Typological variation of the adjectival class:
Markedness and iconicity at the semantics ⇔ syntax interface[*]

David Beck
University of Michigan, Ann Arbor

1. Adjectives and parts of speech systems

Ask a layperson what they know about grammar and you are likely to get an answer that has something to do with parts of speech; ask a linguist what they know about parts of speech and the answer is quite likely to be much less enlightening. Parts of speech systems or, as I will refer to them here, lexical classes are among the most frequently overlooked aspects of linguistic analysis, yet they are at the same time among the most fundamental elements of language. Lexical classes play a key role in most, if not all, syntactic theories, they are the cornerstones of lexicography, and they are crucial elements in morphological analysis—yet precise and rigorous definitions of these classes have never been successfully formulated. More often than not, lexical classes are treated merely as primitives, either in terms of input to rules, determinants of underlying phrase structure, governors of inflectional patterns, or as sources of valency and subcategorization frames. Thus, class-designations such as "verb", "adjective", and "noun" are the linchpins of semantic, syntactic, and morphological structure, but the terms themselves are rarely defined and their properties, both formal and functional, are often taken for granted. While it is certainly possible to carry out linguistic analysis without a clear definition of the basic units involved (as long as the identity of these units can be agreed upon), any theory which proceeds without a full understanding of its own primitives rests on uncertain foundations.

Most current linguistic theories—whose main proponents are speakers of and researchers in European languages—are modeled on languages with parts-of-speech systems organized into the four major classes of verb, noun, adjective, and adverb. Cross-linguistic investigation shows that not all languages fit this pattern: while nouns and verbs appear to be essentially universal, languages that have few or no adjectives are a typological commonplace.[1] Even languages that do have adjectives vary tremendously in terms of their absolute numbers and the sets of meanings singled out by language-specific definitions of an adjectival class. Any credible attempt to define parts of speech must come to terms with both of these types of variation, and to date most proposals for a criterial definition of adjective have failed on both counts. As I will argue in this paper, traditional approaches to defining lexical classes in general (and adjectives in particular) have confined themselves to one of the semantic (Section 1.1), morphological (Section 1.2), or syntactic (Section 1.3) levels of analysis and, as a result, have been unable to account for cross-linguistic variation in parts-of-speech systems. More recent approaches (discussed in Section 1.4) have tried to deal with this variation in terms of prototypicality and markedness (*e.g.* Hopper & Thompson 1984; Hengeveld 1992a, 1992b) and the mapping between semantic and functional classes (Croft 1991). While this work again falls short of allowing us to formulate a criterial, cross-linguistically viable definition of adjective, it does point us in the direction of the proposal sketched out in Section 2, which posits that a criterial definition of adjective be based on the mapping between semantic and syntactic representation and the notion of iconicity (Section 2.1). In this approach, parts of speech are treated as a feature of the lexicon, which mediates between semantic and (deep) syntactic structure, and are treated as serving primarily as input to rules used for building syntactic structure (*cf.* Jesperson 1924), each class having its own unmarked syntactic role (Section 2.2); differences in lexical class may subsequently have consequences for the

morphological realization of lexical items. Adjectives are a marked class in that their unmarked syntactic role, that of modifier of a noun, is inherently non-iconic (Section 2.3). These observations allow us to formulate a definition of adjective that at once accounts for the attested typological variation and at the same time properly constrains it, casting new light on some proposals for parts-of-speech typology current in the literature (Section 3).

1.1. Semantic characterizations

The most familiar and intuitively appealing of the traditional approaches to defining parts of speech is the semantic characterization, which groups words as nouns, verbs, or adjectives based on their meanings. Semantically, nouns and verbs are highly consistent across languages and roughly conform to traditional semantic characterizations in which nouns are said to be those lexical items denoting "people, places, and things" (here, this class will be referred to loosely as semantic names—*cf*. also Wierzbicka's 1988 term *kind*) and verbs are those which denote "actions and states" (loosely, semantic predicates). Similarly, adjectives are often characterized as those words which denote properties, qualities, and (in some languages) states; Thompson (1988, based on the work of Dixon 1982) terms the prototypical semantic domain of adjectives, "property concepts"—words expressing one of the following categories:

(1) Classes of property concepts (with English examples)

DIMENSION — *big, little, long, wide …*
PHYSICAL PROPERTIES — *hard, heavy, smooth …*
COLOUR
HUMAN PROPENSITY — *jealous, happy, clever, generous, proud …*
AGE — *new, young, old …*
VALUE — *good, bad, pure, delicious …*
SPEED — *fast, slow, quick …*

(Thompson 1988: 168)

