were apparently really hungry’ – Mithun 1999: 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 properties to accept LAx and/or LPx. Instead, the derivational affix has changed the property of the higher order stem. Correspondingly, we have to assume that the higher order stem’s property overrides the properties 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, accept 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. (4)–(6) and Bhat 2000; Croft 2005; Peterson 2007; Sasse 1993b). 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.

Depending on the distributional criteria (see below), only a subset of determiners, possessives, TAM and voice markers in the language may qualify as LA and LP markers. As suggested with *masok putih* in section 1.1, 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 & Thompson 1984; Peterson 2005, 2007; Sasse 1993b). 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 & Sadler 2004). Importantly, however, TAM marking on N or DET appears to be restricted to a proper subset of 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 & van Everbroeck 2003). In most of the 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 F¹⁰. This makes the particular LA marking, of course,

¹⁰ Hurford (2007) avoids the DP vs. NP debate (e.g., Bošković 2007; Bruening 2008; Lulu & 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

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 in the applicability of gender marking as such, not in 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 & Osada 2005).

1.3. (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: 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 as an LA or LP 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:

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.

- (e) Tagalog (Himmelmann 2007)
- | | | | | | | | |
|------------|----------------|------------|------------|-------------------|-----------|------------|-------------|
| <i>ang</i> | <i>langgam</i> | <i>rin</i> | <i>ang</i> | <i>t-um-ulong</i> | <i>sa</i> | <i>mga</i> | <i>bata</i> |
| 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: 280)
- | | | |
|-----------|------------------|------------------------|
| <i>iñ</i> | <i>ho-kaR-te</i> | <i>iñ-aʔ-y-oʔj</i> . |
| 1SG | 3-SG.HUM-OBL | 1SG-GEN-y-AVV.PAST.1SG |
| I | him/her | mine made |
- 'I adopted her.'

Here we have [[*iñ-aʔ* POSS]-*y-oʔj* VC/TAM], 'mine made'. POSS marks argument; VC and TAM mark predicate. As the outmost markers are predicate markers, *iñ-aʔ-y-oʔj* is a predicate¹¹. Observe also that the rule that the outmost marker determines element's identity is in concordance with the rule, specified in section 1.2, that the higher order stem's property to accept LAX/LPx overrides the properties of lower order stems.

Arguments can also be incorporated within (or converted to) predicates, as in the following example (a more detailed analysis is given in section 2.3):

- (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 *this*][P *is* [A *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

¹¹ 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).

- (h) Russian
eta moya kniga
 this my book

analyzable as [A *eta*][P [A *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 dolzhna 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).

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' acceptance of LAs/LPs 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's (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 trilemma¹². Needless to say, failures 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 other¹³. A necessary result of the situation is that the scholars applying inclusive criteria find N/V in all languages they investigate (Baker 2003; Evans & Osada 2005; Hopper & 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).

1.4. 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's (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: 67–68). Most researchers would reject this kind

¹² Assuming that (10) is a subcase of (11), the two are not necessarily mutually exclusive.

¹³ Typologically, it would be perhaps wiser to apply inclusive criteria (cf. the penultimate paragraph of section 1.2).

of reasoning. Indeed, it has been stressed that the pervasiveness criterion must be applied (Baker 2003; Evans & 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 & Rijkhoff 2005; Hengeveld, Rijkhoff & Siewerska 2004).

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

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 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 onto. I make no claim as to the reality of any of these types except *N/V/F* (see section 1.5.1). The present typology predicts, however, that types *N* and *V* (lacking both *F* and *V* or *N*, respectively) are not expected to be found among the languages of the world (as they would violate (12) and (13)). 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 1.5.2 and 1.5.3, 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.

1.5.1. Type *N/V/F*

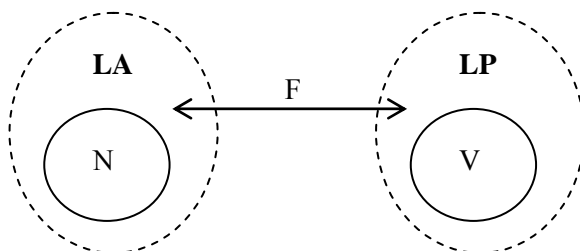


Fig. 1. Type *N/V/F*.

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, $F = (LA - N) \cup (LP - V)$. As we see in (90)-(91) in section 2.3.2, the union set of *F*, *N* and *V* is only a proper subset of the union set of *LA* and *LP*; accordingly, figures 1–5 describe only proper subsets of *LA* and *LP*. *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) of 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*, *engage*) and flexibles (*round*, *love*, *kill*,

walk, run) (Bierwisch 2001; Don & van Lier 2007; Farrell 2001; Jespersen 1924; Vogel 2000).

1.5.2. Type *N/F*

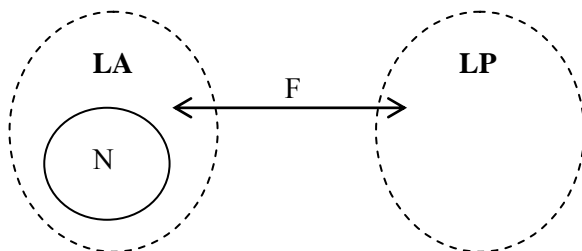


Fig. 2. Type *N/F*.

Type *N/F* (Fig. 2) has been proposed for Maori (Biggs 1971) and Niuean (Massam 2005). However, Bauer (1993: 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 (1993b) for a short overview on languages that have been claimed to belong to the logically impossible types *N* and *V*.

I.5.3. Type V/F

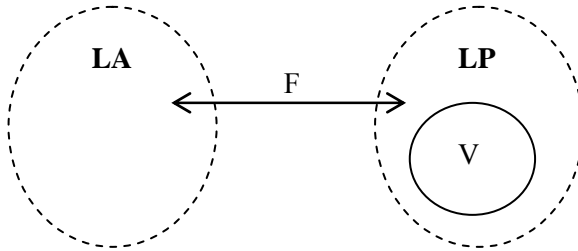


Fig. 3. Type V/F.

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 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: 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 (1993b, 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 languages already at the stem level. In Cayuga, only nouns take possessive prefixes and the noun suffix *-a?*, and only verbs are TAM-marked (Mithun 2000). Presently, the dominant opinion is that both Cayuga and Tuscarora have nouns (Anward 2001). Sasse (1993b, 2001) suggests that Cayuga has nouns (that may, however, constitute a closed class), while giving a more hesitating estimate elsewhere (Sasse 1993a). 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.

1.5.4. Type *N/V*

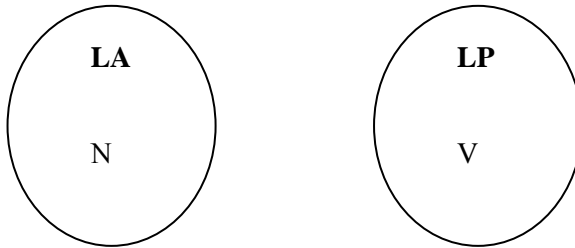


Fig. 4. Type *N/V*.

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.).

1.5.5. Type *F*

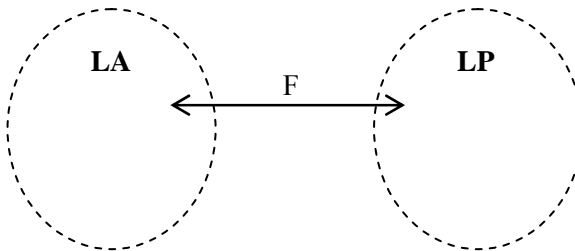


Fig. 5. Type *F*.

A number of languages have been proposed to belong to this type¹⁴. The condition of type *F* is that the language has no stems that map to either LA but

¹⁴ Languages of this type are alternatively labeled as ‘precategoryal’ or having ‘flexible parts of speech’ (Don & van Lier 2007; Evans & Osada 2005; Hengeveld 1992; Hengeveld & Rijkhoff 2005; Rijkhoff 2002). The differences between these notions, if applicable, and type *F* are largely terminological.

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 & Kinkade 1998: 36, on Salishan). At least the following languages have, sometimes tentatively, been proposed to belong to type *F*:

1. Wakashan languages, spoken in Northwestern North America, e.g. Nootka, Makah, Nitinat and Kwakiutl (Boas 1947; Swadesh 1939; Whorf 1945; cf. Jacobsen 1979).
2. Salishan languages, spoken in Northwestern North America, esp. Straits Salish, Squamish and Upper Chehalis (Czaykowska-Higgins & Kinkade 1998; Jelinek & Demers 1994; Kinkade 1963; cf. Sasse 1993b).
3. Munda languages, spoken in India, esp. Mundari and Kharia (Bhat 2000; Hengeveld & Rijkhoff 2005; Peterson 2005, 2007; Rijkhoff 2003).
4. Several Malayo-Polynesian languages (Tongan, Samoan, Tagalog, Kambera¹⁵, 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 & Hovdhaugen 1992, as reviewed in Baker 2003: 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.

¹⁵ However, Klammer (1998, 2005) has offered convincing evidence that Kambera is of type *N/V/F*.

1.5.6. The five logically possible language types: 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 1.4).

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 fn. 13, the inclusive criteria are preferable. As argued in section 1.4, 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* 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.

1.6. 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:

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 1.5.5 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*?

1.7. Summary

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 syntactic 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 syntactic 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 property of stems to accept LA markers but not LP markers.
- (26) V = the property of stems to accept LP markers but not LA markers.
- (27) F = the property of stems to accept 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 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.

2. THE NOUN/VERB AND PREDICATE/ARGUMENT STRUCTURES[†]

2.1. Introduction: Language fossils

As Ray Jackendoff has noted, the claim that language does not fossilize is not entirely correct (Jackendoff 1999). While it is true that “linguistic behavior does not fossilize” (Hauser, Chomsky & Fitch 2002: 1571), this does not preclude the existence of language fossils as structural-functional types¹⁶. It is plausible that such ‘language fossils’ have been maintained in modern language since no better functional analogues have been and possibly could not be developed. The evolution of language, especially the evolution of grammar and vocabulary has largely been driven by a need for higher precision and explicitness (cf. Heine & Kuteva 2002, 2007; Newmeyer 2004). By hypothesis, this development has left some structural features bearing on the core communicative functions of language intact. These structural features are ‘language fossils’. Language fossils are focussed on speed or some other basic aspect of communication (see below), and form a subset of natural language structures that function independently of the “higher” principles of grammar which craft the meaning of complex propositions. Jackendoff proposes the following set of possible language fossils: 1) situation-specific symbols that cannot be integrated into syntax (except in quotes) like *hey!*, *ouch!*, *wow!* and *shh!*, 2) noun-noun compounds like *snowman*, *bellboy*; principles like 3) grouping (modifiers appear adjacent to what they modify), 4) Agent First (as in *he hit me*) and 5) Focus Last (as in *the book was dull*). For a related example, imperative mood verb forms (e.g. *Run!*, *Go!*), like the first type of Jackendoff’s language fossils, are usually shorter than all others and syntactically independent, which is an evidence of their adaptation for rapid reaction and articulation/parsing speed, important in crisis situations.

2.2. The linguistic predicate/argument distinction and its relations to other similar distinctions

To Jackendoff’s list of language fossils one more can be added: the linguistic predicate/argument (LP/A) distinction. It is unclear whether the noun/verb (N/V) distinction is universal in the world’s languages (Anderson 2004; Bach 2004; Laudanna & Voghera 2002), but it is generally agreed upon that the

[†] This chapter is based on Luuk (in press-a).

¹⁶ Observe that the use of the word ‘fossil’ here is different from that of in Bickerton’s (1990) ‘fossils of language’. For Bickerton, the fossils were not structural types but primitive ‘languages’, such as pidgins, the ape ‘language’, the ‘language’ of under-tuos and the ‘language’ of Genie.

syntactic functions of nouns and verbs are linguistic argument (LA) and linguistic predicate (LP), respectively (e.g., Bhat 2000; Croft 2005; Peterson 2007; Sasse 1993b), and that the LP/A distinction is universal in the world's languages. Without further measures taken (see section 2.3), nouns function as LAs but not LPs, and verbs function as LPs but not LAs. In chapter 1 we saw that those languages that putatively lack the N/V distinction are claimed to have, instead, a class of lexical items (stems and words) that can function as both LAs and LPs (Czaykowska-Higgins & Kinkade 1998; Jelinek & Demers 1994; Kinkade 1963; Swadesh 1939). LA and LP refer to NL arguments and predicates **only**. This is important, as LA and LP do not have to correspond to first (or second) order predicate logic arguments and predicates. There is no confusion over this matter: whenever LA or LP is mentioned, I am referring to NL not predicate logic arguments and predicates. Throughout this work, all references to predicate logic arguments and predicates are clearly marked. Below I give an overview of the relationships between the LP/A and other similar distinctions.

1. **The logical predicate/argument distinction.** Conventionally, capital letters denote predicates, lower-case letters denote arguments, and predicates precede the arguments they take in predicate logic. Thus, *John loves Mary* would be, e.g., *Lj,m* and *Plato is a man* would be *Mp*. Similarly, *A man dies* would be *Dm*. A discrepancy between NL and predicate logic is obvious: in the above sentences, the LA *a man* is the same, while in predicate logic, *M* and *m* are nonequivalent – *M* is a predicate and *m* is an argument. Worse, there is no obvious way to derive *M* from *m* or vice versa – at least not in FOPL and SOPL (which are distinct from higher order logics in not allowing predicates as arguments of higher-order predicates). Thus, there is a discrepancy between the predicate/argument distinction of NL and that of FOPL and SOPL.

2. **The linguistic predicate/argument distinction.** We arrive at the definitions of LA and LP in section 2.3.2. LAs are nouns, flexibles bearing LA markers, adjectives bearing LA markers, the products of rules (b), and XPs. LPs are verbs, flexibles bearing LP markers, adjectives bearing LP markers, the products of rules (a), and predicate phrases. XP and predicate phrase are distinct from the others in that they are projections of other LAs and LPs, respectively. Thus, LA and LP designate the original LAs, the original LPs, and projections thereof. A more detailed account, given in the following four sections, is impossible at this point (but see points 5–6 below). Presently, it suffices that we observe the differences between the LP/A and all the other distinctions discussed in this section.

3. **The perceptual predicate/argument distinction.** According to Hurford (2001, 2003b, 2003c), the perceptual predicate/argument structure is a neuroanatomically motivated construct. The perceptual predicate/argument structure differs from the logical and the linguistic ones in being more elementary, e.g. *A man runs* would translate into [*A a man*][*P runs*], *Rm* and *MAN(x)&RUN(x)* in the linguistic, the logical and the perceptual P/A analyses,

respectively (MAN and RUN are perceptual predicates and x is the perceptual argument).

4. **The subject/predicate distinction.** Although weakly similar to the LP/A structure, this structure allows only for the subject as argument, i.e. it is incomplete as compared to LP/A which allows for *n*-place predicates. Other differences include VPs (as compared to predicate phrases in LP/A) for predicates and a disregard for the circumstance that predicatives can be arguments, as in *John* [*P is* [*A a man*]]. A crucial difference between VP and predicate phrase is that only the former can incorporate objects in transitive constructions (cf. (32)).

5. **The noun/verb distinction.** Nouns and verbs form a proper subset of LAs and LPs, respectively. N and V are defined as follows. N = the property of stems to accept LAx but not LPx. V = the property of stems to accept LPx but not LAx. LAx and LPx are defined as follows. LAx = (a subset of) determiners, possessives and LP/A word order constraints. LPx = (a subset of) tense-aspect-mood, voice markers and LP/A word order constraints. I fully appreciate the difference between lexical and phrasal categories; however, lexical categories project phrasal categories and, in this case, it is the properties LA and LP that are projected (from N and V to XP and predicate phrase, respectively). Thus, the lexical and the phrasal categories share these properties, which is the reason why both N and XP are labeled LA and both V and predicate phrase are labeled LP. Cf. below.

6. **The NP/VP distinction.** NPs or, more precisely, XPs are a proper subset of LAs. As the concept of VP is fundamentally vague or, if we adhere to its common interpretation V+AUX+O, incompatible with the LP/A structure (see the discussion of (32) in 2.3), I have substituted it with ‘predicate phrase’, which cannot include objects, i.e. LAs. Predicate phrases are a proper subset of LPs. XPs and predicate phrases are the LAs and LPs that are projections of the original LAs and LPs, e.g. N and V (cf. points 2 and 5 above).

7. **The topic/comment distinction.** In English, there is a strong tendency of conflating the subject XP with topic and LP with comment, as in *John is ill*, *I run*, *The people were lied to* etc. The tendency, however, is not universal, cf. *Martians landed!*, or *There is a pig in the garden*, which, according to Hurford (2007), have no topic (only comment). Cf. *I lost everything* and *Everything is lost* – in these sentences, only *I* qualifies as topic, the rest falls under comment. Thus, a major difference between LP/A and topic/comment is that topic/comment do not correspond to specific word classes (such as N and V in the case of LP/A).

2.3. Noun-argument and verb-predicate correspondences in natural language

James Hurford (2003b, 2003c) has pointed to a problem with Aristotle's predicate/subject (or predicate/argument – P/A) structure. Essentially, the

problem is that the same kind of term can fill both the argument and the predicate slot. For example, a term such as *a man* can be the argument of *A man died* and the predicate of *Plato is a man*. First and second order predicate logics, on the other hand, are more distanced from the surface forms of natural languages, and the same terms cannot be both arguments and predicates (cf. section 2.2). Hurford concludes that it remains to provide an explanation for the typical structure of modern languages, organized around the noun/verb dichotomy. Below I will provide the first part of an explanation by showing how the correspondence between nouns and verbs and NL predicates and arguments can be established. As any predicate logic is an artificial system which NL has no need to correspond to, the second part, establishing the correspondence between NL and predicate logic arguments and predicates, is a technical matter of representing NL predicates and arguments in predicate logic. Here are some natural language (NL) examples analyzed with respect to their P/A structure ([A ...] marks argument and [P ...] predicate). Observe that these are still linguistic not logical predicates and arguments – A and P are used instead of LA and LP only for the sake of brevity:

- (31) [A *John*][P *runs*]
- (32) [A *John*][P *loves*][A *her*]
- (33) [A *John*][P *is*][A *a man*]

It is arguable whether *loves her* in (32) should be analyzed [P *loves*][A *her*] or [P *loves*][A *her*]. The latter concurs better with NL notation, where [P *loves*][A *her*] is interpreted as a VP. However, VP and NP are themselves logical abstractions and thus may be no better than first and second order predicate logic. I prefer the simpler interpretation [A *John*][P *loves*][A *her*], which has the logical form *LOVES*(x,y), and the order of arguments (x,y) stipulated by case, with the first argument (Subject) *John* taking nominative and the second argument (Object/Patient) *her* taking accusative. From the viewpoint of logic, the difference between [P *loves*][A *her*] and [P *is*][A *a man*] makes sense, as *love* is a verb that has necessarily 2 arguments. At the same time, I would not say that *be* (or *is* – see fn. 17) is a verb that has necessarily two arguments (cf. *I think, therefore I am*). Similarly, *loves her* is, uncontroversially, a verb followed by an object. Few, if any, linguists would claim that *is a man* is a verb followed by an object. It is standardly assumed that, if copula is a verb at all, it is a special kind of verb (see fn. 17), and I am not aware of predicate complements being analyzed as objects.

I suggest that in English (and a myriad of other languages), copula heads secondary predicate, as in (33).¹⁷ Alternatively, in some languages like Russian,

¹⁷ Similar rules of type-shifting have been proposed in, e.g., Partee (2002) and Chierchia (1998) within the framework of formal semantics, which views NPs as generalized quantifiers. This, in turn, necessitates type-shifting for referential and predicative readings of NPs. Working in predicate logic, Partee and Chierchia align linguistic

Maltese and Moroccan Arabic, secondary predicate can be specified by juxtaposing two arguments (cf. section 1.3). Here is an example from Russian:

- (34) [A *On*][P[A *sportsmen*]].
 Lit: He sportsman.
 'He is a sportsman.'

The Russian *sportsmen* is an argument, as it can be marked by LA markers, e.g. determiners (*etot sportsmen* – this sportsman) but not LP markers, e.g. TAM (**sportsmen-al* – sportsman-PAST.M – **sportsmanned*). In addition, if we reverse the order of words in (34), we get *sportsmen on* – ‘the sportsman is he’, where *sportsmen* functions exactly as *on* did in (34). In English, the suffix *-ing* derives a secondary argument from a verb or flexible stem:

- (35) [A *He*][P *started*][[*fir*]A *-ing*].

The resulting word, a gerund, is a V_s/F_s -*ing* construction in English (the notation is explained in fn. 18). As English gerund marking is identical to that of imperfective participle (or ‘present participle’), the two should not be confused. (36) is an example of gerund and (37) of the participle:

- (36) *This is *a/the running I expected from you.*
 (37) *John is (*a/*the) running.*

Gerunds and infinitives (e.g., *to*+ V_s/F_s constructions) are secondary arguments derived from stems of primary predicates V and F. Here is an example of English infinitive:

- (38) [A *I*][P *want*][A *to*[P *go*]].

predicates and arguments differently not only from what has been proposed here but also from one another. According to Chierchia (1998), bare lexical nouns can be predicates, arguments, or both, depending on the language (he suggests, however, argumental to be the default type acquisition-wise). In traditional formal semantics, however, there would be no difference in meaning at all between *cat* and *be a cat*, *asleep* and *be asleep*, etc. (Partee, 2002). According to Partee (2002), the English *be* is a potentially universal operator that is always available to turn an $\langle\langle e, t \rangle, t \rangle$ (“generalized quantifier”) meaning into an $\langle e, t \rangle$ (“predicative”) meaning. However, it should be noted that, according to Becker (2004), English has two copulas: one is a verb (*be*), while the other is an INFL head and is not raised from a V position (*is*, *am*, etc.). A more traditional analysis would view copulas as special kinds of verbs (Afarli & Eide 2000; Bowers 1993; Eide & Afarli 1999). To avoid positing a separate word class for it, the copula *is* (if it exists) or *be* (if *is* does not exist) is analyzed as a subcase of verb in the present paper, as it accepts LP (TAM) but not LA marking.