In order for such a semantic characterization to serve as the sole basis for a definition of a part of speech (*e.g. adjective* = a word expressing a property concept), it would have to be the case that all words conforming to this semantic characterization be treated as members of this class by the syntactic rules of a given language. Property-concept words, however, show a great deal of intralinguistic and cross-linguistic variation in their syntactic behaviour and it is not always possible to predict their lexical classification in a given language from their meaning. For instance, in Hausa, the DIMENSION word 'wide' is expressed as a noun *faɗi*, though it is clearly a property concept (Wetzer 1996: 178), whereas in Bemba the HUMAN PROPENSITIES 'strong', 'brave', and 'wise' are expressed by the verbs *ashipa*, *akosa*, and *aceenjela* respectively (Schachter 1985). The PHYSICAL PROPERTY 'hard' surfaces as a noun, *ṣauɓi..*, in Hausa, as an adjective, *duro*, in Spanish, and as a verb, *x̌tadis*, in Lushootseed in spite of the fact that they are all property-concept words and, hence, by a naïve semantic definition, should be adjectives.

The problem of formulating a predictive semantic definition of adjectives is particularly obvious in languages that have a small class of adjectives and divide the remaining property-concept words between nouns and verbs. Such languages may have as few as half a dozen true adjectives (*i.e.* unmarked modifiers of nouns) and, while all of these are typically expressions of property concepts, other members of the same semantic class are not realized as adjectives but as

nouns or verbs. In the Bantu language Venda, for instance, there are only twenty adjectives, listed in (2):

(2)   Venda

hulu 'big, large'   swa 'young, new'   rema 'black'
tluku 'small'   lala 'old'   tshena 'white'
vhi 'bad'   tete 'soft'   tswu 'black'
lapfu 'long'   khwivhilu 'red'   hulwane 'important'
denya 'thick'   sekene 'thin'   nlu 'wet'
vhisi 'raw, green'   pfufhi 'short'   setlha 'yellow'
vhuya 'good-natured'   tswuku 'red'

(Dixon 1982: 4 – 5)

Dixon (1982) observes that in this type of language, the meanings of the reduced class of adjectives seem to cluster consistently around notions relating to his DIMENSION, AGE, VALUE, and COLOUR (DAVC) categories. Of the remaining property concepts, PHYSICAL PROPERTIES in reduced-class languages tend to be expressed by verbs and HUMAN PROPENSITIES tend to be expressed by nouns.

As useful as these observations are, they fall short of a criterial semantic definition. In spite of the fact that adjectives in closed-class languages do tend to express DAVC meanings, the number of meanings whose expressions are unmarked modifiers in such languages can range from dozens down to a mere handful. For example, Venda chooses—out of the potentially much larger set of DAVC meanings—only twenty to be expressed as adjectives. Igbo, on the other hand, realizes only a subset of seven of these (plus 'good', which is not an unmarked modifier of nouns in Venda) as adjectives:

(3)   Igbo

úkwú 'large'   ọ́húrụ́ 'new'   ójí'í 'black'   ọ́ma 'good'
ntà 'small'   ócyè 'old'   ọ́cá 'white'   ọ́jọ́ọ́ 'bad'

(Dixon 1982: 4; Schachter 1985: 15)

The remainder of the set of DAVC meanings in (2) that are not expressed as adjectives in (3) are thus presumably divided up in the Igbo lexicon between the lexical classes of verb and noun. It seems impossible for any purely semantic definition of adjective to be able to single out only the seven words shown in (3) for Igbo and not include all twenty Venda words in (2), let alone for it to include 'good' in one of these languages and exclude it in the other. While it is true that semantic characterizations of parts of speech help to identify likely candidates for inclusion in particular lexical classes, work such as Dixon's shows us that semantics is not the whole story, particularly in the realm of adjectives. As a result, many investigators have turned away from semantic definitions altogether, while others have tried to modify this approach by treating semantic classes of meaning as semantic prototypes for lexical classes (*e.g.* Hopper & Thompson 1984; Croft 1991). Even these approaches run into some difficulties, however, when confronted with the problem of defining the class of adjectives in purely semantic terms.

1.2. Morphological diagnostics

The second type of definition that has enjoyed wide currency is an essentially morphological one that seeks to define parts of speech in terms of the grammatical categories for which they are marked. The simplest approach along these lines posits certain basic morphological categories which are purported to be diagnostic of particular parts of speech, both within and across languages. Thus, nouns are defined as those lexical items that have grammatical gender (Sp. *perro* 'dog$_{MASC}$' : *casa* 'house$_{FEM}$') and are inflected for number (Eng. *dog* : *dogs*) and case (Rus. *mašina* 'car$_{NOM}$' : *mašinu* 'car$_{ACC}$'); verbs mark tense (Eng. *she runs*; *she ran*), aspect (Rus. *kričat'* 'shout$_{IMPF}$' : *kriknut'* 'shout$_{PERF}$'), voice (Bella Coola *tixis* 'he cut it' : *txim* 'it was cut$_{PASS}$' : *txa* 'he cut$_{AP}$'), and mood (Sp. *dices* 'you speak' : *di* 'speak!'); and adjectives mark comparison (Eng. *big* : *bigger* : *biggest*) and in many languages show agreement for number, gender, and case (Rus. *novyj*$_{MASC:SG:NOM}$ *muzej*$_{MASC:SG:NOM}$ 'new museum'; *novye*$_{MASC:PL:NOM}$ *muzei*$_{MASC:PL:NOM}$ 'new museums'; *novaja*$_{FEM:SG:NOM}$ *kniga*$_{FEM:SG:NOM}$ 'new book'; *v novoj*$_{FEM:SG:PREP}$ *knige*$_{FEM:SG:PREP}$ 'in the new book').