In all languages, infinitives and gerunds function as linguistic arguments. A difference between infinitives and all other primary and secondary LAs in English is that infinitives cannot be marked by LA markers and do not project XPs, as it would violate the constraint of adpositions being peripheral constituents of LAs and LPs (cf. section 2.3.1). Below are two examples:

(39) *I liked the swimming.*

(40) **I liked the to swim.*

The adposition *to* cannot appear in the middle of LA, as in (40). As verbs and flexibles can take XP complements, so can gerunds and infinitives, as they are derived from V_s/F_s . Examples are given in (75)-(80). In Russian, gerunds and infinitives are used similarly to English:

(41) [A *On*][P *ljubit*][A *chitat'*]. He likes to read

(42) [A *On*][P *ljubit*][A *chtenie*]. He likes reading

It is well known that whole clauses can function as arguments under the main clause predicate (Dixon 2006). These clausal arguments are called complement clauses (or argument clauses), and are generated by appending a clause to a complementizer, as in *I know **that** he will come*; *I know **where** you hid it*; *I know **when** it happens*; *I know **what** you did*. Alternatively, *that* can be omitted, as in *I know he will come*. Observe that the clauses themselves are comprised of LAs and LPs and are predicative in nature (plausibly because they are headed by V or F). Correspondingly, the term ‘predicate clause’ can be used for all main and unmarked clauses (as explained above, the marking is done by complementizers or clause order). Some examples follow (AC = argument clause, PC = predicate clause):

(43) [PC [A *I*][P *know*][AC *that* [A *John*][P *knows*]].

(44) [PC [A *He*][P *said*][AC [A *winning*][P *was easy*]].

(45) [PC [A *Who*][P *knows*][AC *what* [P *happened*]]?

ACs and PCs are necessarily higher-order units than LAs and LPs, as PCs and ACs are necessarily comprised of LAs and LPs. At the same time, I am aware of only one possibility how an LA or LP could consist of a PC or AC. The possibility is complex XP, like *the things (that) you said* or *the man who was ill*. In complex XPs, an AC is adjoined to an LP (or an LP followed by an LA) to produce a higher-order LA (the XP). The rule for this is as follows (see fn. 18):

(46) {LP+}[LA+]AC \Rightarrow LA₃ (*He is/knows [the man] who was ill*)

(46) stipulates that ACs preceded by LPs (which are optionally followed by an LA) are rendered as tertiary LAs. Below are some examples:

- (47) *I know [AC₁ who was ill].*
 (48) *I know [[A₁ the man][AC₁ who was ill]].*
 (49) *He started [[A₂ the nagging][AC₁ that eventually made me sick]].*
 (50) *He insulted [[A₁ the writer][AC₂ (that)[A₁ the man][AC₁ who was ill]] had seen before].*
 (51) *He pondered over [[A₁ the insults][AC₃ (that)[A₁ the writer][AC₂ who had seen [A₁ the dentist][AC₁ who was ill]] had written]].*

Three levels of clausal embedding, as in (51), is about the maximum in normal (i.e. non-metalinguistic, non-juridical and non-nonsensical) natural language usage (cf. Karlsson 2007a, 2007b). In (47)–(51), the outmost [...] designates LA₃. LA₃s are normal LAs in every respect. For example, they can be incorporated into PCs as well as into secondary ACs. Below are two examples:

- (52) [PC [A₃ *The man who was ill*] *recovered*.]
 (53) [PC *John said* [AC₂ (that) [A₃ [A₁ *the man*][AC₁ *who was ill*]] *recovered*].]

In (46), the LA on the left of the double arrow can be an LA₃. Below is an example:

- (54) *I know [[A₃ who was [A₁ this man]][AC₁ who died]].*

Here, as in (47), the outmost [...] designates AC and LA₃ (by rule (46)). In (54), *who was this man* qualifies as an LA₃; cf. *I know [who was [A₁ this man]]*.

The fact that PC is headed by a predicate does not imply that PC is a predicate (cf. section 3.2). PC is a sentence. Remember what Meillet and Bloomfield said about the sentence (Graffi 2001): sentence is the largest unit of grammar (cf. fn. 19). Thus, it cannot, by definition, have any arguments (as these arguments would have to be grammatically marked with respect to sentence, which would mean that sentence could not be the largest unit of grammar). If it cannot have arguments, it cannot be a predicate.

The situation, then, looks as follows. We have two sets of rules: (a) generates linguistic predicates from adjectives and linguistic arguments, and (b) generates linguistic arguments from verb and flexible stems (resulting in infinitives and gerunds), argument clauses and linguistic arguments. Below are four examples (as the rules are to some extent language-specific, the list may be incomplete)¹⁸:

¹⁸ / means “or”; \Rightarrow indicates generation transfer; {...} specifies a necessary but non-transferred condition for the rule to hold; [...] designates an optional but transferred condition; T_s is a stem of a word of type T in the given language; T_n is an element of type T on derivational level n, where n is the natural number which corresponds to the

- (a.1) {LA+} COP+ADJ/LA \Rightarrow LP₂ (*He is rich/a sportsman.*)
 (a.2) {LA+} ADJ/LA \Rightarrow LP₂ (*On bogaty/sportsmen.* ‘He is rich/a sportsman’ in Russian)
 (b.1) INF/GER = LA₂ (*to see / seeing*)
 (b.2) {LP+}[LA+]AC \Rightarrow LA₃ (*He knows [the man] who was ill*)

As LA, LP, ADJ, INF and GER can take complements and/or adjuncts, the rules include only the necessary components for secondary predicates and secondary and tertiary arguments. The full complexity of syntax can be accounted for by XPs and predicate phrases, which are the projections of LAs and LPs, respectively (cf. fn. 10 and section 2.3.1). (a.2) is a variant of (a.1) for the languages that lack copulas or can omit them in predicative constructions. Russian is different from English in that infinitives cannot be converted to secondary predicates. **Zhitj umeretj* (literally, ‘To live to die’) and **On umeretj* (‘He to die’) are impossible in Russian, and would have to be rephrased (e.g., *Zhitj znachit umeretj* – ‘To live means to die’) in order to be grammatical. However, Russian gerunds comply with rule (a.2). Thus, *Eto chtenie* (lit., ‘This reading’ ~ This is reading) is acceptable. In addition to (a.1), a few English adverbs can be used predicatively on their own in COP+ADV constructions (e.g. *back*, as in *He is back*). As there are very few such adverbs, all COP+ADV constructions that can end sentences should be viewed as idioms. Below are some examples that illustrate rules (a)–(b):

- (55) [A₁ *John*][P₂ is [A₁ *a painter*]].
 (56) [A₁ *The winner*][P₂ is [A₁ *him*]].
 (57) [A₁ *The winner*][P₂ is [A₁ *John*]].
 (58) [A₂ *Seeing*][P₂ is [A₂ *believing*]].
 (59) [A₁ *This*][P₂ is [A₂ *good thinking*]].
 (60) [A₂ *Running*][P₂ is *healthy*].
 (61) [A₂ *To run*][P₂ is *healthy*].
 (62) [A₁ *He*][P₂ is [A₂ *to win*]].
 (63) [A₂ *To live*][P₂ is [A₂ *to die*]].
 (64) [A₁ *It*][P₂ is *clear*][A₃ [AC₁ *that* [A₁ *John*][P₂ is *ill*]]].
 (65) [A₁ *He*][P₂ is [A₃ [A₁ *the man*][AC₁ *who* [P₂ is *ill*]]]].
 (66) [A₁ *I*][P₁ *asked*][A₃ [A₁ *him*][AC₂ *who* [P₂ is [A₃ [A₁ *the man*][AC₁ *who* [P₂ was *ill*]]]]]].

element's derivational level (primary, secondary, tertiary, etc.). COP, GER and INF can be given cross-linguistically universal structural-functional definitions as well. COP is a special kind of verb consistently used in predicative constructions, translatable as ‘be’. GER and INF are productive nominalizations of V/F stems, translatable as ‘V_s/F_s+ing’ and ‘to+V_s/F_s’, respectively.

2.3.1. Adjectives, adverbs and adpositions in the linguistic predicate/argument structure

Ns are prototypical LAs, Vs are prototypical LAs, and F can be both depending on the marking, but there are still two major word types that need to be accounted for – namely, adjectives and adverbs. As adjectives and adverbs are usually adjuncts of N, V, F or ADJ, their LP/A status is determined by their lexical heads. In effect, adjectives and adverbs can become parts of XPs and predicate phrases. As pointed out in sections 2.2 and 2.3, VPs typically include LAs, i.e. VP is unsuitable for the category of predicate phrase if we are to maintain the LP/A distinction. Predicate phrases include phrasal verbs (e.g., *carry out*), numerous idioms (e.g., *kick the bucket*), V/F constructions modified by adverbs, as in (67)–(68), adjectives and infinitives modified by adverbs in copular constructions ((69)–(70)), and gerunds, nouns and flexibles modified by adjectives (that can, in turn, be modified by adverbs) in copular constructions ((71)–(72)). Observe also that modifiers (adjuncts) do not have to be adjacent to their modified heads (73):

- (67) *He* [P *talked interestingly*].
- (68) *They* [P *reluctantly agreed*] *to close the case*.
- (69) *This* [P *is very interesting*].
- (70) *The patient* [P *is soon to die*].
- (71) *This* [P *was almost perfect skiing*].
- (72) *This* [P *was a very interesting potato/talk*].
- (73) *I* [P *know*] *this place* [P *well*].

In (73), *know* (*x*) *well* should be analyzed as one predicate phrase not two predicates. XPs can contain adjectives and adverbs as adjuncts of N/F/GER and ADJ, respectively (cf. (74)). Infinitives and gerunds can have adverbs as adjuncts. In addition, as verbs and flexibles can take XP complements, infinitives and gerunds can have XP complements as well (cf. (75)–(80)):

- (74) [A₁ *a nearly perfect evening/walk/skiing*]
- (75) *John wanted* [A₂ *to close* [A₁ *the unpleasant case*] *quickly*].
- (76) *Our goal is* [A₂ *to buy* [A₁ *a BMW*] *quickly*].
- (77) [A₂ *To close* [A₁ *the unpleasant case*] *quickly*] *is our goal*.
- (78) [A₂ *Buying* [A₁ *a BMW*] *quickly*] *is our goal*.
- (79) *You'd better start* [A₂ *buying* [A₁ *a BMW*] *quickly*].
- (80) *I want* [A₂ [A₁ *you*] *to buy* [A₁ *a BMW*] *quickly*].

In (74), the adjective *perfect* and the adverb *nearly* are incorporated into the LA (and XP) *a nearly perfect evening/walk/skiing*. Under the NP analysis, the LA would be the projection of the N/F/GER *evening/walk/skiing*, under the DP analysis it would be the projection of the article *a*. The fact that the LA is neutral with respect to the licensing head justifies the use of XP as a theory-

neutral term for LA projections. In (75) and (77), the XP *the unpleasant case* is the complement of the verb *close*, the adjective *unpleasant* is the adjunct of the noun *case*, and the adverb *quickly* is the adjunct of the verb *close*.

Semantically, adjectives and adverbs are predicative, i.e. predicating a quality or circumstance. The reason why they are perceived as semantically predicative is probably because the main function of adjectives and adverbs is to modify N and non-N, respectively, and modification presupposes arguments (the modifier and the modified), thus being intrinsically predicative. However, as adjectives have both predicative (*the sky [is blue]*) and argumental (*[the blue sky]*) use, they could be viewed as a category that lies functionally in between LA and LP. In addition, adjectives have a distinct metonymic use of representing objects by their salient properties, as in *the good, the bad and the ugly*. In this (and only in this) use, standalone adjectives accept LA markers (both DET and POSS) in English. As LA markers mark LAs by definition, adjectives in the metonymic use are LAs by definition. However, this is a clearly marginal use, which does not interfere with definition (4). Moreover, the metonymic use of adjectives is sharply distinct from the others both grammatically (due to LA markers) and functionally, as adjectives so used do not stand for properties but for the objects that have these properties. This is supported by the fact that adjectives' supposed property to accept LA markers does not depend on the adjective but on the object it refers to. For example, a woman but not an evening could be referenced as '*the beautiful*', although *a/the beautiful evening* is a common expression.

In English, predicative adjectives (like *good* in *x is good*) are encoded differently from verbs, i.e. they do not accept LP marking. Along with many other languages, English has a non-verbal encoding of predicative adjectives. However, in a number of languages (e.g. North-East Ambae, Korean, Lao, Qiang and Semelai), predicative adjectives are encoded exclusively by LP markers (Hajek 2004). There are also many languages of mixed type, with both verbal and non-verbal encoding options available for predicative adjectives (Stassen 2008). Languages with exclusively verbal encoding of predicative adjectives are concentrated in the Asia-Pacific region (Hajek 2004; Stassen 2008). In Lao, for example, TAM marking on adjectives seems to encompass the whole TAM paradigm of the language. Obviously, this poses the problem of distinguishing adjectives from verbs in Lao. As a solution, Enfield (2004) has suggested that Lao adjectives are a subclass of verbs. In all other languages, TAM marking on adjectives is restricted to proper subsets of the TAM paradigms of the languages (Hajek 2004). Crucially, whenever adjectives are marked by LP or LA markers, they are LPs or LAs, by definition.

Adpositions are LA-LP relation modifiers. Typically, they modify relations between XPs and clause predicate. Cf. the following examples:

- (81) *He walked [A to the house].*
- (82) *He walked [A into the house].*
- (83) *He walked [A inside the house].*

- (84) *He walked [A behind the house].*

As LA-LP relation modifiers, adpositions should be analyzed as peripheral constituents of either LAs or LPs. The prepositional phrases (81)-(84) are examples of the former and phrasal verbs are examples of the latter:

- (85) *The plan [P was carried out].*
 (86) *He [P logged in].*

2.3.2. The linguistic predicate/argument conversion system. Noun-argument and verb-predicate correspondences

The LP/A conversion system works as follows. Rule (b.1) converts primary LP stems to secondary LAs. Rules (a) convert primary, tertiary and (partly) also secondary LAs to secondary LPs. If primary and secondary LPs are incorporated in ACs they can be converted to tertiary LAs by rule (b.2):

- (87) *We know [A₃ [AC₁ that John [P₁ sleeps]]].*
 (88) *We know [A₃ [AC₁ that John [P₂ is ill]]].*

Depicted graphically, the LP/A conversion system is as follows (cf. fn. 18):

- (89) $(A_1, A_{2*}, A_3) \text{ — (a) } \rightarrow P_2$
 $P_{1s} \text{ — (b.1) } \rightarrow A_2$
 $(P_1, P_2) \text{ — (b.2) } \rightarrow A_3$

Arrows indicate possible conversions and on the arrows are written the rules which stipulate the conversions. However, Russian infinitives cannot be converted to secondary predicates (see section 2.3). As infinitives are of type A₂, the restriction that not all elements of this type can be converted to secondary predicates is signalled by A_{2*} in (89). Three inferences can be made from (89). First, all LPs can be converted to LAs but not all LAs can be converted to LPs. Second, as (a.1) and (a.2) are essentially a single rule (a) that is slightly different for different languages, tenses and moods (cf. section 1.3), there are more rules for converting predicates into arguments than vice versa. Third, there are three levels of arguments but only two levels of predicates. Together with the inherent asymmetry of P/A structure (essentially, arguments being elements and predicates being operations with these elements – see point 1 in section 3.3), the inferences suggest that LAs may be evolutionarily more fundamental than LPs, a point taken up in chapter 3.

If we are concerned only with the core components of LAs and LPs (i.e. if we ignore adjuncts and adpositions), the lexicon we are interested in reduces to {N, V, F, ADJ, LAx, LPx}. N, F bearing LAx and ADJ bearing LAx are

primary LAs. V, F bearing LPx and ADJ bearing LPx are primary LPs. Now we can have an overview of what count as LAs and LPs:

- (90) LA = {N, F+LAX, ADJ+LAX, (b), PJ(LA)}
 (91) LP = {V, F+LPx, ADJ+LPx, (a), PJ(LP)}

LAs are nouns, flexibles bearing LA markers, adjectives bearing LA markers (cf. section 2.3.1), and the products of rules (b). In addition, LAs include projections of LAs (i.e. XPs). LPs are verbs, flexibles bearing LP markers, adjectives bearing LP markers, and the products of rules (a). LPs also include projections of LPs (i.e. predicate phrases as outlined in section 2.3.1). Finally, although all languages have LAs and LPs, it does not entail that all languages have all the eight elements from the sets in (90) and (91). For example, there is no consensus whether all languages have nouns and verbs (Anderson 2004; Bach 2004; Laudanna & Voghera 2002), and the claims of adjectives not being universal (Beck 2002; Hengeveld 1992; Junker 2003; McCawley 1992; Rijkhoff 2000; Sasse 1993b) are more frequent than those of the contrary (Baker 2003; Dixon 2004). All this has, of course, no affect on the universality of (90)–(91).

I conclude that noun-argument and verb-predicate correspondences, though complicated by grammar, can be established for NL (cf. (90)–(91)). This is a crucial step towards establishing complete correspondence between first order predicate logic and NL notations of P/A structure. It is certainly correct to argue that first order predicate logic is an artificial system which NL has neither need nor pressure to correspond to. However, it is likely that more potential correspondence problems can, like the one posed by Hurford (2003b, 2003c), be overcome by having multiple levels of predicates and arguments in natural language.

2.4. Summary

It has long been noticed that, in natural language, the same kind of term can fill both the argument and the predicate slot, whereas in first and second order predicate logic the same same terms cannot be arguments and predicates. This is what Hurford (2003b, 2003c) called the ‘Aristotle problem’; Hurford concluded that it remains to provide an explanation for the typical structure of modern languages, organized around the noun/verb dichotomy. I suggest that the ‘Aristotle problem’ has remained unsolved because natural language’s predicate/argument structure has been assumed to be isomorphous to that of first order predicate logic. I argue that natural language has its own intrinsic predicate/argument structure, which is more complex than that of first and second order predicate logics – for instance, it has up to three levels of arguments and predicates, and a rule system that allows (among other things) to convert predicates into arguments and vice versa. Besides making things more complex, this property of the rules – linguistic predicates being convertible into

linguistic arguments and vice versa – provides also a key for the correspondence between natural language's and predicate logic predicate/argument structures. Essentially, the solution to the 'Aristotle problem' lies in the following three assumptions:

- (c) In natural language, predicate/argument categorizations can be overridden by rules (a)–(b).
- (d) From (c) it follows that linguistic arguments and predicates of **different levels** can correspond to predicate logic arguments and predicates.
- (e) The generic linguistic predicate/argument system as reproduced below (observe that there may be more rules analogical to (a)–(b) in the world's languages; the notation is explained in the Abbreviations and in fn. 18):

$$(92) \quad LA = \{N, F+LAX, ADJ+LAX, (b), PJ(LA)\}$$

$$(93) \quad LP = \{V, F+LPX, ADJ+LPX, (a), PJ(LP)\}$$

$$(a.1) \{LA+\}COP+ADJ/LA \Rightarrow LP_2 \quad (He \text{ is rich/a sportsman.})$$

$$(a.2) \{LA+\}ADJ/LA \Rightarrow LP_2 \quad (On \text{ bogatyι/sportsmen. 'He is rich/a sportsman' in Russian})$$

$$(b.1) INF/GER = LA_2 \quad (to \text{ see / seeing})$$

$$(b.2) \{LP+\}[LA+]AC \Rightarrow LA_3 \quad (He \text{ knows [the man] who was ill})$$

An important property of this system is flexibility, as evidenced by (92), (93), (a)–(b), and by the fact that a language does not require all the categories specified in (7), (8), (92)–(b) in order to have a working LP/A system with conversion possibilities. As the backbone of natural language syntax, the generic LP/A system is expected to be robust. By (7), (8), (92)–(b), a language without nouns, verbs, adjectives, determiners, copulas and voice oppositions could, in principle, have an LP/A system as powerful and flexible as that of languages with all these categories. Observe also that rules (a)–(b) are not tied to the underspecification hypothesis (3) which posits F. A refutation of the underspecification hypothesis (presumably by neurolinguistic experiments) would not affect rules (a)–(b). The only difference would be that F would have to be removed from the sets (92)–(93).

The gist of the 'Aristotle problem' is that nouns and verbs do not correspond to first and second order predicate logic arguments and predicates. As I have shown, nouns and verbs correspond to linguistic arguments and linguistic predicates (cf. (92)–(93)). Technically, as natural language does not have to correspond to an artificial system (e.g. predicate logic), the final part of establishing the correspondence reduces to representing linguistic arguments and predicates in first order predicate logic by premises (c)–(e).

3. THE EVOLUTION OF THE NOUN/VERB AND LINGUISTIC PREDICATE/ARGUMENT STRUCTURES[‡]

3.1. Introduction

Unless we make the absurd assumption that the whole complexity of natural language was there from the beginning, complex linguistic structures must have evolved from simpler ones. This chapter deals with the problem of the course of the evolution of the noun/verb and linguistic predicate/argument structures. It is not immediately obvious how something could be said about this issue, as these structures (and especially the second one) are among the most fundamental in syntax. However, as it is established that the evolution of language must have passed a pre-syntactic stage (the latter has received rather detailed treatments in, e.g., Bickerton 1990; Bowie 2008; Dessalles 2008; Jackendoff 1999; Johansson 2006), it is not unreasonable to suspect that something could be conjectured about the evolution of the noun/verb and linguistic predicate/argument structures as well. At the very least, it is possible to evaluate the evidence for and against specific hypotheses.

3.2. The evolution of the noun/verb and linguistic predicate/argument structures

It has been claimed that the sentence / noun phrase (S/NP) or, in a reformulated version, the S/XP distinction is universal in the world's languages (Carstairs-McCarthy 1999; Hurford 2007). I propose that the universality of S/XP derives from the universality of LP/A. The majority of the definitions of S are centered around LP (grammatical predicate)¹⁹. Thus there can be no S before there is LP, and there can be no LP before there is the LP/A distinction (cf. section 3.3, points 1 and 6). Therefore the LP/A distinction is more fundamental than the S/XP distinction. S/XP follows from LP/A, as LP provides the basis for S, and LA for XP. This answers the question of the origins of S and XP. The separate

[‡] This chapter is based on Luuk (in press-a).