Cross-linguistically, however, there is considerable variation with respect to the grammatical categories encoded on lexical items belonging to all three classes. For instance, nouns in Totonac and most Salish languages are not inflected for number, and recent research has suggested that, cross-linguistically, plurality is a category potentially applicable to both nouns and verbs (Dolinina & Beck 1998). Interior Salish languages and Totonac both lack nominal gender, while the coastal Salish languages have gender but generally lack case, as do Chinese and (outside the pronominal paradigms) English and Spanish. Verbs in Salish, Chinese, and many other languages do not inflect for tense, and verbs in Hebrew in all tenses are inflected for gender, as they are in the Russian past tense. In Lushootseed, meanings corresponding to Indo-European tenses and moods can be applied to nouns, as in:

Lushootseed

(4) (a) tu+q̓iyaƛ̕əd ti tu+sč̓istxʷ+s
    PST+slug   D PST+husband+3PO
    'Slug had been her former (*i.e.* deceased) husband'[2]

(Hess 1993: 84)

(b) xʷiʔ kʷi gʷə+pišpiš
    NEG   D   SBJ+cat
    'there are no cats'

(Hess 1993: 123)

(c) ƛ̕u+ləč̓+əd   ti   ƛ̕u+ƛ̓isəd   ʔə šəbad
    HAB+flip+ICS D   HAB+arrow   PR enemy
    'he would flip the habitual arrows of the enemy away'

(Hess 1993: 83)

In (4a) the meaning 'past'—a tense in many languages—is applied to two different nouns, the predicate nominal *q̓iyaƛ̕əd* 'slug' and the syntactic subject *sč̓istxʷ* 'husband', whereas in (4b) the subjunctive, generally classified as a mood, is applied to *pišpiš* 'cat' to indicate its non-existence. (4c) illustrates the application of what is usually glossed as the marker of habitual aspect to a noun, *ƛ̓isəd* 'arrow', the same marker that appears affixed to the verb *ləč̓* 'flip'.

Although there are no cross-linguistically universal morphological categories that can be used to define parts of speech, a more promising approach is to define a set of grammatical categories that are cross-linguistically typical of one or the other parts of speech and then to decide, on a language by language basis, which of these is diagnostic of lexical classes in a given grammatical system. While there is some difficulty with languages like Chinese and Vietnamese which have little or no morphology, definitions along these lines generally single out the same core classes of words as do semantic and syntactic definitions. This, however, is in itself an indication of the fundamental shortcoming of purely morphological definitions: they offer no account of their own success. This success can, in fact, be attributed to a tacit reliance on semantic and syntactic assumptions about the meanings and distributions of parts of speech. Tense in Salish, for instance, might be dismissed as a diagnostic for verbhood given its appearance on the nouns *q̓iyaƛ̓əd* 'slug' and *sč̓istxʷ* 'husband' in (4a), but this presupposes the semantically-driven assumption that these two words are, in fact, nouns.[3] In less problematic cases, say, gender-marking of both verbs and adjectives in Russian, it may be possible to devise more rigorous morphological criteria—*i.e.*, there is a set of words which always bear gender-marking and are marked for the case of their syntactic governors (adjectives), while there is another set which bears marking for tense and aspect (verbs) but can only be marked for the gender of their first syntactic actant in the past tense. However, even if such morphological definitions can be crafted on a language-specific basis, on the cross-linguistic front they do nothing to explain why it is that the bulk of those words singled out as adjectives in Russian morphology express the same meanings and have virtually the same syntactic distribution as those words singled out by the language-specific tests for adjectives in Hebrew, English, Upper Necaxa Totonac, and Japanese.

More telling against purely morphological definitions is the fact that even intralinguistically lexical-class boundaries drawn on purely inflectional bases often give problematic results. Most languages, for instance, have lexical items considered to be a member of a given class which do not have all of the inflections that might be considered criterial for membership in that class. Thus, the English word *significance*—which patterns syntactically and semantically with nouns and does not share any inflectional categories with verbs or adjectives—can be neither a plural (**significances*) nor a possessor (**significance's*), whereas plurality and possessive inflections are commonly cited as morphological indicators of nounhood in English (Lyons 1977: 426). In Russian, a number of words such as *pirožnoe* 'pastry' are declined as if they were adjectives showing agreement with a neuter noun and thus pattern morphologically with more run of the mill adjectives such as *bol′šoj* 'big' and *xorošij* 'good', as shown in (5):