¹⁹ Importantly, S (sentence) is not equivalent to LP (cf. section 2.3.2). For example, *is a man* in *John is a man* is an LP but not S. Similarly, *John runs* is an S but not LP (as it is comprised of an LA and LP). In contemporary linguistics, sentence is seldom defined, except within various X-bar theories, where it is usually regarded as the maximal projection of V (Gazdar, Klein, Pullum & Sag 1985; Kayne 1981; Marantz 1980) or the maximal projection of INFL (Chomsky 1981, 1995), in the latter case, it is commonly analyzed as IP. However, these are not the only definitions of sentence that have been circulating. Both Bloomfield and Meillet have defined sentence as “a form which is part of no other form” (Graffi 2001: 1843) – thus, sentence is the largest unit of grammar.

question of why do all languages provide parallel XP and S structures to express essentially the same proposition, as in [S: *the flower is delicate*][XP: *the delicacy of the flower*] and [S: *the rat runs*][XP: *the running of the rat*], assumes the universality of construction correspondences for this semantic “equivalence”, e.g.

$$(94) \quad LA_1 + LP_2 = LA_1 + POSS + LA_1'$$

as in [LA_1 *the flower*]+[LP_2 *is delicate*] = [LA_1 *the flower*]+[POSS ‘s’]+[LA_1' *delicacy*], and possibly also some conversion rules (e.g. from LP_2 to LA_1' and/or vice versa)²⁰. An answer to the third question of why such correspondences are universally found in the world's languages, is that they provide a structural foundation for semantic diversity. In the S/XP case, the semantic diversity, as argued by Hurford (2007), is the topic/comment structure (Hurford aligns XP with topic and S with comment).

I have shown that S/XP can be reduced to LP/A. What could be said about the origins of the LP/A structure? Obviously, it is impossible to offer a detailed overview of its emergence. However, if Hurford (2003b, 2003c) is correct about the cognitive/perceptual P/A structure being common among mammals, we have to assume that, in certain species at least, the cognitive/perceptual P/A structure was mapped onto conceptual structure, resulting in a corresponding conceptual P/A structure. This conceptual P/A structure was, in turn, projected onto language as the LP/A structure in the earlier stages of language evolution (cf. section 2.2). At this point, not much more can be conjectured on this issue.

In mathematics, function is a proper subset of relation, which in turn can be expressed as a predicate (Hummel 2000). The pervading universality of P/A structure strongly suggests that the structure is not an arbitrary but an evolutionary result with neural implications, which would qualify it as a hardwired fossil – distinct from language fossils that do not have to be hardwired in the brain (cf. section 2.1). It has been claimed that ventral and dorsal pathways in the brain handle the predicate and argument processing, respectively (Hurford 2001, 2003b, 2003c). Importantly, the claim is about the P/A structure in perception and cognition, not about the P/A structures in logic and NL (Hurford 2003c). If the cognitive/perceptual P/A structure is a hardwired fossil, it is not only the case that the linguistic and the logical P/A structures subsume the cognitive/perceptual P/A structure, but also that the cognitive/perceptual P/A structure may have implications for conceptualization (e.g., in the form of a corresponding conceptual P/A structure). According to Hurford (2006a, 2003b, 2006b), even a number of non-human species have the

²⁰ In English, a conversion rule stipulates that LA_1' must be derived from LP_2 in (94). Thus, *the man is interesting* = *the interestingness of the man*, but *the man is interesting* ≠ *the interest of the man*, because *interest* is not derived from *interesting* (obviously, it is vice versa). Similarly, *time passes* = *the passage/passing of time*, but *time passes* ≠ *the pass of time*.

cognitive/perceptual P/A structure. The main difference between the logical/linguistic and the cognitive/perceptual P/A structures is that the latter is more elementary. Some arguments of the former are analyzed as predicates taking an argument in the latter, e.g.

- (95) CAME(man)
 (96) MAN(x) & CAME(x)

Capital letters denote predicates, small letters denote arguments (x is a deictic argument variable), & is a connective (conjunction), (95) is the linguistic and (96) the cognitive/perceptual interpretation. (96) hints at the possibility of a ‘protolanguage’ without grammar and the LP/A distinction, where words would have the logical form of PREDICATE(x) and could be concatenated regardless of their order. This possibility is further supported by the fact that the LP/A distinction is made by grammar. Thus, necessarily, there was no such distinction before grammar. The fact that all human languages have grammar and the LP/A distinction could point to an equivalence relation between grammar and the LP/A distinction. However, it is not known whether the first grammar rule stipulated the LP/A structure (see below). Thus the equivalence relation between NL grammar and the LP/A structure is merely a possibility.

Word order is the simplest grammatical device²¹. Irrespective of whether one takes the analytic or synthetic view on the emergence of syntax, the first rule sufficient for a primitive NL grammar was probably a word order constraint (Heine & Kuteva 2002; Johansson 2006). The synthetic view assumes that the original words of protolanguage were strung together to make the phrases and sentences of full language. The analytic view assumes that the original words of protolanguage were dissected into parts which came to express the atomic meanings of full language (for more details on this distinction, see Hurford 2000a). Like Hurford (2000b), Bickerton (2000) and Jackendoff (1999), I prefer the synthetic view. However, my definition of ‘word’ in section 1.2 is universal. Thus, when Wray, in arguing for the analytic view, says that “if, besides *tebima* meaning *give that to her*, *kumapi* meant *share this with her*, then it might be concluded that *ma* had the meaning *female person + beneficiary*” (Wray 2000: 297), *ma* is specified as a word in the sense it is defined in section 1.2. This in turn stipulates a word order constraint as the first grammar rule. The first word order constraint **automatically** results in two

²¹ Newmeyer (p.c.) suggests pure juxtaposition regardless of order to be even simpler. While this is true, it is dubious whether this qualifies as a ‘grammatical device’. If it did, we would have to face the consequence that protolanguage had grammar (as it is generally assumed that protolanguage allowed for such concatenations – Jackendoff 1999, 2002; Bickerton 1990, 1998, 2000; Johansson 2006). However, the dominant opinion seems to be that protolanguage lacks grammar rather by definition (cf. Tallerman 2007; Wray 2000; Jackendoff 1999; Bickerton 1990). Thus, it is terminologically preferable to separate protolanguage from grammar/syntax. Hence, word order as the simplest grammatical device.

grammatically distinct categories for words²². Theoretically, a word order rule would suffice for the LP/A distinction. However, we do not know whether the first grammatical categories were word types (like *man go*) or just semantic roles (like *man forest* interpreted as ‘a man go to the forest’). Distinct word types would be a more plausible suggestion owing to a greater transparency of interpretation, a greater potential for combinations, and the fact that movement is a perceptually salient property (which is associated with LPs – see point 4 in section 3.3). I conclude that LA and LP would be the most obvious but not the only candidates for the first word categories. With different argumentation and terminology (nouns and verbs instead of linguistic arguments and predicates), other authors have arrived at a similar conclusion (Heine & Kuteva 2002; Hurford 2003a; Newmeyer 2003).

Suppose that the initial grammatical distinction between LAs and LPs was made solely by word order. Such language would have been much cruder than the modern variety. Still, it would have allowed to express **events** – objects/properties caught in actions/changes. In NL, the LP/A structure seems to be a precondition of expressing events in the above defined sense.

3.3. Eleven arguments for the evolutionary primacy of LA over LP

Below I present eleven arguments suggesting that LAs are evolutionarily more fundamental than LPs. But first, we must eliminate a possible source of confusion. It is not a contradiction that LAs could have evolved before LPs. LAs and LPs are complementary but that does not entail interdependence. For instance, dual and plural are complementary without being interdependent (dual implies plural but not vice versa). LA does not imply the LP/A distinction. In the absence of the LP/A distinction, only one thing had to be different: as LA was the only word type, there was no need for LA marking. There is no way of knowing what were the first words but the following arguments should make it clear that they most likely approximated nouns. The eleven arguments for the evolutionary primacy of LA over LP are as follows:

1. LPs presuppose LAs they act upon. A predicate applies to a variable, whose value is provided beforehand (Hurford 2003b). This is the reason why a language without LAs is almost inconceivable, whereas a language without LPs seems accessible enough. Nouns are prototypical LAs and verbs are prototypical LPs. One can utter *ship Amsterdam tomorrow* and be understood that “a ship will arrive in or depart to Amsterdam tomorrow” but a nounless

²² Unless we are dealing with a phonological (e.g., that the word beginning with a vowel comes first) or lexical constraint (e.g., that the word standing for the concept ‘tree’ comes first). I am not aware that there were any such constraints in the world's languages but I cannot preclude this possibility.

English construction expressing the same, though possible, is not likely to be univocally understood. J. L. Borges has explored the possibility of a nounless language in one of his short stories (Borges 1964). The sample text he produces relies heavily on imagination and adjectives, whereas a verbless language can do with nouns alone. Asymmetry is inherent to P/A (and hence, to the LP/A – Luuk) structure (Hurford 2003c). In section 2.3.2, we found that all LPs can be converted to LAs but not vice versa, that there are three levels of arguments but only two levels of predicates, and that there are more rules for converting predicates into arguments than vice versa. Budd (2006) has suggested that in complex systems with asymmetrical dependencies, the functionally necessary core component must have evolved first in relation to the ‘unnecessary’ ones. Among words, LAs are the prime candidates for the functionally necessary core component. See also point 6 below.

2. Children’s early productive vocabularies are dominated by nouns, and infant comprehension of object names appears earlier than comprehension of relational terms (Fisher 2002; Gentner & Boroditsky 2001; Gleitman 1993; Waxman & Lidz 2006). Although it has been argued that early noun dominance is not universal cross-linguistically, the evidence for this is still weaker than the evidence against it (Gentner & Boroditsky 2001; Gopnik 2000).

3. A virtual experiment (Steels, Kaplan, McIntyre & Looveren 2002) has identified a condition favoring nouns (i.e. LAs) for the first words – the condition that agents must have parallel non-verbal ways to achieve goals of interactions (e.g. pointing). Actions/changes are difficult to point to – other than, perhaps, by imitating or carrying them out. Accordingly, as compared to the first LAs, the first LPs would have been more elaborate in gestural modality. This in itself does not rule out the possibility that LPs came first, as it has, for instance, been proposed that language began as a “mixture of isolated grunts and gestures” (Bickerton 2003: 81). However, the fact that language opted for vocal not gestural modality still favors LAs over LPs for the first words.

4. LAs appeal to geometrical and LPs appeal to kinaesthetic properties of images. As Pylyshyn has argued, the intrinsic properties of images are geometrical rather than dynamic, both because the spatial intuitions are among the most entrenched, and because there is evidence that geometrical and optical-geometrical constraints are built into the early-vision system. While we can easily imagine the laws of physics being violated, it seems nearly impossible to imagine the axioms of geometry or geometrical optics being violated (Pylyshyn 2002). Prototypically, nouns are associated more with geometrical and verbs with kinaesthetic properties. A quick look into Webster’s (1988) English dictionary supports this intuition. For example, of the first 10 nouns starting with the letter K (*Kaiser, kale, kaleidoscope, kalology, kalong, kampong, kangaroo, kaolin, kapok, karma*) only two (*kalology* and *karma*) do not appeal to geometrical properties but none appeal to kinaesthetic properties. After

excluding the flexibles (which are expected to evoke both geometrical and kinaesthetic properties), only five verbs starting with the letter K remain (*kindle, knap, knead, kneel, know*). If we add the first five verbs starting with the letter L (*lacerate, lambaste, lave, learn, legislate*), we see that only three (*know, learn, legislate*) do not appeal to kinaesthetic properties. Perhaps not incidentally, these are also the verbs that do not appeal to geometrical properties. It would seem that kinaesthetic properties presuppose geometrical properties. Indeed, it is impossible to imagine movement without or outside space-time. This asymmetric dependency – the kinaesthetic properties of images depending on the geometric ones but not vice versa – together with the tendency of nouns to evoke geometric properties and the tendency of verbs to evoke kinaesthetic as well as geometric properties, suggests that nouns are cognitively more fundamental than verbs and verbs are cognitively more complex than nouns. This, in turn, suggests that nouns (LAs) may evolutionarily predate verbs (LPs).

5. Words of different grammatical category can be selectively harmed. It has been established that while some patients with language disorders show a worse performance with verbs than nouns, others show the opposite pattern. Noun superiority is frequently found in association with Wernicke's and Broca's aphasia and verb superiority with anomia (Mondini, Luzzatti, Zonca, Pistarini & Semenza 2004). Selective impairment of verbs is more frequent than selective impairment of nouns (Arevalo et al. 2007; Luzzatti et al. 2002). There are two mutually nonexclusive explanations for this: 1. Extensive damage to the left hemisphere language areas induces the emergence of right hemisphere lexical abilities that are limited to high frequency concrete nouns (Crepaldi et al. 2006). 2. Selective impairment of verbs is a function of argument structure complexity that is regularly associated with verbs. It has been shown that the impairment is greater with 3-place than 2-place verbs, and 2-place than 1-place verbs (Kim & Thompson 2000; Luzzatti, Aggujaro & Crepaldi 2006). Moreover, production of argumental nouns like the Italian *passeggiata* 'a walk', *risata* 'laughter', *pugnalata* 'a stab', etc. is impaired at an equal level with production of argumental verbs (Collina, Marangolo & Tabossi 2001)²³. I point out that all these findings are consistent with two hypotheses. (1) The N/V double dissociation in aphasia is an effect of the conceptual P/A double dissociation in the brain (the circumstance that argumental nouns are impaired at an equal level with argumental verbs refers to the conceptual P/A rather than the LP/A double dissociation). (2) The N/V double dissociation is an effect of argument structure complexity. It is difficult to disentangle (1) from (2), as they have many correlated features. I conclude that the fact that the processing of

²³ Observe that many putative argumental "nouns" or "verbs" are, in fact, flexibles (e.g. English *walk, stab, run* etc., Italian *pianto, urlo, passeggiata, pugnalata* etc.). Cf. *pianto* 'crying' – *piangere* 'to cry', *urlo* 'a yell' – *urlare* 'to yell', *passeggiata* 'a walk' – *passeggiare* 'to walk', *pugnalata* 'a stab' – *pugnalare* 'to stab'.

LPs is more specialized and/or resource demanding than the processing of LAs suggests that the latter may be evolutionarily more fundamental²⁴.

6. In all natural languages, LP is the cornerstone of syntax. Cf. Ross (1972: 325): “nouns are more inert, syntactically, than adjectives and adjectives than verbs”. NL syntax is based on the principle that LPs take arguments that are differentiated by analytic (adpositions, word order) and/or synthetic (morphological) case markers. Thus, there seems to be an equivalence relation between NL syntax and LP (i.e., if a system has LPs, it has NL syntax; and if it has NL syntax, it has LPs). In addition, the utility of LAs without syntax is obvious but the utility of LPs without syntax is dubious (although imperatives can be syntactically independent, as they are optimized for producing and parsing speed – see section 2.1). The hypothesis that LP is equivalent to syntax, together with the axiom that there was no syntax in the beginning (Jackendoff & Pinker 2005), favors LA over LP for the first words. Bickerton remarks that symbol and syntactic structure can be dissociated – the latter without the former is useless, whereas the former is useful per se. He further argues that this logico-pragmatical dissociation has a historical counterpart: “a variety of factors /---/ suggest that, in the evolution of our species, symbolism may have preceded syntax by as much as two million years” (Bickerton 2003: 81). It is a possibility, then, that the historical dissociation between symbol and syntax is distantly reflected in NL structure in the form of the LP/A distinction.

7. In analyzing the syntactic functions of major parts of speech, it has been frequently suggested that the function of nouns (including pronouns and proper names) is the most basic one. For a simplified language model, it has been found that noun is the only constituent class that all sentences have in common at the highest level of constituent-structure (Lyons 2004). Referring to Jespersen, Lésniewski and Ajdukiewicz, Lyons conveys that nouns are “categories of the first degree” and that “all other parts of speech are derived, complex categories. Categories of the second degree combine with categories of the first degree (according to the principles of well-formedness /---/) to form sentences /---/” (Lyons 2004: 219–220). In analyzing semantic classes (situation, event, place, time etc.) Anward writes that “while the semantic class of person/thing seems lexicalizable by nouns, other semantic classes can be lexicalized in several ways” (Anward 2001: 730).

8. Nichols has formulated two important principles of historical morphology: 1. Headward migration: “If any adposition or piece of affixal morphology moves,

²⁴ It has been also found that nouns are more readily recognized than verbs in homographic priming tests (Laudanna & Voghera 2002). In homographic priming tests, the test word is primed with a word the stem of which is phonologically and ortographically identical but semantically and grammatically different from the test word, as with the Italian *colpire* ‘to hit’ and *colpa* ‘fault’.

it will go from dependent to the head of the constituent, not vice versa” (Nichols 1986: 86). 2. Reduction: the original dependents get cliticized and eventually become morphological markers of their head. Principle 1 suggests that the initial marking is more likely to appear on dependent. Together, the principles suggest a morphological migration pattern from dependent to head (e.g., from N to V). The fact that, cross-linguistically, verbal morphology appears to be richer than nominal morphology, is consistent with this. Although the evidence for it circumstantial, it is not unreasonable to suspect that the morphology appeared on older elements first. As morphology obscures lexical items’ form and meaning, the latter have to be sufficiently conventionalized before any morphology can attach to them. It is plausible that older elements are more conventionalized than younger ones. Second, statistically, the longer an element has been around, the more chances it has had to attract morphology. Thus, the default assumption would be that the element that became a dependent is older than the element that became its head. An analysis of constituent types and their head-dependent relations confirms this. Cf. the following table (based on Helmbrecht 2001: 1425):

Table 1. Constituent types and head-dependent relations

Constituent	Head	Dependent
1. NP	N	ADJ
2.	ADP	N
3. Clause	V	N
4.	AUX	V

From Heine and Kuteva (2002, 2007) it follows that, in three pairs (2, 3, 4), the dependent element is older than the head element. In one pair (1), the situation is the other way around. Thus, the evidence for the dependent element being older than the head element is stronger than the evidence for the contrary. Combined with the considerations put forth by Heine and Kuteva (2002, 2007), this adds up to a modest evidence that, in pair 3, N is older than V.

9. The first stage of spontaneous adult second language acquisition features noun-based utterance organization and lacks verb/argument structure (Klein & Perdue 1997; Perdue 1996). It has also been established that the priority of setting a (static) time reference for a situation viewed as a whole before giving it an aspectual (dynamic) perspective is characteristic of early untutored second language varieties and adult home signs (Benazzo 2006).

10. There are more nouns than verbs, and more productive noun than verb derivation in the world's languages (Gentner 1981; Gentner & Boroditsky 2001; cf. Jacobsen, 1979). This also suggests that nouns are more fundamental, i.e. possibly predate verbs.

11. In the world's languages, there is at least one example of LP marking on an LA marker (TAM on DET in Chamicuro – Nordlinger & Sadler 2004). I know of no examples of LA marking (DET or POSS) on an LP marker (TAM or voice). This asymmetry – LA markers being more independent and having a more substantial role in the lexicon than LP markers – begs an explanation. A plausible explanation is that LA markers are generally older than LP markers²⁵. Lexical items' form and meaning have to be sufficiently conventionalized before they can be modified by markers (cf. point 8). The conjecture that LA markers are generally older than LP markers lends some additional support to the hypothesis that LAs predate LPs.

Assuming that the initial function of language was to give orders, it could be argued that the first words might have approximated imperatives, i.e. verbs or flexibles (e.g. *Run!*, *Help!* or *Catch!*). I have already presented five rather direct arguments against this (see points 1, 3, 4, 6, 7) but let me add two more. First, as words are always interpreted in the context, orders can be also given by other word types, e.g., nouns (*Tiger!*, *Tree!*) or adverbs (*Up!*, *Quickly!*). Second, a language is not necessary for giving orders. In many cases, the desired reaction can be elicited simply by drawing listeners' attention to a potentially dangerous situation or object. Many species have developed natural alarm signal systems the effect of which approximates the intended effect of orders like *Run!*, *Flee!* or even *Help!*. Vervet monkeys have developed an alarm call system which distinguishes four different predation patterns (snakes, birds of prey, large cats, primates), eliciting a different flee response for each (Hauser 1997). Indeed, it has been even argued that these calls might approximate words like *leopard* and *eagle* (Zuberbuhler, Cheney & Seyfarth 1999). While this seems far-fetched, it goes to show that even non-human animals have calls which are functionally equivalent to fairly specific orders. Returning to humans, an inarticulate cry is usually sufficient to draw their attention to a dangerous situation. The resulting reaction (e.g., running, fleeing or helping) is an instinctive and/or learned behavior which does not require any semantics – a common knowledge of dangerous situations and how they can be neutralized is sufficient.

As for other arguments for the contrary – the primacy of LP over LA – I have found only one, presented by Hengeveld (1992) and Rijkhoff (2002, 2003). This argument, however, relies on a specific assumption, viz., that if a language has a reduced number of nouns, it lacks the category noun. I find this viewpoint inconsistent. If a language has any number of nouns (other than zero, of course), it has the category noun by definition. The lack of a typological trait must be pervasive for it to be declared absent in a particular language (Baker 2003; Croft 2000; Evans & Osada 2005; Itkonen 2000). Indeed, in their recent

²⁵ As a single word order constraint could mark both LA and LP, we cannot assume that LA markers are, as a whole, older than LP markers. Assuming the lexicon of two word types, LA and LP, a word order constraint could be sufficient for the LP/A distinction (cf. section 3.2).

papers, Hengeveld and Rijkhoff have revised their views on this issue (Hengeveld & Rijkhoff 2005; Hengeveld et al. 2004).

The claim that N is evolutionarily more fundamental than V is not new. Previously, this has been tentatively suggested within the framework of grammaticalization (Heine & Kuteva 2002, 2007). In reconstructing early language, Heine and Kuteva (2007) propose that at stage 1 there was only one lexical category, namely “nouns” (time-stable, referential units expressing primarily thing-like concepts). The present paper agrees with this, while presenting a number of new arguments, gathered from a variety of domains, to support the claim. To my knowledge, only argument 9 and one point (the possibility of a verbless language) from argument 1 have been suggested before, by Heine and Kuteva (2007). By itself, none of the eleven arguments above is sufficient to establish the primacy of LA over LP, but taken together, the evidence is overwhelming.

As for the referential origins of the LP/A dichotomy, one might speculate that all LPs are ultimately derived from the LAs used or involved in these actions – for instance, *stone* → *to stone (somebody)*, etc. This, however, is merely a speculation as the referential source for the first LPs might have been autonomous. Whichever is the case, if one assumes an equivalence relation between grammar and the LP/A distinction, LP had to evolve before grammar (i.e. in protolanguage) to support that distinction. Nowak and Krakauer (1999) have hypothesized that the N/V distinction reflects the natural way humans perceive reality – specifically, conceptualizing it as a series of **events**: somebody does something, something happens to somebody, etc. – simple constructions that, minimally, should require nouns and verbs (or, as I more conservatively suggest, LAs and LPs) to be expressed in modern language. Thus, the grammar of NL seems to have evolved to reflect the underlying “grammar” of the world we live in. Observe, however, that while it is reasonable to assume that events were conceptualized before the emergence of NL, it is not self-evident that they were conceptualized the way Nowak and Krakauer (1999) propose. In addition, there is the problem of metalanguage: Nowak and Krakauer's formulation “somebody does something, something happens to somebody” already includes, i.e. tacitly assumes N and V.

3.4. Summary

Two inferences can be made from rules (a)–(b) in section 2.3. First, as (a.1) and (a.2) are essentially a single rule (a) that is slightly different for different languages, tenses and moods (cf. section 1.3), there are more rules for converting predicates into arguments than vice versa. Second, there are three levels of arguments but only two levels of predicates. Together with the inherent asymmetry of P/A structure (arguments being elements and predicates being operations with these elements), the inferences suggest that LAs may be evolutionarily more fundamental than LPs. In a slightly modified version (with

N and V instead of LA and LP), this hypothesis has been tentatively proposed within the framework of grammaticalization (Heine & Kuteva 2002, 2007). Section 3.3 analyzes evidence for and against this hypothesis, and reveals a number of new arguments that lend additional support to the hypothesis. The conclusion is that the evidence for the evolutionary primacy of LA over LP is overwhelming, the evidence for the evolutionary primacy of LP over LA is absent, and the evidence for interdependence between LP and LA is feeble.

SUMMARY AND CONCLUSIONS

The present thesis deals with the noun/verb and linguistic predicate/argument structures. The linguistic predicate/argument structure is the universal backbone of natural language syntax. I have managed to find only one citation where the universality of the linguistic predicate/argument structure is seriously contested (Gil 1994: 194); however, Gil's analysis of Riau Indonesian in no way precludes the possibility that the language has linguistic predicates and arguments (see section 1.1). I claim that, in addition to the universality of the LP/A structure, LAs and LPs are universally marked by (a subset of) the following range of grammatical devices: tense-aspect-mood, determiners, possessives, voice and LP/A word order constraints. The circumstance that LAs and LPs are universally marked by (a subset of) these grammatical devices, allows for universal definitions of 'noun', 'verb' and 'flexible' as correlations of LP/A marking and lexical class. The definitions are as follows: N = the property of stems to accept LA markers but not LP markers; V = the property of stems to accept LP markers but not LA markers; F = the property of stems to accept both LA and LP markers. According to the present account, stems bear lexical class, although it would not affect the overall system if they did not. However, in section 1.2 I present some considerations why the hypothesis that functional heads (e.g., TAM or DET) rather than stems bear lexical categories is implausible. In the same section, the word class 'flexible' is established as the parsimonious alternative to zero derivation and homophony. However, 'flexible' (or more generally, underspecification) is the better alternative only insofar as neurolinguistic experiments have not proven the opposite. As present day methods do not allow lexical classes to be identified in the brain, we are left with underspecification as the hypothesis that, differently from zero-derivation and homophony, does not posit hidden structures. Given the functions of the word classes 'noun', 'verb' and 'flexible' – viz., linguistic argument, linguistic predicate, and both –, together with the premise that all languages have linguistic arguments and linguistic predicates, the following system of logically possible language types emerges: *N/V/F*, *N/F*, *V/F*, *N/V* and *F*. After analyzing the evidence for each of these types, I conclude that type *N/V/F* is by far the most common, if not indeed the only one in the world's languages (if the criterion of pervasiveness is applied, stipulating that a language has a linguistic category if it has at least one member of this category). The language types ranking next are (in the order of probability): *F*, *V/F* or *N/V*. Presently, this is the best guess – as mentioned above, it is possible that none of these types exists (if the pervasiveness criterion is applied).

The linguistic predicate/argument structure should not be confused with the logical, the cognitive/perceptual and the conceptual predicate/argument structures. I distinguish and characterize these four predicate/argument structures in sections 2.2, 2.3, 2.4 and 3.2. In sections 2.3–2.4, I show that the 'Aristotle problem' – the circumstance that the same kind of term can fill both the argument and the predicate slot (as *a man* in *A man died* and *Plato is a*

man), resulting in a lack of noun-argument and/or verb-predicate correspondence(s) – arises from confusing the linguistic and the logical predicate/argument structures, or from assuming that the structures are identical or in close resemblance. Section 2.3 argues at length that the linguistic predicate/argument structure is qualitatively different from that of first and second order predicate logics in having up to three levels of predicates and arguments, and sets of rules for generating higher order predicates and arguments as well as for converting predicates to arguments and vice versa. This flexibility of the linguistic predicate/argument system allows, among other things, for noun–LA and verb–LP correspondences, thus reducing the Aristotle problem to the problem of translating the complex LP/A structure into first order predicate logic. Along with more definite answers to the question of which of the five logically possible language types actually exist(ed), the problem of translating the LP/A structure into first order predicate logic is beyond the scope of the present thesis. However, observe that predicate logic is an artificial system which natural language has no a priori need to correspond to – i.e., this problem has more bearing on mathematical logic than linguistics.

Certain asymmetries in the rule system of the linguistic predicate/argument generation and conversion, as well as asymmetries in predicate/argument system in general, suggest that LAs may evolutionarily predate LPs. After connecting the details that can be conjectured on the overall course of the evolution of the LP/A structure, the thesis concludes with the analysis of the hypothesis that LAs may predate LPs. Previously, a similar hypothesis (with nouns and verbs instead of LAs and LPs) has been put forward within the framework of grammaticalization (Heine & Kuteva 2002, 2007). With a number of new arguments from a variety of domains, my analysis confirms that the evidence for LAs being evolutionarily prior to LPs is much stronger than the evidence for the contrary, and that the evidence for the interdependence of LA and LP is feeble. Observe that the conclusion is not that LAs evolutionarily predate LPs – I merely state that there is significantly more evidence of LAs predating LPs than of the other two possibilities.

Finally, I can try to formulate and answer some important questions that, in the hindsight at least, have been a driving force behind this thesis.

1. Is the noun/verb distinction universal in the world's languages? The answer: Undecided. To answer this question, a consensus must be reached on the definitions of 'noun' and 'verb', as well as on the distributional criteria for the linguistic categories. Currently, there is no such consensus (cf. Baker 2003; Croft 2000, 2001; and chapter 1). Even if an agreement on the definitions and distributional criteria for 'noun' and 'verb' could be reached, a positive or negative answer to the question requires descriptive and typological work of colossal proportions.

2. Are rigorous, formal, and cross-linguistically universal definitions of 'noun', 'verb' and 'word' possible? The answer: At least partly, yes. In chapter 1, I give

‘noun’, ‘verb’ and ‘word’ rigorous, cross-linguistically universal, and for ‘noun’ and ‘verb’ also formal (i.e. grammar-grounded) definitions.

3. Is the linguistic predicate/argument structure isomorphous to that of mathematical logic? The answer: No. The LP/A structure is different from those of mathematical logic and cognition/perception in a number of respects. First, the LP/A structure is inherently multi-leveled: there are up to three levels of LAs and LPs. Second, there are specific rules for generating higher order LAs and LPs, as well as for converting LAs to LPs and vice versa. The predicate/argument structure of mathematical logic does not have anything like this (although the mathematical notation presumably allows to implement these rules – this is a problem that lies outside the scope of the present work). Similarly, the predicate/argument structure in cognition/perception does not come anywhere near the complexity of the linguistic predicate/argument structure.

4. Is there a correspondence between nouns and arguments and verbs and predicates? The answer: Yes. Nouns correspond to linguistic arguments and verbs correspond to linguistic predicates by virtue of rules (a)–(b). A correspondence between nouns and FOPL (or SOPL) arguments and verbs and FOPL/SOPL predicates is not a subject matter of this work. However, it must be noted that (1) FOPL and SOPL are artificial systems that natural language has no need to correspond to, and that (2) noun – FOPL/SOPL argument and verb – FOPL/SOPL predicate correspondences are a matter of translating rules (a)–(b) into FOPL/SOPL (which is beyond the scope of the present study). The rules are as follows (the notation is explained in the Abbreviations and in fn. 18):

- | | |
|--|--|
| (a.1) {LA+} COP+ADJ/LA \Rightarrow LP ₂ | (<i>He is rich/a sportsman.</i>) |
| (a.2) {LA+} ADJ/LA \Rightarrow LP ₂ | (<i>On bogaty/sportsmen.</i> ‘He is rich/a sportsman’ in Russian) |
| (b.1) INF/GER = LA ₂ | (<i>to see / seeing</i>) |
| (b.2) {LP+}[LA+]AC \Rightarrow LA ₃ | (<i>He knows [the man] who was ill</i>) |

5. Can something be conjectured about the evolution of the linguistic predicate/argument structure? The answer: Yes. Previously, Heine and Kuteva (2002, 2007) have delivered certain arguments as to the evolutionary primacy of nouns over verbs within the framework of grammaticalization. In chapter 3, I submit a number of new arguments in favor of the hypothesis that LAs (e.g. nouns) are evolutionarily earlier than LPs (e.g. verbs). Importantly, noun and verb (as well as LA and LP) are complementary but that does not entail interdependence. In fact, a number of asymmetries suggest that (1) LPs depend on LAs but not vice versa, and that (2) LAs are onto- and phylogenetically older than LPs.

Importantly, the present thesis is not confined to answering the questions I detailed above. Some of the more original results are obtained by positing the word class 'flexible' for the stems that can fill both nominal and verbal functions, such as the English *love*, *hate*, *walk* etc. Together with three universal assumptions (12)-(14), 'noun', 'verb' and 'flexible' allow for the framework of five logically possible language types: $N/V/F$, N/F , V/F , N/V and F . At least one of these types must, by definition, be realized in the world's languages but, perhaps inevitably, the present study cannot answer which (if any) of the types besides $N/V/F$ is realized. However, I have made efforts to narrow down the search space with references to particular languages and language families (cf. section 1.6), as well as by facilitating a possible future search with the pervasiveness criterion for linguistic categories (section 1.4). Importantly, the definitions of N, V and 'word' in chapter 1 as well as the results obtained on the N/V , P/A and LP/A structures in chapters 2 and 3 do not depend on the correctness of the underspecification hypothesis that establishes the word type 'flexible' in chapter 1.

SUMMARY IN ESTONIAN

Nimisõna/verb ja predikaat/argument struktuurid

Käesolev töö keskendub mõningaile nimisõna/verb ja predikaat/argument struktuuridega seotud probleemidele. Lähema vaatluse all on tavakeele predikaat/argument struktuur, kuid kuna predikaat/argument struktuurid esinevad ka matemaatilises loogikas ning ilmselt ka tajus (Hurford 2001, 2003b, 2003c), käsitlen ka tavakeele, loogika ja taju predikaat/argument struktuuride erinevuste ja seoste probleemi. Nimisõna/verb struktuur on mõistagi omane vaid tavakeelele. Pole siiski selge, kas nimisõna/verb struktuur esineb kõigis tavakeeltes. Õigupoolest on see küsimus üks vaieldumaid sõnaklasside tüpoloogias (vrld Baker 2003; Broschart 1997; Evans & Osada 2005; Hopper & Thompson 1984; Jacobsen 1979; Peterson 2007; Whorf 1945), olles ühtlasi käesoleva doktoritöö esimese peatüki algpunktiks. Kolm nõ põhiprobleemi²⁶, millega doktoritöö tegeleb, on järgmised:

1. Kas nimisõna/verb struktuur esineb kõigis keeltes?
2. Kas eksisteerib vastavus nimisõna/verb ja predikaat/argument struktuuride vahel?
3. Kas ja mida saab järeldada nimisõna/verb ja predikaat/argument struktuuride evolutsiooni kohta (lisaks Heine ja Kuteva (2002, 2007) poolt väljatoodule)?

Kõik need probleemid on seni lahendamata. Doktoritöö ülesehitus on järgmine: esimene peatükk tegeleb esimese probleemiga, teine teisega ja kolmas kolmandaga. Alljärgnevalt esitan nende peatükkide eestikeelsed lühikokkuvõtted, millele järgneb üldkokkuvõte. Käesolevas eestikeelses ülevaates pole ma piirdunud ainult originaalis esinevate näidetega, vaid püüdnud seal, kus see põhjendatud on, tuua näiteid ka eesti keelest. Eestikeelse kokkuvõtte struktuur on sarnane ingliskeelsele originaalile, olles viimasest mõistagi lühem. Eestikeelses kokkuvõttes on (ala)peatükid nummerdatud rooma, mitte araabia numbrita nagu originaalis; viimasega võrreldes on numeratsioon muutunud ka mõne alapeatüki ärajätmise arvelt.

²⁶ “Nõ”, kuna sõltuvalt vaatenurgast võib põhiliseks osutada midagi hoopis muud. Lisaks nimetatud põhiprobleemidele tuleb tegelda veel hulga alamprobleemidega, mis võivad (sõltuvalt lugeja rõhuasetusest, huvidest jne) põhiprobleemidest lõppkokkuvõttes isegi olulisemaks osutada.

I. NIMISÕNAD, VERBID JA FLEKSIIBLID (JA NENDE TÛPOLOOGILISED IMPLIKATSIOONID)

i.i. Nimisõnad, verbid ja fleksiiblid

On selge, et enne nimisõna/verb struktuuri (mitte)universaalsuse üle otsustamist tuleb 'nimisõna' ja 'verb' defineerida (Rijkhoff 2002). Nimisõna ja verbi definitioone on palju, kuid teaduslikule rangusele ja keelteülesele universaalsusele pretendeerivaid definitioone nende seas vähe. On siiski kaks aspekti, milles peaaegu kõik autorid ühel meelel on: 1. Nimisõna ja verb on leksikaalsed kategooriad. 2. Nimisõna ja verbi süntaktilised funktsioonid on vastavalt argument ja predikaat (Alfieri 2007; Anderson 2004; Anward 2001; Bhat 2000; Broschart 1997; Croft 2005; Helmbrecht 2001; Jacobsen 1979; Peterson 2007; Ramat 1999; Sasse 1993b)²⁷. Üldlevinud arvamuse kohaselt on predikaat ja argument tavakeeles universaalsed. Olen leidnud ainult ühe allika, kus see arvamus kahtluse alla seatakse. Gil (1994: 194) kirjutab riau indoneesia kohta järgmist:

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 – Luuk] as either **masok** (**putih**) “the white one is going in” or **putih** (**masok**) “the going is of the white one”.

Kuid samal leheküljel on *masok putih* analüüsitud kui [EVENT[EVENTmasok] [THINGputih]. Õigupoolest on see analüüs piisav, et kehtestada *masok* predikaadi ja *putih* argumendina (ehkki küll veel mitte lingvistilise predikaadi ja argumendina). Asjaolu, et *masok putih* on grammatiline ilma P ja A nähtava märkimiseta näib kummastav. On siiski võimalik, et P ja A märgitakse fraasis *masok putih* sõnajärgiga. Nt Gil (2000) väidab, et lausealgne positsioon on iseloomulik predikaatidele tagalogi keeles, kus (nii nagu ka riau indoneesia keeles) on eriti nõrk nimisõna/verbi eristus.

Nagu öeldud, on teaduslikule rangusele ja keelteülesele universaalsusele pretendeerivaid 'nimisõna' ja 'verbi' definitioone vähe. Vaatleme kaht hästituntud, kuid väga erinevat näidet. Croft (2000, 2001) defineerib 'nimisõna' ja 'verbi' kui semantilise klassi ja propositsioonilise akti funktsiooni prototüüpilise korrelatsiooni. 'Nimisõna' puhul on propositsioonilise akti funktsioon 'osutus' ja semantiline klass 'objekt (mitte-relatsiooniline, staatiline, pidev, mittegradueeritav)'. Ehkki küll universaalsed, on need definitioonid ähmased, kuna põhinevad prototüüpidel (nt *piinlikkus*, ehkki küll nimisõna, pole prototüüpiline objekt, kuna pole ei mitte-relatsiooniline, staatiline, pidev, ega ka mittegradueeritav). Teine probleem on selles, et definitioonid pole formaalsed,

²⁷ Lisaks süntaktilistele funktsioonidele on nimisõnal ja verbil ka pragmaatilised (nt 'diskursuses manipuleeritud osaline' ja 'raporteeritud sündmus' – Hopper & Thompson 1984) ja denotatiivsed (nt 'asi' ja 'protsess' – Langacker 2004) funktsioonid.

st neil puuduvad grammatilised implikatsioonid. On muidugi võimalik, et nimisõna ja verbi keelteüleselt [*cross-linguistically*] universaalsed ja formaalsed definitsioonid on võimatud. Baker (2003: 23) annab verbile järgneva definitsiooni: “X on verb siis ja ainult siis kui X on leksikaalne kategooria ja X-1 on spetsifitseerija [*specifier*]”. Paraku võivad spetsifitseerijad olla ka nimisõnad ja adjektiividel (Bennett 1995; Putseys 1989; Wehrli 1988); samuti pole selge, mis on spetsifitseerija (täpsemalt vt alaptk 1.2).

Enne nimisõna ja verbi defineerimist võiksime vaadelda keelt, kus need kategooriad väidetavalt puuduvad. Üks kuulsamaid on järgmine nootka k näide (Swadesh 1939: 78–9):

<i>mamu:k-ma</i>	<i>qu:ʔas-ʔi</i>
work-INDIC	man-the
'The man is working.'	
<i>qu:ʔas-ma</i>	<i>mamu:k-ʔi</i>
man-INDIC	work-the
'The working one is a man.'	

Nagu näeme, on kõneviisi marker ja artikkel neis kahesõnalistes lausetes vahetatavad, ja nii *mamu:k* kui ka *qu:ʔas* võivad funktsioneerida nii predikaadi kui ka argumendina. Kuna sõnatüved on P/A struktuuri suhtes sümmeetrilised, pole nad nimisõnad ja verbid, kuna nimisõna ja verb kodeerivad vastavalt argumenti ja predikaati. Ülaltoodud näiteis on argument ja predikaat kodeeritud vastavalt määratleja [*determiner*] -*ʔi* ja TAM markeri -*ma* poolt. Inglise keeles on olukord sama, vrd *he worked* [work-PAST] ja *he did the work* [DET work], kus TAM marker kodeerib samuti predikaati ja määratleja argumenti. Sõnatüvede nagu *jooks, run, walk, love* jne puhul on ainult 3 võimalust:

1. Nulltuletus: Nimisõna nulltuletus verbist või vastupidi, või mõlema nulltuletus eelkategoriaalsest tüvest.
2. Homofoonია: 2 identset tüve, nimisõna ja verbi oma, tuletust pole.
3. Alamääratus: Tüvi on alamääratud (paindlik, sümmeetriline, universaalne) nimisõna/verb eristuse suhtes, tuletust pole.

Kõik need hüpoteesid on mittetestitavad, kuid 3 on kõige parsimoonsem (ökonoomsem). Täpsemalt on need hüpoteesid ‘mittetestitavad tänapäevaste meetoditega’ – kui leksikaalseid üksusi oleks võimalik ajus identifitseerida, saaks neid hüpoteese testida (vt alaptk 1.2, eriti joonealune märkus 7). Et kõik 3 hüpoteesi on mittetestitavad, võtan aluseks hüpoteesi 3 kui kõige ökonoomsema. Hüpoteesist 3 järeldub, et eksisteerib liik predikaate ja argumente kodeerivaid sõnatüvesid, mis on N/V eristuse suhtes alamääratud. Laenates termini seotud mõistest, ‘paindlikud sõnaliigid’ [*flexible parts of speech*] (nt Don & van Lier 2007; Hengeveld 1992; Hengeveld & Rijkhoff 2005; Rijkhoff

2002), nimetan need tüved 'fleksiiibliteks' (F). Seega iseloomustab nii inglise kui ka eesti keelt kolmene N/V/F eristus N/V asemel.

Defineerin 'nimisõna', 'verbi' ja 'fleksiiibli' järgmiselt: **N = sõnatüve omadus aktsepteerida lingvistilise argumendi, kuid mitte lingvistilise predikaadi markereid; V = sõnatüve omadus aktsepteerida lingvistilise predikaadi, kuid mitte lingvistilise argumendi markereid; F = sõnatüve omadus aktsepteerida nii lingvistilise argumendi kui ka lingvistilise predikaadi markereid.** Morfoloogiliselt on 'tüvi', 'juur' ja 'sõna' osundatud järgmiselt: juur = alusvorm; tüvi = juur pluss mistahes arv tuletusafikseid (kui rakendatavad); sõna = tüvi pluss mistahes arv muuteafikseid (kui rakendatavad). Nimisõna, verbi ja fleksiiibli definitsioonidest nähtub, et LP/A eristuse universaalsus on sõltumatu N/V eristuse universaalsusest, kuid mitte vastupidi, kui N ja V on defineeritud nende süntaktiliste funktsioonide LA ja LP järgi, nagu nad seda enamasti on (vrd Bhat 2000; Croft 2005; Peterson 2007; Sasse 1993b). Väidan, et lisaks LP/A eristuse universaalsusele on **LA universaalselt märgitav LP/A sõnajärje kitsenduste, määratlejate ja possessiivide (alamhulga) poolt, ning LP on universaalselt märgitav LP/A sõnajärje kitsenduste, TAM ja tegumoe markerite (alamhulga) poolt.** Paljudes keeltes pole kõiki neid kategooriaid (nt määratlejaid), kuid kõigis keeltes on olemas vähemalt üks kategooria LA ja vähemalt üks kategooria LP markereid. Põhimõtteliselt piisaks juba ühest LA ja LP sõnajärje kitsendusest nii LA kui ka LP märkimiseks. Mitmed käsitlused on pakkunud (mõningaid) neid markereid, nagu ka sugu, arvu, käänet ja isikut nimisõna ja verbi või argumendi ja predikaadi indikaatoreiks (Broschart 1997; Croft 1990; Hopper & Thompson 1984; Peterson 2005, 2007; Sasse 1993b). Põhjus, miks ma sugu, arvu, käänet ja isikut nimisõna markerite hulka ei arva, on selles, et paljudes keeltes märgivad nad (ka) teisi sõnaliike peale nimisõna (Blake 1994; Evans & Osada 2005; Polinsky & van Everbroeck 2003). Täpsemalt vt alaptk 1.2.

i.ii. Näiliselt konfliktset argumendi ja predikaadi markerid

Üks variant LA ja LP konfliktset märkimist viitab olukorrale, kus tüvi on **samaaegselt** märgitud nii LA kui ka LP markerite poolt. Kaks näidet, vastavalt tagalogi ja tonga [*Tongan*] keelest:

Tagalogi

ang b-um-ib-ilí
DET buy-ARV-IMPF.REAL-buy
'the one buying'

Tonga (Broschart 1997: 136)

<i>na'e</i>	<i>kau</i>	<i>faiakó</i>	<i>('a)</i>	<i>e</i>	<i>Siasí</i>
PAST	PL.HUM	teacher.DEF	ABS	ART	Church

'The Church provided the teachers.'