(5)   Declension of Russian *pirožnoe* 'pastry'

	pirožnoe 'pastry_{NEU}'	bol′šoe 'big_{NEU}'	xorošee 'good_{NEU}'
NOM	pirožn+oe	bol′š+oe	xoroš+ee
ACC	pirožn+oe	bol′š+oe	xoroš+ee
GEN	pirožn+ogo	bol′š+ogo	xoroš+ego
DAT	pirožn+omu	bol′š+omu	xoroš+emu
INST	pirožn+ym	bol′š+im	xoroš+im
PREP	pirožn+om	bol′š+om	xoroš+em

Semantically, however, *pirožnoe* is more closely related to morphological nouns such as *pirog* 'pie' and *tort* 'cake'; syntactically, *pirožnoe* patterns as a noun as well, undergoing quantifica-

tion by numerals (*pjat′ pirožnyx* 'five pastries'), serving as subject or object (*Oni s″eli pirožnye* 'They ate the pastries'), the complement of a preposition (*Ja ne mogu žit′ bez pirožnyx* 'I can't live without pastries'), or the head of a relative construction (*pirožnoe, kotoroe ja kupil* 'the pastry that I bought'). What's more, unlike true adjectives but like nouns, *pirožnoe* can not serve as a modifier (**pirožnoe testo* 'pastry dough'), nor does it have comparative (**pirožnee*) or superlative (**samoe pirožnoe*) forms. Given that, first and foremost, lexical classes serve as input to syntactic rules—that is, they characterize lexical items for the purpose of the rules used in the organization of syntactic structures—the classification of *pirožnoe* as an adjective is at best inconvenient, insofar as it is recognized by the syntax as a noun, as reflected in both its functions and its distribution.

Thus, while morphology often supplies important clues as to lexical class membership, morphological definitions—like semantic characterizations—tend to falter when confronted with both cross-linguistic variation in lexical class membership and with intra-linguistic idiosyncrasies. Morphological facts, particularly inflectional categories, however, may reflect a word's underlying semantic and syntactic properties and so may, indirectly, be an indicator of its lexical class. Given the existence of a class of nouns in a language, for instance, it may turn out that all nouns in that language must be inflected for number and so inflectional marking for number can be used by the linguist as a indicator that a word may be a member of the nominal class. This type of indicator, however, is not a definition but a diagnostic, and is purely language-specific: number is not a universal inflectional category for nouns and in some languages it is an inflectional category for other parts of speech as well. Indeed, the fact that number is frequently marked on nouns follows from the semantics of the prototypical noun, but the appearance of number with other lexical classes such as adjectives is merely agreement and so follows from their syntactics. Morphology in this sense becomes a somewhat superficial phenomenon, depending not so much on the universal characteristics of lexical classes so much as how these classes are treated by the morphosyntax of a given language.

1.3.   Syntactic distribution

As noted in the previous section, syntactic distribution is often more closely related to lexical class membership than are inflectional patterns; this type of observation has led to a number of attempts at defining parts of speech (often called "syntactic categories" in such definitions) in purely distributional terms. The most elementary of these approaches defines each part of speech strictly on the basis of the syntactic roles in which it is permitted to appear: nouns are defined as lexical items that can be the subjects of a sentence, verbs can be syntactic predicates, and adjectives are modifiers (*cf.* Chomsky 1965). Such naïve definitions falter, however, when confronted with lexical items appearing in non-prototypical or extended uses. Words defined as nouns this way, for instance, can comfortably serve as attributives of other nouns, as in *jazz musician* or *gas stove*, while certain adjectives can act as syntactic subjects (*e.g. The rich fear the poor*; *Louder is better*). As we will see below, such difficulties can be dealt with by appealing to criteria of markedness, but without a rigorous apparatus to deal with this issue (such as that proposed in section 1.4) naïve syntactic or distributional definitions fail.

An additional cross-linguistic difficulty with distributional definitions of parts of speech comes from languages with reduced lexical inventories—that is, languages which appear, on a distributional basis, to lack one or more lexical class distinctions. Consider, for instance, the ex-

amples in (6) from the Salishan language Lushootseed, which show the distributional overlap between verbs and nouns ((6a) and (b)), and between verbs and adjectives ((6c) and (d)):

Lushootseed
(6)   (a) <u>ʔu+ʔɬəd</u> tiʔiɬ <u>pišpiš</u>   ʔə tiʔəʔ sʔuladxʷ
            PNT+eat   D   cat   PR   D   salmon
            'that cat ate a salmon'

    (b) <u>pišpiš</u> tiʔiɬ <u>ʔu+ʔɬəd</u> ʔə tiʔəʔ sʔuladxʷ
            cat   D   PNT+eat   PR   D   salmon
            'that one eating the salmon [is] a cat'

(based on Hess 1993: 133)