Pakun, et näiline konflikt on lahendatav lihtsa reeglga, et elemendi identiteedi määrab selle süntaktiliselt kaugeim marker. Seega saame [DET *ang* [VC/TAM *bumibilí*]] and [TAM *na'e* [*kau* [*faiakó* DET]]] ('provided the teachers'). Ülaltoodud näiteis märgivad DET ja TAM kui tüvede 'buy' ('ost') ja 'teacher' ('õpetaja') süntaktiliselt kaugeimad markerid tüved vastavalt argumendiks ja predikaadiks (rohkem näiteid leiata alapeatükist 1.3).

Lisaks LA ja LP samaaegsele märkimisele on võimalik ka nende **distributiivne** konflikt. Kui lekseem aktsepteerib kõiki (või vähemalt ligikaudu võrdsel määral) LA ja LP markereid selles keeles, on see lekseem F. Olulised erinevused aktsepteeritud LA ja LP markerite proportsioonis kujutavad endast aga tüpoloogilist probleemi. Sellisel juhul võib kaaluda vähemalt kolme erinevat distributiivset kriteeriumi:

1. Kui tüvi aktsepteerib vähemalt üht LA ja vähemalt üht LP markerit, on see tüvi F.
2. Kui tüvi aktsepteerib kõiki LA kuid ainult mõningaid LP markereid, on see tüvi N (ja kui vastupidi, siis on see tüvi V).
3. Kui tüvi aktsepteerib LA markerite enamust ja LP markerite vähemust, on see tüvi N (ja kui vastupidi, siis on see tüvi V).

Nimetan kriteeriumi 1 eksklusiivseks ning kriteeriume 2 ja 3 inklusiivseiks. Kuigi kriteeriumid 2 ja 3 võivad paista juuksekarva lõhkiajamisena, võib nendevaheline valik mõningail juhtudel siiski paratamatu olla. Nt põhineb Jacobseni vakassi [*Wakashan*] keelte analüüs kas kriteeriumil 2 või 3. Seetõttu klassifitseerib ta lekseeme, mis aktsepteerivad vakassi keeltes nii LA kui ka LP markereid mitte F-i, vaid N-i või V-na. Klassifikatsioon põhineb mh asjaolul, et – kuigi nii Jacobseni N kui V võtavad TAM markereid – on TAM märkimine N-l piiratud duratiivse aspektiga ega esine tulevikus, samas kui TAM märkimine V-l hõlmab kogu TAM paradigmat.

On selge, et suutmatus neid kriteeriume üksteisest eristada ja piirduda korraga vaid ühega neist korraga külvavad N/V tüpoloogias suurt segadust. On üpris tavaline, et inklusiivseid kriteeriume pruukivad uurijad kritiseerivad eksklusiivset kriteeriumi kasutanud teadlaste **tulemusi** ja vastupidi – samas kui tegelikult oleks põhjust kritiseerida ainult kriteeriumi. Kuna pole alust üht kriteeriumi *a priori* teistele eelistada²⁸, leiavad inklusiivseid kriteeriume kasutavad uurijad N/V eristuse kõigis keeltes, mida nad uurivad (Baker 2003; Evans & Osada 2005; Hopper & Thompson 1984; Jacobsen 1979), samas kui

²⁸ Tüpoloogiliselt mõistlikum oleks siiski ilmselt eelistada inklusiivseid kriteeriume.

eksklusiivse kriteeriumi kasutajad leiavad N/V eristuse paljudes keeltes puuduvat (Gil 1994; Itkonen 2000).

i.iii. Tüpoloogilise omaduse hõlmavus

Kui mingis keeles on ainult üks nimisõna, kas siis selles keeles esineb kategooria 'nimisõna'? Kui kasutada tüpoloogilise omaduse hõlmavuse kriteeriumi, siis esineb, kui aga mitte, siis ei pruugi esineda. Nt Hengeveld (1992) loobub oma tüpoloogias hõlmavuse kriteeriumist. Selle tulemusena puudub tuskarora keeles, mida ühel leheküljel kirjeldatakse kui "väheste nimisõnadega" keelt, järgmisel leheküljel kategooria 'nimisõna' (Hengeveld 1992: 67–68). Selline loogika näib arusaamatu, ja paljud uurijad on rõhutanud hõlmavuse kriteeriumi kohustuslikkust (Baker 2003; Evans & Osada 2005; Itkonen 2000); veelgi enam: oma hilisemates töodes lähtuvad ka Hengeveld ja Rijkhoff ise hõlmavuse kriteeriumist (Hengeveld & Rijkhoff 2005; Hengeveld et al. 2004).

i.iv. Viis loogiliselt võimalikku keeletüüpi

Kordame olulisimat punkti ja vaatame, mis sellest järeldub.

1. Igas keeles on LA-d ja LP-d.
2. Tuletus eelnevast: igas keeles on vähemalt üks leksikaalne klass, mis kujutub LA-le ja vähemalt üks leksikaalne klass, mis kujutub LP-le.
3. On kolm võimalust, kuidas leksikaalne klass saab kujutuda LA-le või LP-le: see võib kujutuda LA-le, see võib kujutuda LP-le, see võib kujutuda mõlemale. N-i, V ja F-i definitsioonide järgi (vt punkt i.i) vastavad need kolm võimalust vastavalt N-ile, V-le ja F-ile.
4. Punktidest 2 ja 3 järeldub viie loogiliselt võimaliku keeletüübi olemasolu: $N/V/F$, N/F , V/F , N/V ja F .

Esialgu pole selge, milline või millised neist viiest tüübist reaalselt eksisteerivad. Samas on ilmne, et vähemalt üks neist – ja nimelt tüüp $N/V/F$ (vt punkt i.iv.i.) – peab eksisteerima. See loogilis-tüpoloogiline süsteem ütleb mh, et tüübid N ja V (kus mõlemad ülejäänud sõnaliigid kolmikust $\{N, V, F\}$ puuduvad) on loogiliselt võimatud. Ennustus pole väärtusetu, kuna keeletüüpe V ja N on varem postuleeritud (Hengeveld 1992; Itkonen 2000; Rijkhoff 2003; Sasse 1993b). Alapeatükkides 1.5.2 ja 1.5.3 näitan, et väited keeletüüpide V ja N olemasolu kohta ei pea paika. Kuigi selle kohta on raske midagi kindlat väita, näib siiski, et – kui rakendada hõlmavuse kriteeriumi, – on tõenäoline, et eksisteerib kas ainult tüüp $N/V/F$ või tüübid $N/V/F$ ja F . See on siiski vaid oletus. Pangem tähele, et mõlemad juhud eeldavad F-i kui keelteüleselt universaalset sõnaliiki. Lisaks pole keeletüübi $N/V/F$ olemasolule – nagu kogu sellele loogilis-tüpoloogilisele süsteemile – kirjanduses varem tähelepanu pööratud.

i.iv.i. Tüüp N/V/F

Võimalik, et kõik maailma keeled kuuluvad sellesse tüüpi (kui rakendada hõlmavuse kriteeriumi). Nt on nii inglise kui ka eesti k puhul täidetud tingimus, et keeles esineb vähemalt üks N (nt *prairie*, *maja*), V (*agree*, *maga*) ja F (*run*, *jooks*) tüvi (vrd Bierwisch 2001; Don & van Lier 2007; Farrell 2001; Jespersen 1924; Vogel 2000).

i.iv.ii. Tüüp N/F

On pakutud, et maoori ja niuea [*Niuean*] keel võiksid kuuluda sellesse liiki (Biggs 1971; Massam 2005). See pole siiski selge, kuna vaid ühest V tüvest piisaks, et need keeled kuuluksid N/V/F liiki (vt ka Bauer 1993: 259).

i.iv.iii. Tüüp V/F

Pole teada, et selline keel eksisteeriks, kuid kajuuga [*Cayuga*], samoa ja santali keel näivad olevat kõige tõenäolisemad kandidaadid sellesse tüüpi (Neukom 2001; Sasse 1993a; anonüümne retsensent).

i.iv.iv. Tüüp N/V

Pole selge, kas selline tavakeel eksisteerib, kuid kunstlikud keeled nagu ido ja esperanto on näited N/V tüüpi keelest. Irokeesi ja bantu keeled, kus sõnaliik F võib täielikult puududa, on tõenäolisimad kandidaadid (Baker 2003; Hengeveld 1992; anonüümne retsensent).

i.iv.v. Tüüp F

Sellesse tüüpi kuulub väidetavalt palju keeli. Samas pole ühegi puhul selge, kas N ja V neis ikka täielikult puuduvad. Tüüpiline hinnang kõlab sellele vaatamata u nii:

"(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 & Kinkade 1998: 36, sališi [*Salishan*] keelte kohta).

Vähemalt 18ne keele puhul on oletatud selle kuuluvust tüüpi F (täpsemalt vt alaptk 1.5.5).

i.iv.v. Viis loogilist keeletüüpi: kokkuvõte

Pole selge, kas tüübid N/F, V/F, N/V ja F eksisteerivad, kuna see sõltub järgmistest objektiivsetest kriteeriumitest:

1. Morfoloogilise analüüsi tasand (juur, tüvi, sõna või fraas).

2. N-i, V ja F-i definitsioonid.
3. Kas N-i, V ja F-i definitsioonide interpreteerimisel kasutatakse sama distributiivset kriteeriumi (vt punkt i.ii).
4. Kas rakendatakse tüpoloogilise omaduse hõlmavuse kriteeriumi (vt punkt i.iii).

Ainult tingimusel, et kõik 4 kriteeriumit on samad, on lootust konsensusele viie loogiliselt võimaliku keeletüübi eksisteerimise, ja seega ka N/V eristuse universaalsuse küsimuses. Käesolevas töös on morfoloogilise analüüsi tasandiks, millega N-i, V ja F-i definitsioonid seotakse, sõnatüvi (vt punkt i.i). Distributiivsed kriteeriumid ei mõjuta tüüpide *N* ja *V* loogilist võimatust. Nagu joonealuses märkuses 13 öeldud, on inklusiivsed kriteeriumid eelistatavad eksklusiivsele. Nagu punktis i.iii väidetud, tuleks rakendada ka tüpoloogilise omaduse hõlmavuse kriteeriumi. Nende definitsioonide ja kriteeriumite korral on kõige tõenäolisem oletus, et kõik maailma keeled kuuluvad tüüpi *N/V/F* või tüüpidesse *N/V/F* ja *F*. Rõhutan veel: see on kõigest oletus. Kindel vastus eeldaks valitud kriteeriumite konsensuslikkust, ja tohutut deskriptiiv-tüpoloogilist uurimistööd, mis ületaks kaugelt doktori- (ja isegi elutöö) mahu.

II. NIMISÕNA/VERB JA PREDIKAAT/ARGUMENT STRUKTUURID

ii.i. Sissejuhatus: keelefossiilid

Jackendoff (1999) on postuleerinud 'keelefossiilide' kui evolutsiooniliselt fundamentaalsete struktuur-funktsionaalsete tüüpide olemasolu keeles. Näiteks toob ta N-N liitsõnad nagu *snowman* või *bellboy*, ja printsiibid nagu grupeerimine (modifitseerijad on selle kõrval, mida nad modifitseerivad), Agent Esimesena ja Fookus Viimasena. Pakun, et ka lingvistilist predikaat/argument struktuuri on põhjust vaadelda keelefossiilina.

ii.ii. Nimisõna-argument ja verb-predikaat vastavused tavakeeles

Hurford (2003b, 2003c) on osundanud ühele probleemile Aristotelese predikaat/argument struktuuris. Probleem seisneb selles, et sama liiki term [*term* – ilmselt loogikalises, mitte 'termini' tähenduses] võib täita nii argumendi kui ka predikaadi rolli. Nt on term *mees* lause *Mees suri* argument ja lause *Platon on mees* predikaat. Esimest järku predikaatloogika on aga, nagu Hurford ütleb, "distantseerunud tavakeele pindmistest vormidest", mistõttu sama (liiki) term ei saa olla nii argument kui ka predikaat. Hurford ütleb, et tavakeeltele tüüpiline N/V eristuse ümber koondunud struktuur vajab seletust. Näidates, kuidas

nimisõna–LA ja verbi–LP vastavused tavakeeles kehtestatakse, annan allpool seletuse esimese osa. Et predikaatloogika on kunstlik keel, millele tavakeel ei pea vastama, siis seletuse teine pool – vastavuse kehtestamine tavakeele ja esimest järku predikaatloogika argumentide ja predikaatide vahel – on tehniline küsimus, mis taandub tavakeele argumentide ja predikaatide esitamisele predikaatloogikas.

Vaatleme mõningate lausete lingvistilisi argumente ja predikaate (põhjalikuma käsitluse leiad alapeatükist 2.3):

[A *John*][P *jookseb*]
 [A *John*][P *joob*][A *vett*]
 [A *John*][P *on*][A *mees*]]

Väidan, et eesti, inglise ja paljudes teistes keeltes spetsifitseerib koopula sekundaarse predikaadi, nagu kolmandas näitelauses. Mõnedes keeltes (nt vene, malta ja maroko araabia) saab sekundaarset predikaati spetsifitseerida (ka) kaht argumenti kõrvuti seades. Näide vene keelest:

[A *On*][P[A *sportsmen*]].
 Sõnasõnalt: Ta sportlane
 'Ta on sportlane'

Vähemalt vene keeles on see võimalik ainult olevikus ja indikatiivses kõneviisis (teistes ajad ja kõneviisid nõuavad koopulat).

Inglise keeles tuletab sufiks *-ing* verbi või fleksiibli tüvest sekundaarse argumendi:

[A *He*][P *started*][[*fir*]A *-ing*].

Sellisel saadud sõna on gerund. Gerundid ja infinitiivid on primaarsete predikaatide V ja F tüvedest tuletatud sekundaarsed argumendid. Inglise k infinitiivi näide:

[A *I*][P *want*][A *to*][P *go*]].

Vene, eesti jt keelte gerund ja infinitiiv toimivad sarnaselt inglise keele omadele²⁹:

²⁹ Eesti k on *-mine* liidet varem analüüsitud kui nimisõnatuletust (Erelt, Erelt & Ross 2000; Erelt et al. 1995). Gerund on nimisõnatuletusega ekvivalentne ses mõttes, et mõlemad annavad tulemuseks LA. *-mine* liite absoluutne produktiivsus – see, et ta liitub igale V/F tüvele – viitab siiski tema kuulumisele V/F paradigmasse, nagu ka see, tegu on 'tegevusliitega' (tegevuslikkus tuleb tõenäolisemalt aluseksolevast V/F tüvest kui liitest, mis pidi ju nimisõna tuletama). Seetõttu vaatlen *-mine* vormi kui gerundi.

[A On][P ljuibit][A chitat']. [A Talle][P meeldib][A lugeda].
 [A On][P ljuibit][A chtenie]. [A Talle][P meeldib][A lugemine].

Esimene on infinitiivi, teine gerundi näide.

'Argumendil' on keeleteaduses veel üks tähendus: nimelt võivad osalauseid funktsioneerida argumentidena pealause predikaadi all (Dixon 2006). Sellisel juhul nimetatakse neid argumentlauseiks (komplementlauseiks); neid genereeritakse lisades argumentlause sidendile [*complementizer*] osalause, nagu näidetes *I know **that** he will come; Ma tean, **et** sa tuled; I know **where** you hid it; Ma tean, **kuhu** sa selle peitsid* jne. Argumentlause sidendid *that* ja *et* võib ka ära jätta, vrd *I know he will come; Ma tean, sa tuled*. Osalauseid koosnevad ka ise LA-dest ja LP-dest ja on oma olemuselt predikatiivsed (vb seetõttu, et nende peasõna on verb või fleksiibel). Seega võib kõiki pea- ja märkimata lauseid (nagu öeldud, toimib märkimine argumentlause sidendiga või osalauseste järjekorraga) nimetada 'predikaatlauseiks'. Näited (AC = argumentlause, PC = predikaatlause):

[PC [A I][P know][AC that [A John][P knows]].]
 [PC [A Ma][P tean][AC et [A John][P teab]].]

Sellest, et PC-d ohjab predikaat, ei järeldu, et PC ise oleks predikaat (täpsemalt vt alaptk 2.3). AC-d ja PC-d on paratamatult kõrgemat järku üksused kui LA-d ja LP-d. On vaid üks võimalus, kuidas LA või LP võib sisaldada PC-d või AC-d. See võimalus on liit-XP³⁰ (nt *mees kes oli haige*). Liit-XP-s liidetakse AC LP-le (millele võib, aga ei pruugi järgneda LA). Reegel on seega järgmine:

{LP+}[LA+]AC ⇒ LA₃ (*Ta teab [meest] kes oli haige*)

Reegel ütleb, et AC-d, millele eelnevad LP-d (millele valikuliselt järgnevad LA-d) transleeritakse kolmandat järku LA-deks.

Olukord on seega järgmine: meil on kaks komplekti reegleid: (a) genereerib lingvistilisi predikaate adjektiividest ja lingvistilistest argumentidest, ja (b) genereerib lingvistilisi argumente verbide ja fleksiiblite tüvedest (tulemuseks infinitiivid ja gerundid), argumentlauseist ja lingvistilistest argumentidest. Siin on neli näidisreeglit (kuna reeglid on osaliselt keelespetsiifilised, võib nimekiri olla mittetäielik)³¹:

³⁰ XP on Hurfordi (2007) teorianeutraalne vaste NP/DP-le – täpsemalt vt joonealune märkus 10.

³¹ / tähendab "või"; ⇒ näitab genereerimist (transleerimist); {...} tähistab reegli seisukohalt vajalikku kuid mittetransleeritavat tingimust; [...] tähistab valikulist kuid transleeritavat tingimust; T_s on liiki T kuuluva sõna tüvi antud keeles; T_n on liiki T kuuluv element tuletustasandil n, kus n on tuletustasandile (primaarne, sekundaarne, tertsiaarne jne) vastav naturaalarv.

- (a.1) {LA+} COP+ADJ/LA \Rightarrow LP₂ (*He is rich/a sportsman.*)
 (a.2) {LA+} ADJ/LA \Rightarrow LP₂ (*On bogaty/sportsmen.* ‘He is rich/a sportsman’ in Russian)
 (b.1) INF/GER = LA₂ (*to see / seeing*)
 (b.2) {LP+} [LA+] AC \Rightarrow LA₃ (*He knows [the man] who was ill*)

Kuna LA, LP, ADJ, INF ja GER võivad võtta komplemente ja/või adjunkte, sisaldavad reeglid ainult vajalikke komponente. Täielik ülevaade süntaksist eeldab XP-d ja predikaatfraasi, mis on vastavalt LA ja LP projektsioonid (vt joonealune märkus 10 ja punkt ii.ii.i). (a.2) on variant (a.1)-st keeltele, kus koopula kas puudub või võib selle (mõnes) predikatiivses konstruktsioonis ära jätta (täpsemalt vt alaptk 2.3). Järgnevalt mõned näited, mis illustreerivad reegleid (a)–(b):

[A₁ John][P₂ is [A₁ a painter]].
 [A₁ John][P₂ on [A₁ kunstnik]].
 [A₂ Seeing][P₂ is [A₂ believing]].
 [A₂ Nägemine][P₂ on [A₂ uskumine]].
 [A₂ To live][P₂ is [A₂ to die]].
 [A₂ Elada][P₁ tähendab][A₂ surra].
 [A₁ It][P₂ is clear][A₃ [AC₁ that [A₁ John][P₂ is ill]]].
 [P₂ On selge][A₃ [AC₁ et [A₁ John][P₂ on haige]]].

ii.ii.i Adjektiivid, adverbid ja adpositsioonid lingvistilises predikaat/argument struktuuris

Nimisõnad on prototüüpilised LA-d, verbid prototüüpilised LP-d ja fleksiiblid võivad olla mõlemat sõltuvalt markeeringust, kuid milline on adjektiivide, adverbide ja adpositsioonide roll lingvistilises predikaat/argument struktuuris? Kuna adjektiivid ja adverbid on tavaliselt N, V, F või ADJ adjunktid, on nende LP/A staatus määratud nende leksikaalse peasõna poolt. Sel viisil saavad adjektiividest ja adverbidest XP-de ja predikaatfraaside osad. Kuna VP tavaliselt sisaldab LA-d (tüüpiline VP on nt *kana* [VP nägi [LA *poissi*]]), ei sobi VP predikaatfraasiks, kui LP/A eristust soovitakse säilitada. Predikaatfraasi mõningad näited on toodud allpool (täpsemalt vt alaptk 2.3.1):

He [P *talked interestingly*].
Ta [P *rääkis huvitavalt*].
This [P *is very interesting*].
See [P *on väga huvitav*].
This [P *was almost perfect running*].
See [P *oli peaaegu täiuslik jook*].
I [P *know*] *this place* [P *well*].
Ma [P *tean*] *seda kohta* [P *hästi*].

Nagu näha, ei pea modifitseerijad (adjunktid) asuma nende poolt modifitseeritavate peasõnade kõrval. Kahes viimases näites tuleks tarindeid *know* (x) *well* ~ *tean* (x-i) *hästi* analüüsida kui üht predikaatfraasi, mitte kaht predikaati.

XP-d võivad sisaldada adjektiive ja adverbe vastavalt N/F/GER ja ADJ adjunktidenä:

[A₁ *a nearly perfect evening/walk/skiing*]
[A₁ *peaaegu täiuslik õhtu/jooks/suusatamine*]

Infinitiivid ja (nt ingl keeles ka) gerundid saavad võtta adverbe adjunktidenä. Kuna V ja F võivad omada XP-sid komplementidenä, võivad seda ka infinitiivid ja gerundid (sest INF ja GER tuletatakse V ja F tüvedest):

[A₂ *Buying* [A₁ *a BMW*] *quickly*] *is our goal*.
[A₂ *To buy* [A₁ *a BMW*] *quickly*] *is our goal*.
Meie eesmärk on [A₂ [A₁ *BMW*] *kiire ostmine*].
Meie eesmärk on [A₂ *osta* [A₁ *BMW*] *kiiresti*].