    (c) <u>ƛ̓u+ləkʷ+əd</u>   tiʔəʔ <u>haʔɬ</u>   sʔɬəd
            HAB+eat+ICS   D   good   food
            '[he/she] would eat the good food'

    (d) <u>haʔɬ</u>   tiʔəʔ sʔɬəd <u>ƛ̓u+ləkʷ+əd</u>
            good   D   food   HAB+eat+ICS
            'the food [he/she] would eat [is/was] good'

(based on Bates *et al.* 1994: 105)

The example in (6a) illustrates a sentence whose predicate is the verb *ʔɬəd* 'eat$_{INTR}$' and whose syntactic subject is *tiʔiɬ pišpiš* 'that cat'; in (6b) the same word, *pišpiš* 'cat', serves as syntactic predicate. In (6d) the syntactic predicate is *haʔɬ* 'good', which is shown acting as a modifier in (6c). A naïve distributional definition of a verb as "a lexical item that can act as a syntactic predicate" would not only pick out the syntactic predicates of (6a)—*ʔɬəd* 'to eat$_{INTR}$'—and (6c)—*ləkʷəd* 'eat$_{TRANS}$'—as verbs, but would pick out the syntactic predicates of (6b)—*pišpiš* 'cat'—and (6d)—*haʔɬ* 'good'—as well. This type of argument can and has been used as evidence that Lushootseed, and Salish in general, lacks an underlying lexical distinction between verbs, nouns, and adjectives (*e.g.* Kuipers 1968; Kinkade 1983). As shown in Beck (1999), the use of verbs as actants in Lushootseed is a structurally marked one (see Section 1.4.1), the expression *tiʔiɬ ʔuʔɬəd ʔə tiʔəʔ sʔuladxʷ* 'the one eating the salmon' in (6b), for instance, being (as its gloss suggests) a relative clause headed by the deictic *tiʔiɬ*.[4] Thus, the naïve distributional criterion gives the wrong results for Lushootseed: if Lushootseed does have nouns and verbs, then clearly something other than straightforward distribution has to be used in the definition of lexical classes.

    Just as with semantics and morphology, then, syntactic distribution, in and of itself, fails as an adequate means of defining parts of speech. Intra-linguistically, words of a given lexical class are frequently capable of appearing in syntactic roles which are typical of, or even diagnostic of, other parts of speech; cross-linguistically, we frequently find variation with respect to the syntactic roles open to different parts of speech, as illustrated by the Lushootseed examples in (6). Nevertheless, it is true that there are certain widespread commonalities in the syntactic behaviour of lexical classes, just as there are prototypical semantic domains and inflectional categories associated with them. This type of observation has lead some researchers to treat syntactic variation in the distribution of parts of speech in terms of the *markedness* of a given syntactic role for

members of a particular lexical class: in essence, such approaches—like the naïve syntactic definition—identify certain syntactic roles as being typical or unmarked for a given part of speech and then allow, in one way or another, for the appearance of that part of speech in other, marked, roles in the sentence. Thus, the appearance of the normally adjectival *red* as a syntactic subject in *Red is my favorite colour* could be argued to be an example of a marked or extended use of *red* in a basically nominal syntactic role. However, to do this without recourse to purely stipulative definitions of parts of speech it is necessary to show that the behaviour of an element in an extended position is in some way marked. These are complex issues to be taken up in more detail in the section that follows.

1.4. Extended roles and syntactic markedness

As we have seen, neither semantic characterizations (Section 1.1) nor naïve syntactic distribution (Section 1.3) are sufficient in and of themselves to allow for a criterial definition of lexical classes. Morphological properties of words turn out to be useful as diagnostics for lexical-class membership on a language-specific basis, but fail both as universal and intralinguistically comprehensive definitions (Section 1.2). Of the three levels of linguistic description, it is the morphological which shows the greatest cross-linguistic variation in terms of its marking of lexical class distinctions and so ultimately seems to be the least useful in terms of finding a criterial definition. This leaves us with the semantic and the syntactic levels, both of which have been the focus of more recent attempts to define parts of speech. The variation that we have seen described in the sections above is dealt with in different ways by different authors, but in general two approaches have become predominant in the literature. Semantic approaches have by and large moved in the direction of treating variation in class-membership in terms of *prototypicality* and graded class membership (*e.g* Hopper & Thompson 1984). More recent syntactic approaches have dealt with distributional variability in terms of *markedness*—that is, they have sought to define parts of speech in terms of their unmarked or typical syntactic distribution. This will be our focus here, beginning with a discussion of what markedness means and how it will be measured throughout the remainder of our discussion (Section 1.4.1). Section 1.4.2 introduces the notion of markedness as the measure of the prototypicality of certain types of mapping between semantic class and "pragmatic" role as put forward by Croft (1991). While Croft's proposal has some weaknesses, it does point us in the direction of what seems to be the correct approach to forming accurate definitions of lexical classes that at once account for typological similarities in parts of speech systems and predict the attested variation in these systems in the world's languages.