Adjektiividel on nii predikatiivne (*ilm on ilus*) kui ka argumentaalne (*ilus ilm*) kasutus. Eesti ja inglise k kodeeritakse predikatiivsed adjektiivid verbidest erinevalt, st neid ei märgita LP markeritega. Paljudes teistes keeltes (nt kirde-ambae, korea, lao, kangi [*Qiang*] ja semelai k) kodeeritakse predikatiivseid adjektiive ainult LP markeritega (Hajek 2004). Paljudes keeltes on võimalikud mõlemad kodeerimisvõimalused (Stassen 2008). Kõigis teistes keeltes peale lao näib adjektiivide TAM-märgimine piirduvat TAM paradigma pärisalamhulgaga (Hajek 2004)³². Kokkuvõttes: millal iganes on adjektiivid märgitud LA või LP markerite poolt, on nad juba definitsiooni järgi LA-d või LP-d.

Adpositsioonid on LA-LP seose modifitseerijad. Mõned näited:

Ta jalutas [A *maja taga*].
Ta jalutas [A *maja ees*].
Ta jalutas [A *maja sees*].

LA-LP seose modifitseerijaina tuleks adpositsioone analüüsida kui LA või LP perifeerseid osi. Kaassõnafraasid (vt kolm eelmist näidet) on näited esimestest ja ühendverbid viimastest:

Ta [A *visati välja*].
Ta [A *logis sisse*].

³² See on põhjus, miks lao adjektiive on analüüsitud verbide alamliigina (Enfield 2004).

ii.ii.ii. Lingvistiline predikaat/argument konverteerimissüsteem. Nimisõna-argument ja verb-predikaat vastavused

Lingvistiline predikaat/argument konverteerimissüsteem töötab järgmiselt. Reegel (b.1) konverteerib primaarseid LP tüvesid sekundaarseteks LA-deks. Reeglid (a) konverteerivad primaarseid, tertsiaarseid ja (osaliselt) ka sekundaarseid LA-sid sekundaarseteks LP-deks. Kui primaarsed ja sekundaarsed LP-d on argumentlause koosseisus, saab neid reegli (b.2) järgi konverteerida tertsiaarseteks LA-deks:

We know [A_3 [AC_1 *that John* [P_1 *sleeps*]]].

Me teame [A_3 [AC_1 *et John* [P_1 *magab*]]].

We know [A_3 [AC_1 *that John* [P_2 *is ill*]]].

Graafiliselt kujutatuna näeb LP/A konverteerimissüsteem välja selline:

$(A_1, A_2^*, A_3) \rightarrow (a) \rightarrow P_2$

$P_{1s} \rightarrow (b.1) \rightarrow A_2$

$(P_1, P_2) \rightarrow (b.2) \rightarrow A_3$

Nooled tähistavad võimalikke konversioone ja nooltele on kirjutatud reeglid, mis neid konversioone võimaldavad. Mitte kõik sekundaarsed argumentid (nt vene infinitiivid – täpsemalt vt alaptk 2.3) pole konverteeritavad sekundaarseteks predikaatideks. Kuivõrd infinitiivid kuuluvad liiki A_2 , tähistab A_2^* piirangut, et mitte kõik seda tüüpi elemendid pole konverteeritavad sekundaarseteks predikaatideks. LP/A konverteerimissüsteemist saab teha kolm järeldust. Esiteks, kõiki LP-sid saab konverteerida LA-deks, kuid mitte vastupidi. Teiseks, kuna (a.1) ja (a.2) on sisuliselt sama reegel, mis on erinevate keelte, aegade ja kõneviiside jaoks kergelt erinev (vt alaptk 1.3), on rohkem predikaate argumentideks kui argumente predikaatideks konverteerivaid reegleid. Kolmandaks, keeles on kolm tasandit argumente, kuid ainult kaks tasandit predikaate. Koos P/A struktuuri sisemise asümmeetriaga (argumentid kui elemendid ja predikaadid kui operatsioonid nende elementidega) viitavad need järeldused sellele, et LA-d võivad evolutsiooniliselt LP-dele eelneda. Sellel küsimusel peatume järgmises peatükis.

Kui vaadelda ainult LA ja LP tuumkomponente (st kui ignoreerida adjunkte ja adpositsioone), taandub meid huvitav leksikon hulgale $\{N, V, F, ADJ, LAx, LPx\}$. N, LAx poolt märgitud F ja LAx poolt märgitud ADJ on primaarsed LA-d. V, LPx poolt märgitud F ja LPx poolt märgitud ADJ on primaarsed LP-d. Nüüd on meil ülevaade sellest, mis on LA-d ja LP-d:

$LA = \{N, F+Lx, ADJ+Lx, (b), PJ(LA)\}$

$LP = \{V, F+LPx, ADJ+LPx, (a), PJ(LP)\}$

LA-d on nimisõnad, LAX poolt märgitud F ja ADJ ning reeglite (b) saadused. LA-de hulka kuuluvad ka LA-de projektsioonid (st XP-d). LP-d on verbid, LPx poolt märgitud F ja ADJ ning reeglite (a) saadused. LP-de hulka kuuluvad ka LP-de projektsioonid (st predikaatfraasid). Sellest, et kõigis piisavalt arenenud tavakeeltes on LA ja LP, ei järeldu, et kõigis neis keeltes esineks kõik kaheksa elementi eeltoodud hulkadest. Nt pole selge, kas kõigis keeltes esinevad nimisõnad ja verbid (Anderson 2004; Bach 2004; Laudanna & Voghera 2002), ja väited adjektiivide mitteuniversaalsusest (Beck 2002; Hengeveld 1992; Junker 2003; McCawley 1992; Rijkhoff 2000; Sasse 1993b) on sagedasemad kui vastupidised väited (Baker 2003; Dixon 2004).

Nagu eeltoodust nähtub, kehtivad tavakeeles nimisõna-LA ja verb-LP vastavused (vrd LA ja LP hulga ülal), kusjuures reeglite (a)–(b) abil on lahendatud ka Hurfordi poolt postuleeritud mittevastavuse probleem. See on oluline samm täieliku vastavuse kehtestamiseks tavakeele ja predikaatloogika P/A struktuuride vahel. Kuna predikaatloogika on kunstlik süsteem, millele tavakeel ei peagi vastama, on järgmine samm – tavakeele P/A struktuuri väljendamine predikaatloogikas – võrdlemisi formaalne, kuuludes sellisena pigem matemaatilisse loogikasse kui keeleteadusesse.

III. NIMISÕNA/VERB JA LINGVISTILINE PREDIKAAT/ARGUMENT STRUKTUURIDE EVOLUTSIOON

iii.i. Ülevaade probleemist

On väidetud, et lause / nimisõnafraas (S/NP) või, reformuleerituna, S/XP eristus on tavakeeles universaalne (Carstairs-McCarthy 1999; Hurford 2007). Pakun, et S/XP universaalsus tuleneb LP/A universaalsusest. Enamus lause definitsioone koonduvad ümber LP e grammatilise predikaadi (erandiks on nt Meillet'ist ja Bloomfield'ist lähtuv “grammatika suurim ühik” või “vorm, mis pole ühegi teise vormi osa” – Graffi 2001: 1843). Võttes aluseks LP-keskse S-i definitsiooni, ei saa olla S-i enne LP-d, ja ei saa olla LP-d enne LP/A eristust (vrd alaptk 3.3, punktid 1 ja 6). Seega on LP/A eristus fundamentaalsem kui S/XP eristus. S/XP tuleneb LP/A-st, kuna LP paneb aluse S-ile ja LA XP-le.

Niisiis on S/XP taandatav LP/A-le. Mida on võimalik öelda LP/A struktuuri päritolu kohta? Mõistagi ei saa anda detailset ülevaadet selle tekkimisest. Kuid kui Hurfordil (2003b, 2003c) on õigus, et pertseptuaalne P/A struktuur esineb paljudel imetajaliikidel, peaksime eeldama, et vähemalt mõningail liikidel projitseerus pertseptuaalne P/A struktuur kontseptuaalsesse struktuuri, andes tulemuseks kontseptuaalse P/A struktuuri. See kontseptuaalne P/A struktuur omakorda projitseeriti mingis keele evolutsiooni varases staadiumis keelde LP/A struktuurina. Esialgu on see kõik, mida selle kohta järeldada võib.

Matemaatiliselt on funktsioon seose pärisalamhulk. Seos omakorda on väljendatav predikaadina. On väidetud, et selgmine ja kõhtmine juhtetee ajus tegelevad vastavalt argumenti ja predikaadi töötusega (Hurford 2001, 2003b, 2003c). See väide käib siiski pertseptuaalse, mitte lingvistilise ja loogikalise P/A struktuuri kohta. Kui pertseptuaalne P/A struktuur peaks tõesti olema ajju sisse ehitatud, siis keeleline ja loogikaline P/A mitte ainult ei eelda pertseptuaalset P/A-d, vaid viimasel on tõenäoliselt implikatsioone ka kontseptualisatsioonile. Põhierinevus loogilise/lingvistilise ja pertseptuaalse P/A vahel on keerukuses. Mõned esimese argumentid on teises analüüsitud kui argumente võtavad predikaadid, nt

CAME(man)
MAN(x) & CAME(x)

Suurtähed märgivad predikaate, väikesed argumente (x on deiktiline muutuja), & on konnektiiv, esimene on lingvistiline ja teine pertseptuaalne tõlgendus. Teine viitab protokeelele³³, kus sõnade loogikaline vorm oleks PREDIKAAT(x) ja sõnu võiks konkateneerida suvalises järjekorras. Seda võimalust toetab ka fakt, et LP/A eristus tehakse grammatiliselt. Seega ei olnud see eristus enne grammatikat juba definitsiooni järgi võimalik. Fakt, et kõigis maailma keeltes on grammatika ja LP/A eristus, võib viidata ekvivalentsusseosele nende kahe vahel. Pole siiski veel selge, kas esimene grammatikareegel kehtestab LP/A struktuuri. Sellest pikemalt allpool.

Sõnajärg on lihtsaim grammatiline vahend. Esimene reegel, millest piisas primitiivseks grammatikaks, oli tõenäoliselt sõnajärje kitsendus (Heine & Kuteva 2002; Johansson 2006). Esimene sõnajärje kitsendus viib **automaatselt** kahe grammatiliselt eristatud sõnakategooria tekkeni³⁴. Kuid me ei tea, kas esimesed grammatilised kategooriad olid sõnaliigid (nagu *mees minema*) või lihtsalt semantilised rollid (nagu *mees mets*, interpreteeritud kui *mees minema metsa* vms). Teised autorid on pakkunud esimesteks grammatilisteks kategooriateks sõnaliike (Heine & Kuteva 2002; Hurford 2003a; Newmeyer 2003), kuid välistatud pole ka semantilised rollid.

Oletame, et algne grammatiline eristus LA ja LP vahel tehti ainult sõnajärgiga. Selline keel oleks olnud palju ramedam kui kaasaegne inimkeel. Siiski oleks see võimaldanud väljendada **sündmusi** – tegevustesse/muutustesse haaratud objekte/omadusi. LP/A struktuur näib olevat (eelkirjeldatud mõttes) sündmuste väljendamise eeltingimus tavakeeles.

³³ Mitte segi ajada algkeelega, nagu protouurali jne. Protokeel (või -keeled) on kõige primitiivsemad hüpoteetilised inimkeeled, kus süntaks ja grammatika kas puudub või on väga rudimentaarne (vt Bickerton 1990; Bowie 2008; Dessalles 2008; Jackendoff 1999; Johansson 2006).

³⁴ V.a. juhul, kui tegu on fonoloogilise (nt, et vokaaliga algav sõna on esimene) või leksikaalse (nt, et mõistet 'puu' tähistav sõna on esimene) kitsendusega. Ma pole teadlik sellistest kitsendustest maailma keeltes, kuid nende olemasolu ei saa välistada.

iii.ii. LA evolutsiooniline eelnevus LP-le: üksteist argumenti

Allpool esitan üksteist argumenti, mis toetavad hüpoteesi, et LA-d on evolutsiooniliselt varasemad kui LP-d. Kuid esmalt tuleb tõrjuda üks eksiarvamus. Väide, et LA-d võisid tekkida enne LP-sid pole loogiliselt vastuoluline. LA-d ja LP-d on komplementaarsed, kuid sellest ei järeldu vastastikust sõltuvust. Näide: duaal ja mitmus on komplementaarsed ilma vastastikuse sõltuvuseta (duaal sõltub mitmusest, kuid mitte vastupidi). LA ei implitseeri LP/A vastavust. Viimase puudumisel oleks ainult üks asi pidanud olema teisiti: kuna LA oli ainus sõnaliik, polnud vajadust selle märkimise järele. Pole teada, mis olid esimesed sõnad, kuid järgnevad üksteist argumenti peaksid tegema selgeks, et kõige tõenäolisemalt nad sarnanesid nimisõnadele.

1. LP-d eeldavad LA-sid, millele nad rakenduvad. Predikaat rakendub argumendile, mis on ette antud (Hurford 2003b). See on põhjus, miks keel ilma LA-deta on peaaegu kujuteldamatu, samas kui keel ilma LP-deta näib üsnagi võimalik. Lausungit *laev Tallinn homme* võib hõlpsasti tõlgendada kui “laev saabub/lahkub Tallinnasse/Tallinnast homme”, kuid samatähenduslik nimisõnade ja pärisnimedeta konstruktsioon oleks ilmselt üpris raskesti mõistetav. Asümmeetria on P/A struktuuri sisemine omadus (Hurford 2003c). Budd (2006) on pakkunud, et keerukates asümmeetriliste sõltuvustega süsteemides on funktsionaalselt vajalikud tuumkomponendid ‘ebavajalikega’ võrreldes varem tekkinud. Sõnade seas on LA-d kindlaimad kandidaadid funktsionaalselt vajalikuks tuumkomponendiks. Vt ka punkt 6 allpool.
2. Laste varases produktiivses sõnavaras domineerivad nimisõnad ja laste arusaamine objektinimedest eelneb nende arusaamisele seoseid väljendavatest mõistetest (Fisher 2002; Gentner & Boroditsky 2001; Gleitman 1993; Waxman & Lidz 2006). Kuigi on väidetud, et nimisõnade varane domineerimine pole keelteüleselt universaalne, on tõendid selle poolt nõrgemad kui tõendid selle vastu (Gentner & Boroditsky 2001; Gopnik 2000).
3. Virtuaalne eksperiment (Steels et al. 2002) on identifitseerinud tingimuse, mis soosib nimisõnu (st LA-sid) esimeste sõnadena – tingimuse, et agentidel peavad olema paralleelsed mitteverbaalsed võimalused interaktsiooni eesmärkide saavutamiseks (nt osutamine). Tegevustele/ muutustele on raske osutada teisiti kui neid imiteerides või sooritades. Seega oleks esimesed LP-d esimeste LA-dega võrreldes olnud detailsemad žestikulaarses modaalsuses. See iseenesest ei välista, et LP-d tulid esimesena (nt on pakutud, et keel sai alguse kui “a mixture of isolated grunts and gestures” – Bickerton 2003: 81). Fakt, et keele jaoks

valiti vokaalne, mitte žestikulaarne modaalsus, soosib esimeste sõnadena siiski LA-sid.

4. LA-d apelleerivad kujundite geomeetrilistele ja LP-d kinesteetilistele omadustele. Pylyshyni järgi on kujunditele olemuslikud pigem geomeetrilised kui dünaamilised omadused. Me suudame kujutleda füüsikaseaduste rikkumist, kuid on peaaegu võimatu kujutleda geomeetria või geomeetrilise optika aksioomide rikkumist (Pylyshyn 2002). Nimisõnu seostatakse prototüüpiliselt pigem geomeetriliste ning verbe pigem kinesteetiliste omadustega (täpsemalt vt alaptk 3.3 punkt 4). Pangem tähele, et kinesteetilised omadused eeldavad geomeetrilisi – nt on võimatu kujutleda liikumist väljaspool aegruumi. Koos nimisõnade seostumisega geomeetriliste ja verbide seostumisega kinesteetiliste omadustega viitab see asümmeetriline sõltuvus (kujundite kinesteetiliste omaduste sõltumine geomeetristest, kuid mitte vastupidi) sellele, et nimisõnad on kognitiivselt fundamentaalsemad kui verbid ning verbid kognitiivselt kompleksemad kui nimisõnad. See omakorda viitab sellele, et nimisõnad (LA-d) võivad evolutsiooniliselt eelneda verbidele (LP-dele).
5. Eri grammatilistesse kategooriatesse kuuluvaid sõnu saab selektiivselt kahjustada. Nimisõnade domineerimine seostub sageli Wernicke ja Broca afaasiatega ja verbide domineerimine anoomilise afaasiaga (Mondini et al. 2004). Verbide selektiivne kahjustus on sagedasem kui nimisõnade selektiivne kahjustus (Arevalo et al. 2007; Luzzatti et al. 2002). Sellele on kaks teineteist mittevälistavat seletust: 1. Vasaku poolkera keelepiirkondade ulatuslik kahjustus kutsub esile leksikaalsed võimed paremas poolkeras, viimased piirduvad aga sagedamini kasutatavate konkreetsete nimisõnadega (Crepaldi et al. 2006). 2. Verbide selektiivne kahjustus on argumentstruktuuri keerukuse funktsioon, argumentstruktuuri keerukus seostub aga regulaarselt verbidega. On näidatud, et kahjustused on suuremad kolme kui kahe argumendiga, ja kahe kui ühe argumendiga verbidel (Kim & Thompson 2000; Luzzatti et al. 2006). Veelgi enam – argumentaalsete nimisõnade nagu itaalia *passaggiata* 'jalutuskäik', *risata* 'naer', *pugnalata* 'noahoop' jne produktsioon kahjustub samal määral argumentaalsete verbide produktsiooniga (Collina et al. 2001)³⁵. Kõik need leiud on kooskõlas kahe hüpoteesiga. (1) N/V topelt-dissotsiatsioon afaasias on kontseptuaalse P/A topelt-dissotsiatsiooni tagajärg (asjaolu, et argumentaalsed nimisõnad kahjustuvad võrdsel määral argumentaalsete verbidega viitab

³⁵ Pangem tähele, et paljud argumentaalsed "nimisõnad" ja "verbid" on tegelikult fleksiiblid: vrd *pianto* 'nutt' – *piangere* 'nutma', *urlo* 'karje' – *urlare* 'karjuma', *passaggiata* 'jalutuskäik' – *passaggiare* 'jalutama', *pugnalata* 'noahoop' – *pugnalare* 'pussitama'.

pigem kontseptuaalse P/A kui LP/A topelt-dissotsiatsioonile). (2) N/V topelt-dissotsiatsioon on argumentstruktuuri keerukuse efekt. Kuna hüpoteesidel (1) ja (2) on palju korreleeritud omadusi, on raske neid teineteisest lahti harutada. Kokkuvõttes: fakt, et LP-de töötlus on spetsialiseeritum ja/või ressursimahukam kui LA-de töötlus, viitab sellele, et viimased võivad olla evolutsiooniliselt varasemad.

6. Kõigis tavakeeltes on LP süntaksi nurgakivi. Tavakeele süntaks baseerub põhimõttel, et LP-d võtavad argumente, mis on eristatud analüütiliste (sõnajärg, adpositsioonid) ja/või sünteetiliste (morfoloogiliste) käändemarkeritega. Seega näib tavakeele süntaksi ja LP vahel kehtivat ekvivalentsusseos (kui süsteemis on LP, kehtib seal tavakeele süntaks, ja kui süsteemis kehtib tavakeele süntaks, on seal LP). Lisaks on LA-d kasulikud ka ilma süntaksita, LP-de kasulikkus ilma süntaksita on aga kaheldav (kuigi imperatiivid võivad olla süntaktiliselt sõltumatud, kuna nad on optimeeritud produktsiooni ja analüüsi kiirusele³⁶). Koos aksioomiga, et alguses oli keel süntaksivaba (Bickerton 2003; Jackendoff & Pinker 2005), soosib hüpotees, et LP ja tavakeele süntaks on ekvivalentsed, esimeste sõnadena LA-sid.
7. Olulisemate sõnaliikide süntaktilisi funktsioone analüüsides on sageli pakutud, et nimisõnade (sh asesõnade ja pärisnimede) funktsioon on fundamentaalseim. Jespersenile, Lésniewskile and Ajdukiewicz'ile viidates väidab Lyons, et nimisõnad on "esimest järku kategooriad", ja et "kõik teised sõnaliigid on tuletatud, kompleksed kategooriad. Teist järku kategooriad kombineeruvad esimest järku kategooriatega (grammatikareeglite järgi /---/), moodustades lauseid /---/" (Lyons 2004: 219–220). Semantilisi klasse (olukord, sündmus, koht, aeg jne) analüüsides märgib Anward (2001: 730), et "kui isiku/asja semantiline klass näib olevat leksikaliseeritav nimisõnade poolt, siis teisi semantilisi klasse saab leksikaliseerida erinevatel viisidel".
8. Nichols on formuleerinud ajaloolise morfoloogia kaks olulist printsiipi: 1. Peasõnasuunaline migratsioon: "Kui adpositsioon või afiksaalse morfoloogia tükk liigub, läheb see laiendilt peasõnale, mitte vastupidi" (Nichols 1986: 86). 2. Reduktsioon: algsed laiendid muutuvad kliitikuteks, ja lõpuks morfoloogilisteks markeriteks peasõnal. Printsiip 1 viitab sellele, et algne märkimine ilmub tõenäolisemalt laiendil kui peasõnal. Koos viitavad need printsiibid morfoloogilise migratsiooni muustrile "laiendilt peasõnale" (nt N-lt V-le). Fakt, et verbaalne morfoloogia näib maailma keeltes olevat üldiselt rikkam kui nominaalne morfoloogia, on sellega kooskõlas. Seda toetavad tõendid on muidugi kaudsed, kuid on põhjust arvata, et morfoloogia ilmus esmalt vanematel

³⁶ Imperatiive vaatleme täpsemalt allpool.

elementidel. Kuna morfoloogia hägustab lekseemi vormi ja tähendust, peavad viimased olema piisavalt konventsionaliseerunud, et morfoloogia saaks lekseemile liituda. On usutav, et vanad elemendid on konventsionaalsemad kui uued. Teiseks, mida kauem mingi element on eksisteerinud, seda rohkem (puht-statistiliselt) on sel olnud võimalusi liita morfoloogiat. Seega oleks vaikumisi eeldus, et element, millest sai laiend, on vanem kui element, millest sai selle peasõna. Moodustajaliikide ja nende peasõna-laiend suhete analüüs kinnitab seda. Vrd järgnev tabel (Helmbrecht 2001: 1425 põhjal):

Tabel 1. Moodustajaliigid ja peasõna-laiend suhted

Moodustaja	Peasõna	Laiend
1. NP	N	ADJ
2.	ADP	N
3. Osalause	V	N
4.	AUX	V

Heine ja Kuteva (2002, 2007) järgi on kolmes paaris (2, 3, 4) laiend vanem kui peasõna. Ühes paaris (1) on olukord vastupidine. Seega on tõendeid laiendi evolutsioonilisest eelnemisest peasõnale rohkem kui tõendeid vastupidisest. Koos Heine ja Kuteva (2002, 2007) poolt esitatud kaalutlustega on see tagasihoidlikuks tõendiks, et paaris 3 on N vanem kui V.

9. Täiskasvanute spontaanse teise keele omandamise esmast staadiumi iseloomustab nimisõnal põhinev lausungi organisatsioon ja verb/argument struktuuri puudumine (Klein & Perdue 1997; Perdue 1996), samuti staatilise ajalise osutuse prioriteet aspektuaalse (dünaamilise) perspektiivi ees (Benazzo 2006).
10. Maailma keeltes on rohkem nimisõnu kui verbe, ja rohkem produktiivset nimisõna- kui verbituletust (Gentner 1981; Gentner & Boroditsky 2001; vrd Jacobsen, 1979). See viitab samuti sellele, et nimi-sõnad on fundamentaalsemad, st võivad verbidele eelneda.
11. Maailma keeltes on vähemalt üks näide LP märkimisest LA markeril (TAM märkimine määratlejal tšamikuuro [*Chamicuro*] k – Nordlinger & Sadler 2004). Ma ei tea ühtki näidet LA märkimisest (DET või POSS) LP markeril (TAM või tegumood). See asümmeetria – LA markerite suurem leksikaalne iseseisvus LP markeritega võrreldes – vajab seletust. Võimalik seletus on, et LA markerid on üldiselt vanemad

kui LP markerid³⁷. Et markerid lekseemi modifitseerida saaks, peab viimane olema piisavalt konventsionaliseerunud (vrd punkt 8 ülal). Järeldus, et LA markerid on üldiselt vanemad kui LP markerid, lisab veelgi toetust hüpoteesile, et LA-d eelnevad LP-dele.

Tean ainult kaht argumenti, mis võiksid toetada LP-de evolutsioonilist eelne- mist LA-dele. Eeldades, et keele algne funktsioon oli jagada käske, võidakse väita, et esimesed sõnad pidid lähenema imperatiividele, st verbidele või fleksiiblitele (nt *Jookse!*, *Hüppa!* või *Püüa!*). Olen juba esitanud viis võrdlemisi otsest argumenti selle vastu (vt punktid 1, 3, 4, 6, 7 ülal), kuid lubage neile lisada veel kaks. Esiteks, kuna sõnu tõlgendatakse alati kontekstis, saab käsklusi anda ka teiste sõnaliikidega, nt nimisõnade (*Tiiger!* *Puu!*) või adverbidega (*Üles!*, *Kiiresti!*). Teiseks, keel pole käskluste andmiseks vajalik. Nt inimese puhul piisab potentsiaalsele ohule tähelepanu juhtimiseks artikuleerimata karjest. Reaktsioon sellele – nt võitlus või põgenemine – on instinktiivne ja/või õpitud käitumine, mis ei eelda mingit keelelist mediatsiooni. Vervetitel [*vervet monkeys*] on välja kujunenud häirekutsungite süsteem, mis eristab nelja röövlusmustrit (maod, röövlinnud, suured kaslased, primaadid), kutsudes iga mustri jaoks esile erineva põgenemistaktika. On isegi väidetud, et need kutsungid lähenevad sõnadele nagu *leopard* või *kotkas* (Zuberbühler et al. 1999). Kuigi viimane näib ebausutav, on siiski ilmne, et ka teised loomad peale inimese omavad kutsungeid, mis on funktsionaalselt ekvivalentsed võrdlemisi spetsiifiliste käsklustega. Teine argument, mis võiks viidata LP primaarsusele LA ees, on esitatud Hengeveldi (1992) poolt. Kuid see argument põhineb eeldusel, et kui mingis keeles on vähe nimisõnu, siis selles keeles nimisõnad puuduvad. Punktis i.iii olen selle arutluskäigu ebaloogilisusele juba tähelepanu juhtinud.

Väide, et N on evolutsiooniliselt fundamentaalsem kui V, pole uus. Varem on seda väidetud grammatiseerimisteoorias (Heine & Kuteva 2002, 2007). Varajast keelt rekonstrueerides pakuvad Heine ja Kuteva (2007), et algstaadiumis oli ainult üks leksikaalne kategooria, nimelt “nimisõna” (ajas stabiilne osutuslik üksus, mis väljendab peamiselt asjataolisi mõisteid). Käesolev töö nõustub sellega, lisades mitmeid uusi argumente selle seisukoha toetuseks. Nii palju kui ma tean, on ainult argument 9 ja üks mõte (verbideta keele võimalikkus) argumendist 1 varem esitatud, Heine ja Kuteva (2007) poolt. Eraldi võttes pole ükski neist üheteistkümnest argumendist piisav veenmaks kedagi LA evolutsioonilises eelnevuses LP-le, kuid koos võetuna on tõendid ülekaalukad.

³⁷ Kuna üksainus sõnajärje kitsendus võib märkida nii LA kui ka LP, ei saa eeldada, et LA markerid on tervikuna vanemad kui LP markerid.

Kokkuvõte

Käesolev doktoritöö tegeleb nimisõna/verb ja predikaat/argument struktuuridega. Esimene peatükk sai alguse küsimusest, kas nimisõna/verb struktuur on maailma keeltes universaalne; teine küsimusest, kas eksisteerib vastavus nimisõna/verb ja predikaat/argument struktuuride vahel. Mõlemad probleemid olid seni lahenduseta. Loodan, et teisele probleemile on käesolev töö suutnud vastata. Esimese probleemi puhul tuleb möönda, et olukorras, kus puuduvad isegi 'nimisõna' ja 'verbi' formaalsed ja keelteüleselt universaalsed definit-sioonid (mis ma küll alapeatükis 1.2 anda olen püüdnud), on keeleteadus tervikuna väga kaugel selle lahendamisest. Kolmas oluline teema oli N/V ja LP/A struktuuride evolutsiooniga seonduv. Kuna sellele on seni võrdlemisi vähe tähelepanu pööratud (ainsad kaalukad erandid on ilmselt Heine ja Kuteva 2002, 2007), õnnestus mul peatükis 3 öelda selle kohta nii mõndagi uut, mis küll eelmiste poolt väidetuga oluliselt vastuollu ei lähe. Kolmanda peatüki originaalseimaks panuseks on tõenäoliselt LP/A struktuuri evolutsiooni üldise käigu visandamine, ja mitmete uute toetavate argumentide lisamine Heine ja Kuteva hüpoteesile, et keele evolutsioonis eelnevad nimisõnad (ja üldisemalt LA-d) verbidele (LP-dele). Allpool annan detailsemad vastused olulisematele küsimustele, millega käesolev töö tegeles.

1. Kas nimisõna/verb eristus on maailma keeltes universaalne? Vastus: Teadmata. Küsimusele vastamine eeldab konsensust 'nimisõna' ja 'verbi' definit-sioonide, nii nagu ka neile aluseks olevate distributiivsete kriteeriumide osas. Tänapäeval sellist konsensust pole (vrd Baker 2003; Croft 2000, 2001; ja ptk 1). Isegi kui selline konsensus oleks saavutatav, eeldaks positiivne või negatiivne vastus deskriptiiv-tüpoloogilist uurimust, mis ületab kaugelt doktori- (ja elutöö) mahu.
2. Kas ranged, formaalsed, ja keelteüleselt universaalsed 'nimisõna', 'verbi' ja 'sõna' definit-sioonid on võimalikud? Vastus: Vähemalt osaliselt küll. Peatükis 1 annan ma 'nimisõna', 'verbi' ja 'sõna' ranged, keelteüleselt universaalsed, ja 'nimisõna' ja 'verbi' jaoks ka formaalsed (st grammatikaga seotud) definit-sioonid, mis on järgmised: N = sõnatüve omadus aktsepteerida LA, kuid mitte LP markereid; V = sõnatüve omadus aktsepteerida LP, kuid mitte LA markereid; F = sõnatüve omadus aktsepteerida nii LA kui ka LP markereid. LA ja LP markerid on defineeritud järgmiselt: LA on universaalselt märgitav LP/A sõnajärje kitsenduste, määratlejate ja possessiivide (alamhulga) poolt; LP on universaalselt märgitav LP/A sõnajärje kitsenduste, TAM ja tegumoe markerite (alamhulga) poolt. Alamhulga tingimuse rakendumine sõltub valitud distributiivsetest kriteeriumidest (vt punkt i.ii). Lihtsõna on defineeritud järgmiselt: minimaalne kõneüksus, mis on mõistetav (kuigi mitte tingimata kasutatav) väljaspool konteksti (vt alaptk 1.2).

3. Kas lingvistiline P/A struktuur on isomorfne matemaatilises loogikas kasutatava P/A struktuuriga? Vastus: Ei ole. LP/A struktuur on loogika ja taju P/A struktuuridest mitmeski mõttes erinev. Esiteks, LP/A struktuur on olemuslikult mitmetasandiline, sisaldades kuni kolm tasandit LA-sid ja LP-sid. Teiseks, spetsiaalsed reeglid (a)–(b) genereerivad kõrgemat järku LA-sid ja LP-sid ning konverteerivad LA-sid LP-deks ja vastupidi. Matemaatilise loogika predikaatstruktuur ei sisalda midagi sellist (kuigi kõiki neid reegleid on loodetavasti võimalik selles implementeerida – mis on pigem matemaatilise loogika kui keeleteaduse ülesanne, kuna tavakeelel pole vaja vastata kunstliku süsteemi poolt seatud tingimustele). Taju P/A struktuur on LP/A struktuuriga võrreldes märksa lihtsama ülesehitusega (vt punkt iii.i). Reeglid (a)–(b) on järgmised (notatsiooni vt joonealusest märkusest 31 ja lühenditest):

- | | |
|--|--|
| (a.1) $\{LA+\}COP+ADJ/LA \Rightarrow LP_2$ | (<i>He is rich/a sportsman.</i>) |
| (a.2) $\{LA+\}ADJ/LA \Rightarrow LP_2$ | (<i>On bogaty/sportsmen.</i> ‘He is rich/a sportsman’ in Russian) |
| (b.1) $INF/GER = LA_2$ | (<i>to see / seeing</i>) |
| (b.2) $\{LP+\}[LA+]AC \Rightarrow LA_3$ | (<i>He knows [the man] who was ill</i>) |

4. Kas nimisõna-argument ja verb-predikaat vastavused eksisteerivad? Vastus: Jah. Nimisõnad vastavad LA-dele ja verbid LP-dele. Vastavused nimisõnade ja FOPL/SOPL argumentide ning verbide ja FOPL/SOPL predikaatide vahel pole selle töö teema, kuid nende kehtestamine sõltub kõige rohkem reeglite (a)–(b) tõlkimisest FOPL-i või SOPL-i (mis ei kuulu praeguse töö, ega võibolla isegi mitte otseselt keeleteaduse ülesannete hulka).
5. Kas on võimalik midagi järeldada lingvistilise P/A struktuuri evolutsiooni kohta? Vastus: Jah. Heine ja Kuteva (2002, 2007) on varem esitanud argumente nimisõnade evolutsioonilisest eelnevusest verbi-dele. Peatükis 3 esitan mitmeid uusi argumente, mis toetavad hüpoteesi, et LA-d (nt nimisõnad) eelnevad evolutsiooniliselt LP-dele (nt verbidele). Lisaks visandan LP/A evolutsiooni üldise kulu, niivõrd kui seda on mulle teadaolevast infost võimalik tuletada.

Käesolev töö ei piirdu nende küsimuste ja vastustega. Mõned olulised tulemused saadakse sõnaliigi ‘fleksiiibel’ postuleerimisest tüvede jaoks, mis võivad täita nii nominaalseid kui ka verbaalseid funktsioone, nagu *love*, *hate*, *kill*, *jooks* jne. Koos kolme universaalse eeldusega (vt punkt i.iv) tingivad ‘nimisõna’, ‘verb’ ja ‘fleksiiibel’ järgmise loogiliselt võimalike keeletüüpide süsteemi: *N/V/F*, *N/F*, *V/F*, *N/V* and *F*. Vähemalt üks neist tüüpidest peab esinema maailma keeltes, kuid käesoleva töö ülesanne pole anda vastust küsimusele, millised neist tüüpidest (kui üldse) peale tüübi *N/V/F* reaalselt eksisteerivad. Samas olen püüdnud kitsendada võimalike tulevaste otsingute ala viidetega

konkreetsetele keeltele ja keelkondadele (vt alaptk 1.6), ja hõlbustada otsinguid tüploogilise omaduse hõlmavuse kriteeriumiga (vt punkt i.iii). On oluline märkida, et N-i, V ja 'sõna' definitsioonid ja peatükkides 2 ja 3 N/V, P/A ja LP/A struktuuride kohta saadud tulemused ei sõltu alamääratuse hüpoteesi (mis kehtestab sõnaliigi 'fleksibel') õigsusest.

REFERENCES

- Afarli, Tor A., Kristin M. Eide 2000: Subject requirement and predication. – *Nordic Journal of Linguistics*, 23, 1, 27–48.
- Alfieri, Luca 2007: The contribution of Proto-Indo-European to a part-of-speech typology. – Paper presented at the Workshop on languages with flexible parts-of-speech systems, 20. April, University of Amsterdam.
- Anderson, John M. 2004: Syntactic categories and notional features. – B. Aarts (ed), *Fuzzy grammar: a Reader*. Oxford; New York: Oxford University Press, 225–237.
- Anward, Jan 2001: Parts of speech. – M. Haspelmath, E. König, W. Oesterreicher, W. Raible (eds), *Language Typology and Language Universals. An International Handbook* (Vol. 1). Berlin; New York: Walter De Gruyter, 726–735.
- Arevalo, A., D. Perani, S. F. Cappa, A. Butler, E. Bates, N. Dronkers 2007: Action and object processing in aphasia: from nouns and verbs to the effect of manipulability. – *Brain and Language*, 100, 1, 79–94.
- Bach, Emmond 2004: Linguistic universals and particulars. – P. G. J. v. Sterkenburg (ed), *Linguistics today: facing a greater challenge*. Amsterdam; Philadelphia: John Benjamins, vi, 47–60.
- Baker, Mark C. 2003: *Lexical categories: verbs, nouns, and adjectives*. Cambridge; New York: Cambridge University Press.
- Bauer, Winifred 1993: *Maori*. London, New York: Routledge.
- Beck, David 2002: The typology of parts of speech systems: the markedness of adjectives. New York; London: Routledge.
- Becker, Misha 2004: Is isn't be. – *Lingua*, 114, 4, 495–516.
- Belardinelli, Marta Olivetti, Rosalia Di Matteo, Cosimo Del Gratta, Andrea De Nicola, Antonio Ferretti, Armando Tartaro, Lorenzo Bonomo, Gian Luca Romani 2004: Intermodal sensory image generation: an fMRI analysis. – *The European Journal of Cognitive Psychology*, 16, 5, 729–752.
- Benazzo, Sandra 2006: The expression of temporality in early second language varieties and adult home signs. – Paper presented at the NIAS workshop, May 19–20.
- Bennett, Paul 1995: *A Course in Generalized Phrase Structure Grammar*. London: UCL.
- Bensafi, M., J. Porter, S. Pouliot, J. Mainland, B. Johnson, C. Zelano, N. Young, E. Bremner, D. Aframian, R. Khan, N. Sobel 2003: Olfactomotor activity during imagery mimics that during perception. – *Nature Neuroscience*, 6, 11, 1142–1144.
- Bhat, D.N.S. 2000: Word classes and sentential functions. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 47–63.
- Bickerton, Derek 1990: *Language and Species*. Chicago: University of Chicago Press.
- Bickerton, Derek 1998: Catastrophic evolution: The case for a single step from protolanguage to full human language. – J. R. Hurford, M. Studdert-Kennedy, C. Knight (eds), *Approaches to the Evolution of Language: Social and Cognitive Bases*. Cambridge: Cambridge University Press, 341–358.
- Bickerton, Derek 2000: How protolanguage became language. – J. R. H. Chris Knight, M. Studdert-Kennedy (eds), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press.
- Bickerton, Derek 2002: Foraging versus social intelligence in the evolution of protolanguage. – A. Wray (ed), *The Transition to Language*. Oxford: Oxford University Press.

- Bickerton, Derek 2003: Symbol and structure: a comprehensive framework for language evolution. – M. H. Christiansen, S. Kirby (eds), *Language Evolution: The States of the Art*. Oxford: Oxford University Press, 77–93.
- Bickerton, Derek 2007: Language evolution: A brief guide for linguists. – *Lingua*, 117, 3, 510–526.
- Bierwisch, Manfred 2001: Repertoires of Primitive Elements: Prerequisite or result of acquisition? – J. Weissenborn, B. Höhle (eds), *Approaches to Bootstrapping* (Vol. 2). Amsterdam: John Benjamins, 281–307.
- Biggs, Bruce 1971: Languages of Polynesia. – *Current Trends in Linguistics*, 8, 1, 466–505.
- Blake, Barry J. 1994: *Case*. Cambridge [England]; New York, NY, USA: Cambridge University Press.
- Bloomfield, Leonard 1942: Outline of Ilocano syntax. – *Language*, 18, 193–200.
- Boas, Franz 1947: Kwakiutl Grammar with a Glossary of the Suffixes. – *Transactions of the American Philosophical Society*, 37, 3, 201–377.
- Borer, Hagit 2003: Exo-skeletal vs. endo-skeletal explanations. – J. Moore, M. Polinsky (eds), *The Nature of Explanation in Linguistic Theory*. Chicago: CSLI, 31–67.
- Borer, Hagit 2005a: Structuring Sense. Volume I: In Name Only. Oxford: Oxford University Press.
- Borer, Hagit 2005b: Structuring Sense. Volume II: The Normal Course of Events. Oxford: Oxford University Press.
- Borges, Jorge Luis 1964: *Labyrinths: Selected Stories and Other Writings*. New York: New Directions Pub. Co.
- Bošković, Željko 2007: What will you have, DP or NP? – Paper presented at the NELS 37, Urbana-Champaign.
- Bowers, John 1993: The syntax of predication. – *Linguistic Inquiry*, 24, 591–656.
- Bowie, Jill 2008: Proto-discourse and the emergence of compositionality. – *Interaction Studies*, 9, 1, 18–33.
- Broschart, Jürgen 1997: Why Tongan does it differently: Categorical distinctions in a language without nouns and verbs. – *Linguistic Typology*, 1, 2, 123–166.
- Bruening, Benjamin 2008: Selectional asymmetries between CP and DP suggest that the DP hypothesis is wrong. – Paper presented at the The 32nd Penn Linguistics Colloquium, University of Pennsylvania.
- Budd, Graham E. 2006: On the origin and evolution of major morphological characters. – *Biological Reviews of the Cambridge Philosophical Society*, 81, 4, 609–628.
- Carstairs-McCarthy, Andrew 1999: *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth*. Oxford: Oxford University Press.
- Chierchia, Gennaro 1998: Reference to kinds across languages. – *Natural Language Semantics*, 6, 339–405.
- Chomsky, Noam 1981: *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam 1995: *The Minimalist Program*. Cambridge, MA: MIT Press.
- Collina, Simona, Paola Marangolo, Patrizia Tabossi 2001: The role of argument structure in the production of nouns and verbs. – *Neuropsychologia*, 39, 11, 1125–1137.
- Cooper, Lynn A. 1995: Varieties of visual representation: How are we to analyze the concept of mental image? – *Neuropsychologia*, 33, 1575–1582.
- Crepaldi, D., S. Aggijuro, L. S. Arduino, G. Zonca, G. Ghirardi, M. G. Inzaghi, M. Colombo, G. Chierchia, C. Luzzatti 2006: Noun-verb dissociation in aphasia: the

- role of imageability and functional locus of the lesion. – *Neuropsychologia*, 44, 1, 73–89.
- Croft, William 1990: *Typology and Universals*. Cambridge [England]; New York: Cambridge University Press.
- Croft, William 2000: Parts of speech as language universals and as language-particular categories. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 65–102.
- Croft, William 2001: *Radical construction grammar: syntactic theory in typological perspective*. Oxford, etc.: Oxford University Press.
- Croft, William 2005: Word classes, parts of speech, and syntactic argumentation. – *Linguistic Typology*, 40, 3, 431–441.
- Crystal, David 2004: English Word Classes. – B. Aarts (ed), *Fuzzy grammar: a Reader*. Oxford; New York: Oxford University Press, 191–211.
- Czaykowska-Higgins, Ewa, M. Dale Kinkade 1998: Salish languages and linguistics. – E. Czaykowska-Higgins, M. D. Kinkade (eds), *Salish Languages and Linguistics: Theoretical and Descriptive Perspectives*. The Hague: Mouton, 1–68.
- Dessalles, Jean-Louis 2008: From metonymy to syntax in the communication of events. – *Interaction Studies*, 9, 1, 51–65.
- Di Sciullo, Anna Maria, Edwin Williams 1987: *On the definition of word* (Vol. 14). Cambridge, MA: MIT Press.
- Dixon, Robert M. W. 2004: Adjective classes in typological perspective. – R. M. W. Dixon, A. Y. Aikhenvald (eds), *Adjective Classes: A Cross-linguistic Typology*. Oxford: Oxford University Press, 1–49.
- Dixon, Robert M. W. 2006: Complement clauses and complementation strategies in typological perspective. – R. M. W. Dixon, A. Y. Aikhenvald (eds), *Complementation: a cross-linguistic typology*. Oxford: Oxford University Press, 1–48.
- Dixon, Robert M. W., Alexandra Y. Aikhenvald 2002: Word: a typological framework. – R. M. W. Dixon, A. Y. Aikhenvald (eds), *Word: A Cross-linguistic Typology*. Cambridge, etc.: Cambridge University Press, 1–41.
- Don, Jan, Eva van Lier 2007: On the lack of word-class distinction: flexibility and conversion. – Paper presented at the Workshop on languages with flexible parts-of-speech systems, April 20, University of Amsterdam.
- Eide, Kristin M., Tor A. Afarli 1999: The syntactic disguises of the predication operator. – *Studia Linguistica*, 53, 2, 155–181.
- Enfield, Nick J. 2004: Adjectives in Lao. – R. M. W. Dixon, A. Y. Aikhenvald (eds), *Adjective Classes: A Cross-linguistic Typology*. Oxford: Oxford University Press, 323–347.
- Erelt, Mati, Tiit Erelt, Kristiina Ross 2000: *Eesti Keele Käsiraamat*. Tallinn: Eesti Keele Sihtasutus.
- Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare 1995: *Eesti Keele Grammatika: Morfoloogia; Sõnamoodustus* (Vol. I). Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Evans, Nicholas, Toshiaki Osada 2005: Mundari: The myth of a language without word classes. – *Linguistic Typology*, 40, 3, 351–390.
- Farrell, Patrick 2001: Functional shift as category underspecification. – *English Language and Linguistics*, 5, 1, 109–130.
- Fisher, Cynthia 2002: The role of abstract syntactic knowledge in language acquisition: a reply to Tomasello (2000). – *Cognition*, 82, 3, 259–278.
- Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, Ivan Sag 1985: *Generalized phrase structure grammar*. Oxford, UK: B. Blackwell.