1.4.1. Criteria for markedness

Before looking at syntactic definitions of parts of speech based on markedness, it is worth taking a little time to clarify what it is precisely that is meant by markedness and what kinds of criteria will (and will not) be allowed in this discussion. Markedness is one of the most widely, and wildly, used terms in linguistics, and its senses range from a very narrow, structure-based notion of relative complexity to an extremely open sense of "unusual" or "unnatural." A recent definition of markedness located somewhere in the middle of the continuum is put forward by Givón (1995), who writes that "three main criteria can be used to distinguish the marked from the unmarked category in a binary grammatical contrast:

(a) **Structural complexity:** The marked structure tends to be more complex (or larger) than the corresponding unmarked one.

(b) **Frequency distribution:** The marked category (figure) tends to be less frequent, thus cognitively more salient, than the corresponding unmarked category (ground).

(c) **Cognitive complexity:** The marked category tends to be cognitively more complex—in terms of mental effort, attention demands or processing time—than the unmarked one."

<div align="right">(Givón 1995: 28)</div>

Of these three criteria, (a) is the least controversial and the most universally accepted: given the contrast between two (comparable) signs A and B, the more complex of the two is the marked one. The second and third items on Givón's list, however, are much less straightforward. Frequency is a very commonly cited criterion for markedness, due largely to the intuitive feeling that the unmarked is the most usual or standard form. While this may often be the case, it is not always so, and Trubetskoy (1969: 262ff) argues explicitly against frequency as a reliable indicator of markedness, offering a number of examples of phonological segments which are marked (in terms of their complexity, etc.) but are statistically more frequent than their unmarked counterparts. The unreliability of frequency as a measure of markedness also becomes obvious if we think in concrete terms. In phonology, for instance, the appearance of a marked phoneme in a high-frequency word (say, a function word, a common morpheme, or a usual expression) could potentially make the instances of that phoneme more frequent than those of its unmarked counterpart. In our own domain of lexical classes, it turns out that in English the predicative use of adjectives is textually more frequent than the attributive use (Thompson 1988)—yet clearly, judged in terms of structural complexity (adjectival predicates require a copula), the former is the more marked of the two constructions. Thus, while frequency in a textual sense may tend to correlate with markedness, it is neither a necessary nor a sufficient criterion for it and will not be used in the course of our discussion.

There is, however, a type of markedness that will be used here that is, at least intuitively, related to the notion of frequency. Consider the following situation: in a particular language, words belonging to the lexical class X appear in six structural environments $\{E_1, E_2, \ldots E_6\}$. In three of the six environments, X displays a set of properties $\{P_1, P_2, \ldots P_7\}$ (*e.g.* inflectional categories, referential meaning, etc.), but in E_3 and E_5 X displays a reduced set of these properties $\{P_1, P_3, P_7\}$ and in E_6 it has only one of these $\{P_2\}$. Environments E_3, E_5, and E_6, then, can be considered as marked structural configurations for X with respect to the remainder of environments $\{E_1, E_2, E_4\}$ in which X displays the greatest range and most consistent set of properties. The markedness of a given environment, then, can be determined by a reduction in number of typical properties of X, which are those which X displays in the largest number of environments. English nouns, for example, may be inflected for number and take a possessor in most of the syntactic roles they occupy (subject, object, object of a preposition), but when they are used attributively in an expression like *pant leg*, nouns can not bear number inflection (*pants leg*) or take a possessor (*a/the his-pant leg*). This sort of loss of categorial properties is often referred to as *decategorialization* (Hopper & Thompson 1984) or *decategorization*.

Markedness can also result from the acquisition of a new property, P_8, in one or more of the environments open to X, provided that either a) the number of environments in which X has the set of properties $\{P_1, P_2, \ldots P_7\}$ is greater than the number of environments where X has the set of properties $\{P_1, P_2, \ldots P_8\}$ or b) X has, in addition to P_8, only a restricted subset of the other

properties $\{P_1, P_2, \ldots P_7\}$ in the marked environment. This is frequently referred to in the literature as *recategorization*. On its own, this type of markedness is trickier to establish, particularly for lexical items that have a very limited number of syntactic roles. Generally, it is only invoked in cases where P_8 is felt to be marked in its application to X for other reasons—that is, that it is typical of another lexical class, Y, or it is marked in terms of complexity with respect to some other environment in which X appears. A classic example of recategorization is the Russian participle or *prichastie*, the modificative form of a verb which takes adjectival inflection for the case, number, and gender of its governing noun. Both de- and recategorization might seem like a kind of contextual frequency in that they are determined by the "frequency" with which a certain set of properties is associated with the members of the set of environments $\{E_1, E_2, \ldots E_6\}$ open to X. Because of the dangers inherent in the term "frequency", however, it is safer to refer to these measures as *contextual* markedness, a term which has the added advantage of reminding us (as noted by Givón above) that the markedness of a given item must not only be determined relative to some other item of a comparable nature, but also must be determined for a specific context.

Givón's third criterion, cognitive complexity, is also somewhat problematic, although if used judiciously it turns out to be a useful one. Terms such as "mental effort, attention demands or processing time" are frequently used in a hand-waving fashion without regard to the fact that—as real-world, neurological events—they are subject to empirical verification. Failing psycholinguistic measurement of complexity based on the criteria proposed by Givón, then, it is important to be very clear what we mean by "cognitive complexity" and to provide plausible reasons to think that this complexity would indeed correspond to increased effort, attention, or processing time. To this end, I wish to propose one, specific type of cognitive complexity that will play a role in the discussion below, something that I will refer to as *non-iconicity*. According to this criterion, a linguistic sign *a* ‹'a', A› (that is, the sign *a* having the signified 'a' and the signifier A) is more cognitively marked (cognitively complex) than a sign *b* ‹'b', B› if A is a less direct reflection of 'a' than B is of 'b'. On its own, of course (like all uses of the term *iconicity*—cf. Haiman 1980) this can be very open-ended, although it will be put to a single, highly constrained and specific use in Section 2.1 below (to which the reader is referred for a concrete example). The rationale for this criterion is simply that a non-iconic sign will be harder to recognize than an iconic sign, and hence is cognitively more complex. Straying from the field of linguistics, an illustration of this might be the mental effort required to recognize a picture (a direct representation matched to visual information) of a familiar object—say, an alarm clock—versus recognizing it from a description (which requires lexical access and linguistic processing). In writing systems, an ideographic system is more difficult to learn in that the representations of words contain no information about their phonological shape, whereas an alphabetic system allows learners to match written representations directly to spoken words. This last example probably gives us as good a formulation of the notion of "direct representation" as we are going to get: the more direct a representation is (that is, the more iconic it is) the more information it contains about the underlying content it represents. Thus, if the signifier B contains more information about 'b' than A does about 'a', *b* can be said to be less marked (and more iconic) than *a*.

The criteria for markedness that will be used here, then, differ somewhat from those put forward by Givón. This discussion makes use of three criteria for syntactic markedness:

(7) (a) **Structural complexity:** A sign X is marked with respect to another sign Y if X is more complex, morphologically or syntactically, than Y.

(b) **Contextual markedness:** An environment E is a marked one for a sign X if E is not a member of the largest subset of environments of X where X shares the greatest number of common properties with other instances of X (hence, the appearance of X in this environment can be said to be marked or an extended use).

(c) **Cognitive complexity**: A sign X is marked with respect to another sign Y if the representation of X is a less direct expression of X's meaning than the representation of Y is of Y's meaning.

An important point to be made about all of these criteria is that they are formulated in terms of contrast—that is, it is not enough to say that X is marked, it is necessary to specify what it is that X is marked in contrast to. Thus, it is essential to keep in mind that markedness is always *contrastive*, an issue which is central to the discussion in the section that follows.

1.4.2. Markedness and prototypical mappings

Approaches based on markedness and the syntactic distribution of parts of speech have the dual advantage of defining the typical syntactic distribution shown by words of a given lexical class, while at the same time allowing for extended uses of lexical items which in themselves can be the source of diagnostic patterns. In one of the best examples of this methodology to date, Hengeveld (1992a, 1992b) makes reference to the "additional" grammatical machinery required to allow a lexical item to appear in an extended syntactic role. Hengeveld refers to such mechanisms as "further measures" and uses this notion to arrive at the definitions in (8):[5]

(8) *verb*—a lexical item that without further measures being taken can be used as a syntactic predicate only.

noun—a lexical item that without further measures being taken can be used as the actant of a syntactic predicate.

adjective—a lexical item that without further measures being taken can be used as the modifier of a noun.

For Hengeveld, "further measures" are defined as those morphosyntactic means which "derive" Functional Grammar predicates from constituents that are not already predicates (1992a: 58). In more conventional terms, "further measures" can be regarded as the morphosyntactic properties acquired or lost by an sign in a non-prototypical syntactic role. These properties can thus be used as diagnostics for markedness of a lexical item in a given syntactic environment.

On the down side, the sets of lexical items singled out by definitions like those in (8) are potentially arbitrary and their membership is unconstrained—that is, under a purely syntactic definition, there is no obvious reason why it is that the meaning 'dog' is a noun in all languages rather than a verb, or why it is that meanings like 'red' show cross-linguistically variable class membership while meanings like 'break' do not. And this brings us back to the issue of the semantics of lexical classes. For, while it is true that purely semantic definitions have met with relatively little success, it is still a fact that there is across languages a highly consistent common core or "focal class" (Lyons 1977: 440) of meanings associated with each of the major lexical classes, and that these semantic prototypes correlate in predictable (although not always predic-

tive) ways with the lexical classes singled out by definitions in terms of syntactic markedness. This type of correlation is made explicit in the work of Croft (1991), who argues that, cross-linguistically, words which are unmarked in one of three principal "pragmatic functions" (roughly, syntactic functions) belong prototypically to a particular class of meanings, as in (9)[6]