- Gentner, Dedre 1981: Some interesting differences between nouns and verbs. – *Cognition and Brain Theory*, 4, 161–178.
- Gentner, Dedre, Lera Boroditsky 2001: Individuation, relativity, and early word learning. – M. Bowerman, S. Levinson (eds), *Language Acquisition and Conceptual Development*. New York: Cambridge University Press, 215–256.
- Gil, David 1994: The structure of Riau Indonesian. – *Nordic Journal of Linguistics*, 17, 179–200.
- Gil, David 2000: Syntactic categories, cross-linguistic variation and universal grammar. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 173–216.
- Gil, David 2007: On the position of Riau Indonesian in a typology of “Flexible-Syntactic-Category” languages. – Paper presented at the Workshop on languages with flexible parts-of-speech systems, April 20, University of Amsterdam.
- Givón, Talmy 2001: *Syntax: an introduction* (rev. ed.). Amsterdam; Philadelphia: John Benjamins.
- Gleitman, Lila R. 1993: A Human Universal: The Capacity to Learn a Language – *Modern Philology*, 90, Supplement, S13–S33.
- Gopnik, Alison 2000: *Theories, language and culture: Whorf without wincing*. New York: Cambridge University Press.
- Graffi, Giorgi 2001: The emphasis on syntax in the early phase of European structuralism: Ries, Jespersen, Mathesius, Guillaume, Tesnière. – S. Auroux, E. F. K. Koerner, H.-J. Niederehe, K. Versteegh (eds), *History of the Language Sciences / Geschichte der Sprachwissenschaften / Histoire des sciences du langage* (Vol. 2). Berlin - New York: De Gruyter, 1838–1847.
- Greenberg, Joseph Harold 1963: *Essays in linguistics*. Chicago; London: Phoenix books.
- Hajek, John 2004: Adjective classes: What can we conclude? – R. M. W. Dixon, A. Y. Aikhenvald (eds), *Adjective Classes: A Cross-linguistic Typology*. Oxford: Oxford University Press, 348–362.
- Haspelmath, Martin 2007: Pre-established categories don't exist: Consequences for language description and typology. – *Linguistic Typology*, 11, 119–132.
- Hauser, Marc D. 1997: *The Evolution of Communication*. Cambridge, MA: MIT Press/BradfordBooks.
- Hauser, Marc D., Noam Chomsky, W. Tecumseh Fitch 2002: The faculty of language: what is it, who has it, and how did it evolve? – *Science*, 298, 5598, 1569–1579.
- Heine, Bernd 1997: *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.
- Heine, Bernd, Tania Kuteva 2002: On the evolution of grammatical forms. – A. Wray (ed), *The Transition to Language*. Oxford: Oxford University Press, 376–397.
- Heine, Bernd, Tania Kuteva 2007: *The genesis of grammar: a reconstruction*. New York: Oxford University Press.
- Helmbrecht, Johannes 2001: Head-marking vs. dependent-marking languages. – M. Haspelmath, E. König, W. Oesterreicher, W. Raible (eds), *Language Typology and Language Universals. An International Handbook* (Vol. 2). Berlin; New York: Walter De Gruyter, 1424–1432.
- Hengeveld, Kees 1992: *Non-verbal predication: theory, typology, diachrony*. Berlin; New York: Mouton de Gruyter.
- Hengeveld, Kees, Jan Rijkhoff 2005: Mundari as a flexible language. – *Linguistic Typology*, 40, 3, 406–431.

- Hengeveld, Kees, Jan Rijkhoff, Anna Siewerska 2004: Part-of-speech systems and word order. – *Journal of Linguistics*, 40, 3, 527–570.
- Herman, L. M., D. G. Richards, J. P. Wolz 1984: Comprehension of sentences by bottlenosed dolphins. – *Cognition*, 16, 2, 129–219.
- Himmelmann, Nikolaus P. 2007: Lexical categories and voice in Tagalog. – P. Austin, S. Musgrave (eds), *Voice and Grammatical Functions in Austronesian Languages*. Stanford: CSLI.
- Hockett, Charles 1966: The problem of universals in language. – J. H. Greenberg (ed), *Universals of Language* (2nd ed.). Cambridge, Mass.: M.I.T. Press, xxvii, 1–22.
- Hopper, Paul J., Sandra A. Thompson 1984: The discourse basis for lexical categories in universal grammar. – *Language*, 60, 703–752.
- Hummel, Kenneth E. 2000: *Introductory Concepts for Abstract Mathematics*. Boca Raton, FL: Chapman & Hall/CRC.
- Hurford, James R. 2006a: Recent Developments in the Evolution of Language. – *Cognitive Systems*, 7, 1, 23–32.
- Hurford, James R. 2000a: The emergence of syntax. – J. R. H. Chris Knight, M. Studdert-Kennedy (eds), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press, 219–230.
- Hurford, James R. 2000b: Social transmission favours linguistic generalization. – J. R. H. Chris Knight, M. Studdert-Kennedy (eds), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge: Cambridge University Press, 324–352.
- Hurford, James R. 2001: Protothought had no logical names. – *New Essays on the Origins of Language*, 117–130.
- Hurford, James R. 2003a: The Language Mosaic and its Evolution. – M. H. Christiansen, S. Kirby (eds), *Language Evolution: The States of the Art*. Oxford: Oxford University Press, 38–57.
- Hurford, James R. 2003b: The neural basis of predicate-argument structure. – *Behavioral and Brain Sciences*, 26, 3, 261–283; Discussion 283–316.
- Hurford, James R. 2003c: Ventral/dorsal, predicate/argument: the transformation from perception to meaning. – *Behavioral and Brain Sciences*, 26, 3, 301–311.
- Hurford, James R. 2006b: Proto-propositions. – *Proceedings of the 6th International Conference on the Evolution of Language*, 131–138.
- Hurford, James R. 2007: The origin of noun phrases: reference, truth, and communication. – *Lingua*, 117, 3, 527–542.
- Itkonen, Esa 2000: Noun vs. Verb? – Paper presented at the Parts of speech in and across languages, August 17–19, Symposium presentation; University of Helsinki.
- Jackendoff, Ray 1999: Possible stages in the evolution of the language capacity. – *Trends in Cognitive Sciences*, 3, 7, 272–279.
- Jackendoff, Ray, Steven Pinker 2005: The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). – *Cognition*, 97, 2, 211–225.
- Jacobsen, William H. 1979: Noun and verb in Nootkan. – B. Efrat (ed), *The Victorian Conference on Northwestern Languages*. Victoria: British Columbia Provincial Museum, 83–153.
- Jelinek, Eloise, Richard A. Demers 1994: Predicates and pronominal arguments in Straits Salish. – *Language*, 70, 697–736.
- Jespersen, Otto 1924: *The philosophy of grammar*. New York: H. Holt and company.

- Johansson, Sverker 2006: Working backwards from modern language to proto-grammar. – Proceedings of the 6th International Conference on the Evolution of Language, 160–167.
- Junker, Marie-Odile 2003: A Native American view of the ‘mind’ as seen in the lexicon of cognition in East Cree. – *Cognitive Linguistics*, 14, 2–3, 167–194.
- Karlsson, Fred 2007a: Constraints on multiple center-embedding of clauses. – *Journal of Linguistics*, 43, 2, 365–392.
- Karlsson, Fred 2007b: Constraints on multiple initial embedding of clauses. – *International Journal of Corpus Linguistics*, 12, 1, 107–118.
- Kayne, Richard 1981: ECP Extensions. – *Linguistic Inquiry*, 12, 93–133.
- Kim, Mikyong, Cynthia K. Thompson 2000: Patterns of comprehension and production of nouns and verbs in agrammatism: implications for lexical organization. – *Brain and Language*, 74, 1, 1–25.
- Kinkade, M. Dale 1963: Phonology and morphology of Upper Chehalis. – *International Journal of American Linguistics*, 29, 345–356.
- Klamer, Marian 1998: *A Grammar of Kambera*. Berlin/New York: Mouton de Gruyter.
- Klamer, Marian 2005: Kambera. – K. A. Adelaar, N. P. Himmelmann (eds), *The Austronesian Languages of Asia and Madagascar*. London: Routledge, 709–734.
- Klatzky, R. L., S. J. Lederman, D. E. Matula 1991: Imagined haptic exploration in judgments of object properties. – *Journal of Experimental Psychology: Learning Memory and Cognition*, 17, 2, 314–322.
- Klein, Wolfgang, Clive Perdue 1997: The Basic Variety. – *Second Language Research*, 13, 4, 301–347.
- Lambert, Silke 1998: Partial overlap of nominal and verbal linkers: the case of Tongan. – Paper presented at the International Conference of the SFB ‘Theory of the Lexicon’, August 17–19, Wuppertal.
- Langacker, Ronald W. 2004: Bounded regions. – B. Aarts (ed), *Fuzzy grammar: a Reader*. Oxford; New York: Oxford University Press, 239–245.
- Laudanna, Alessandro, Miriam Voghera 2002: Nouns and verbs as grammatical classes in the lexicon. – *Rivista di Linguistica*, 14, 1, 9–26.
- Lulu, Wang, Liu Haitao 2007: A Description of Chinese NPs using Head-Driven Phrase Structure Grammar. – Paper presented at the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford.
- Luzzatti, C., S. Aggujaro, D. Crepaldi 2006: Verb-noun double dissociation in aphasia: theoretical and neuroanatomical foundations. – *Cortex*, 42, 6, 875–883.
- Luzzatti, C., R. Raggi, G. Zonca, C. Pistarini, A. Contardi, G. D. Pinna 2002: Verb-noun double dissociation in aphasic lexical impairments: the role of word frequency and imageability. – *Brain and Language*, 81, 1–3, 432–444.
- Luuk, Erkki in press-a: The noun/verb and predicate/argument structures. – *Lingua*.
- Luuk, Erkki in press-b: Nouns, verbs and flexibles: implications for typologies of word classes. – *Language Sciences*.
- Lyons, John 2004: A notional approach to the parts of speech. – B. Aarts (ed), *Fuzzy grammar: a Reader*. Oxford; New York: Oxford University Press, 213–223.
- Marantz, Alec 1980: English S is the maximal projection of V. – Paper presented at the NELS X.
- Marantz, Alec 1984: *On the nature of grammatical relations*. Cambridge, MA: The MIT Press.
- Marantz, Alec 1997: No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. – Paper presented at the 21st Annual Penn Linguistics Colloquium, Philadelphia.

- Massam, Diane 2005: Lexical categories in Niuean. – A. Carnie, H. Harley, S. A. Dooley (eds), *Verb First: On the Syntax of Verb-initial Languages* (Vol. 73). Amsterdam/Philadelphia: John Benjamins, 227–241.
- Mauck, Björn, Guido Dehnhardt 1997: Mental rotation in a California sea lion (*Zalophus californianus*). – *Journal of Experimental Biology*, 200, Pt 9, 1309–1316.
- McCawley, James D. 1992: Justifying part-of-speech assignments in Mandarin Chinese. – *Journal of Chinese Linguistics*, 20, 2, 211–246.
- Mithun, Marianne 1999: *The languages of native North America*. Cambridge, UK; New York: Cambridge University Press.
- Mithun, Marianne 2000: Noun and verb in Iroquoian languages: multicategorisation from multiple criteria. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 397–420.
- Mondini, S., C. Luzzatti, G. Zonca, C. Pistarini, C. Semenza 2004: The mental representation of verb-noun compounds in Italian: evidence from a multiple single-case study in aphasia. – *Brain and Language*, 90, 1–3, 470–477.
- Mosel, Ulrike, Even Hovdhaugen 1992: *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Neukom, Lukas 2001: *Santali*. München: Lincom Europa.
- Newmeyer, Frederick J. 2003: What can the field of linguistics tell us about the evolution of language? – M. H. Christiansen, S. Kirby (eds), *Language Evolution: The States of the Art*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2004: Cognitive and functional factors in the evolution of grammar. – *European Review*, 12, 2, 245–264.
- Nichols, Johanna 1986: Head-marking and dependent-marking grammar. – *Language*, 62, 1, 56–119.
- Nordlinger, Rachel, Louisa Sadler 2004: Tense beyond the verb: encoding clausal tense/aspect/mood on nominal dependents. – *Natural Language & Linguistic Theory*, 22, 597–641.
- Nowak, Martin A., David C. Krakauer 1999: The evolution of language. – *Proceedings of the National Academy of Sciences of the United States of America*, 96, 14, 8028–8033.
- Oxford 1991: *The Oxford English Dictionary* (2nd ed. Vol. XV Ser-Soosy). Oxford: Clarendon Press.
- Partee, Barbara H. 2002: Noun phrase interpretation and type-shifting principles. – P. Portner, B. H. Partee (eds), *Formal Semantics: The Essential Readings*. Chicago: Blackwell Publishing.
- Pawley, Andrew 1966: Samoan phrase structure: Morphology-syntax of a Western Polynesian language. – *Anthropological Linguistics*, 8, 1–63.
- Perdue, Clive 1996: Pre-basic varieties: the first stages of second language acquisition. – *Toegepaste Taalwetenschap in Artikelen*, 55, 135–150.
- Peterson, John 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, John 2007: Languages without nouns and verbs? An alternative to lexical classes in Kharia. – C. P. Masica (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*. Delhi: Motilal Banarsidass, 274–303.
- Polinsky, Maria, Ezra van Everbroeck 2003: Development of gender classifications: Modeling the historical change from Latin to French. – *Language*, 79, 2, 356–390.

- Progovac, Liliana 1998: Determiner phrase in a language without determiners. – *Journal of Linguistics*, 34, 165–179.
- Putseys, Yvan 1989: Sentential complementation and the meanings of *such*. – D. Jaspers, W. G. Klooster, Y. Putseys, P. A. M. Seuren (eds), *Sentential Complementation and the Lexicon*. Dordrecht: Foris.
- Pylyshyn, Zenon W. 2002: Mental imagery: in search of a theory. – *Behavioral and Brain Sciences*, 25, 2, 157–182; Discussion 182–237.
- Pylyshyn, Zenon W. 2003: *Seeing and visualizing: It's not what you think*. Cambridge, Mass.; London: MIT Press.
- Ramat, Paolo 1999: Linguistic categories and linguists' categorizations. – *Linguistics*, 37, 1, 157–180.
- Reisberg, D., M. Wilson, J.D. Smith 1991: Auditory imagery and inner speech. – R. H. Logie, M. Denis (eds), *Mental images in human cognition*. Amsterdam: Elsevier Science Publishers B.V., 59–81.
- Rijkhoff, Jan 2000: When can a language have adjectives? An implicational universal. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 217–257.
- Rijkhoff, Jan 2002: Verbs and nouns from a cross-linguistic perspective. – *Rivista di Linguistica*, 14, 1, 115–147.
- Rijkhoff, Jan 2003: When can a language have nouns and verbs? – *Acta Linguistica Hafniensa*, 35, 7–38.
- Ross, John Robert 1972: The category squish: endstation hauptwort. – Paper presented at the Eighth Regional Meeting, Chicago Linguistic Society, Chicago.
- Sapir, Edward 1978 [1921]: *Language: an introduction to the study of speech*. London: Hart-Davis MacGibbon.
- Sasse, Hans-Jürgen 1993a: Das nomen — eine universale kategorie? – *Sprachtypologie und Universalienforschung*, 46, 187–221.
- Sasse, Hans-Jürgen 1993b: Syntactic categories and subcategories. – J. Jacobs, A. von Stechow, W. Sternefeld, T. Vennemann (eds), *Syntax: an international handbook*. Berlin: Mouton de Gruyter, 646–686.
- Sasse, Hans-Jürgen 2001: Scales between nouniness and verbiness. – M. Haspelmath, E. König, W. Oesterreicher, W. Raible (eds), *Language Typology and Language Universals. An International Handbook* (Vol. 1). Berlin; New York: Walter De Gruyter, 495–509.
- Savage-Rumbaugh, S, SG Shanker, TJ Taylor 1998: *Apes, Language, and the Human Mind*. New York: Oxford University Press.
- Stassen, Leon 2008: Predicative adjectives. – M. Haspelmath, M. S. Dryer, D. Gil, B. Comrie (eds), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 118.
- Steels, Luc, F. Kaplan, A. McIntyre, J. Van Looveren 2002: Crucial factors in the origins of word-meaning. – A. Wray (ed), *The Transition to Language*. Oxford: Oxford University Press.
- Swadesh, Morris 1939: Nootka internal syntax. – *International Journal of American Linguistics*, 9, 77–102.
- Tallerman, Maggie 2007: Did our ancestors speak a holistic protolanguage? – *Lingua*, 117, 579–604.
- Vauclair, Jacques, Joel Fagot, William D Hopkins 1993: Rotation of mental images in baboons when the visual input is directed to the left cerebral hemisphere. – *Psychological Science*, 4, 2, 99–103.

- Waxman, Sandra R, Jeffrey Lidz 2006: Early word learning. – D. Kuhn, R. Siegler (eds), *Handbook of Child Psychology* (6th ed., Vol. 2). Hoboken, NJ: Wiley, 299–335.
- Webster's 1988: *Webster's Dictionary*. Larchmont: Book Essentials Publications.
- Wehrli, Eric 1988: Parsing with a GB-grammar. – U. Reyle, C. Rohrer (eds), *Natural Language Parsing and Linguistic Theories*. Dordrecht: D. Reidel, 177–201.
- Whorf, Benjamin L. 1945: Grammatical categories. – *Language*, 21, 1, 1–11.
- Wierzbicka, Anna 2000: Lexical prototypes as a universal basis. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 285–317.
- Vogel, Petra M. 2000: Grammaticalisation and part-of-speech systems. – P. M. Vogel, B. Comrie (eds), *Approaches to the Typology of Word Classes*. Berlin; New York: Mouton de Gruyter, 259–284.
- Wray, Alison 2000: Holistic utterances in protolanguage: The link from primates to humans. – J. R. H. Chris Knight, M. Studdert-Kennedy (eds), *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*. Cambridge, UK: Cambridge University Press, 285–302.
- Zuberbuhler, Klaus, Dorothy L Cheney, Robert M Seyfarth 1999: Conceptual semantics in a nonhuman primate. – *Journal of Comparative Psychology*, 113, 1, 33–42.

CURRICULUM VITAE

Erkki Luuk

Date and place of birth: 14th February 1971, Paide
Nationality: Estonian
Address: Postimaja pk 149, Tartu 50002, Estonia
Telephone: +372 5806 8518
Email: erkkil@gmail.com

Education

2001 University of Tartu, MA in semiotics and culturology
1994 University of Tartu, diploma (=MA) in sociology
1989 Tartu Secondary School no. 2

Professional experience

2009– University of Tartu, Institute of Computer Science, semantics specialist
2006–2007 University of Tartu, Graduate School of Linguistics and Language Technology, researcher (part-time)
1998–2007 Tartu Higher Art School, lecturer (part-time)
1995–1997 University of Tartu, Department of Psychology, programmer (part-time)

ELULOOKIRJELDUS

Erkki Luuk

Sünniaeg ja -koht: 14. veebruar 1971, Paide
Kodakondsus: eesti
Aadress: Postimaja pk 149, Tartu 50002, Estonia
Telefon: +372 5806 8518
E-post: erkkil@gmail.com

Haridus

2001 Tartu Ülikool, MA semiootikas ja kultuuroloogias
1994 Tartu Ülikool, diplom (=MA) sotsioloogias
1989 Tartu 2. Keskkool

Teenistuskäik

2009– Tartu Ülikool, Arvutiteaduse instituut, semantika spetsialist
2006–2007 Tartu Ülikool, Keeleteaduse ja -tehnoloogia doktorikool, teadur (osaline k.)
1998–2007 Tartu Kõrgem Kunstikool, lektor (osaline k.)
1995–1997 Tartu Ülikool, Psühholoogia osakond, programmeerija (osaline k.)

DISSERTATIONES LINGUISTICAE UNIVERSITATIS TARTUENSIS

1. **Anna Verschik.** Estonian yiddish and its contacts with coterritorial languages. Tartu, 2000, 196 p.
2. **Silvi Tenjes.** Nonverbal means as regulators in communication: socio-cultural perspectives. Tartu, 2001, 214 p.
3. **Ilona Tragel.** Eesti keele tuumverbid. Tartu, 2003, 196 p.
4. **Einar Meister.** Promoting Estonian speech technology: from resources to prototypes. Tartu, 2003, 217 p.
5. **Ene Vainik.** Lexical knowledge of emotions: the structure, variability and semantics of the Estonian emotion vocabulary. Tartu, 2004, 166 p.
6. **Heili Orav.** Isiksuseomaduste sõnavara semantika eesti keeles. Tartu, 2006, 175 p.
7. **Larissa Degel.** Intellektuaalsfäär intellektuaalseid võimeid tähistavate sõnade kasutuse põhjal eesti ja vene keeles. Tartu, 2007, 225 p.
8. **Meelis Mihkla.** Kõne ajalise struktuuri modelleerimine eestikeelsele tekst-kõne sünteesile. Modelling the temporal stucture of speech for the Estonian text-to-speech synthesis. Tartu, 2007, 176 p.
9. **Mari Uusküla.** Basic colour terms in Finno-Ugric and Slavonic languages: myths and facts. Tartu, 2008, 207 p.
10. **Petar Kehayov.** An Areal-Typological Perspective to Evidentiality: the Cases of the Balkan and Baltic Linguistic Areas. Tartu, 2008, 201 p.
11. **Ann Veismann.** Eesti keele kaas- ja määrsõnade semantika võimalusi. Tartu, 2009, 145 p.