(9)   Croft's mapping between semantic class and pragmatic functions

	Noun	Adjective	Verb
Semantic class	object	property	action
Pragmatic function	reference	modification	predication

(Croft 1991: 55)

To account for typological variation in semantic class membership, Croft goes on to identify four semantic properties which typify each class and which can be used as criteria for class membership, each of which has the prototypical values for the respective semantic classes in (10):

(10)   Prototypical values of features for semantic classes

	Objects	Properties	Actions
Valency	0	1	≥ 1
Stativity	STATE	STATE	PROCESS
Persistence	PERSISTENT	PERSISTENT	TRANSITORY
Gradability	NON-GRADABLE	GRADABLE	NON-GRADABLE

(Croft 1991: 65)

Individual languages may depart from these prototypes in that they can differ with respect to which value of a given feature is assigned to a particular meaning and, hence, to which of Croft's semantic classes that meaning is felt to belong.

A second source of potential variation in Croft's system is in the mappings between specific semantic classes and "pragmatic functions", which in (9) represent only the prototypical mappings found in three-class systems. There are, however, languages which seem to neutralize certain parts of speech distinctions and thus allow for alternate mappings. For instance, in Hausa (Chadic), most English adjectives translate as to nouns, meaning that both **Properties** and **Objects** must be mapped onto Croft's pragmatic functions of reference and modification. As a result, modification is realized by what is frequently (mis-)glossed as a possessive construction:[7]

Hausa
(11) (a) mùtuᵐì   mài   àlhe..öì   /   aözìki…   /   hankàli…
           person   ATRB   kindness   /   prosperity /   intelligence
           'a kind/prosperous/intelligent person'

      (b) ità..če… mài   tauöi…   /   lauši   /   nauyi…
           wood   ATRB   hardness /   softness   /   heaviness
           'hard/heavy/soft wood'

12

(c) mùtuḿ   mài   do̥kì…
     person   ATRB   horse
     'a person who has a horse'

<div align="right">(Schachter 1985: 15 – 16)</div>

On the other hand, in the Bantu language Bemba, words denoting both property concepts and actions are verbs and can be used as either syntactic predicates or modifiers inside relative clauses, as in (12):

Bemba
(12) (a) umuuntu   ù+ashipa
           person   RELATIVE:SUBJ:CONCORD+brave
           'a person who is brave'

      (b) umuuntu   ù+alemba
           person   RELATIVE:SUBJ:CONCORD+write
           'a person who is writing'

      (c) umuuntu   á+ashipa
           person   SUBJ:CONCORD+brave
           'the person is brave'

      (d) umuuntu   á+alemba
           person   SUBJ:CONCORD+write
           'the person is writing'

<div align="right">(Schachter 1985: 16)</div>

Thus, Bemba seems to map both the semantic classes of **Properties** and **Actions** onto the two separate pragmatic functions of predication and modification—in effect, neutralizing the distinction between the classes of verb and adjective (actions and property concepts both being verbs), just as Hausa seems to neutralize the distinction between adjectives and nouns (actions and property concepts both being nouns).

   Although Croft (1991) expresses some skepticism about the existence of languages which completely neutralize the prototypical semantics–pragmatics correspondences, he does allow for the possibility that such languages exist, provided that they do not show a "negative markedness pattern … so that, say, lexical roots denoting objects take a non-zero function-indicating morpheme in referring expressions but not in modification" (p. 94). In other words, Croft argues that there are no languages where words denoting objects are marked as referential items but are unmarked modifiers, or where words denoting actions are marked as syntactic predicates but are unmarked subjects. And indeed there are not. What is a problem for Croft is not the existence of languages that his model excludes, but the non-existence of languages that his model allows. The existence of languages of the Hausa-type (neutralization of the object–property distinction) and of the Bemba-type (neutralization of the action–property distinction) is widely known and well documented—however, there do not appear to be any languages that show the third possible type of neutralization, that between objects and actions. That is, even among those languages that are claimed in the literature to neutralize the distinction between nouns and verbs (*e.g.* Tongan—Broschart 1997; Mundari—Bhat 1994; Tuscarora—Hengeveld 1992b), there are no lan-

guages that are claimed at the same time to have a distinctive class of adjectives—that is, while there are languages that lack adjectives but have verbs and nouns, there are no languages that have adjectives and nouns but lack verbs, or which have adjectives and verbs but lack nouns. It is this fact which leads Hengeveld (1992b: 68) to propose the "parts-of-speech hierarchy" in (13):

(13)   Parts of Speech Hierarchy

        Verb > Noun > Adjective > Adverb

(13) is an expression of the implicational relations that hold for the existence of various parts of speech and states that the existence of a lexical class on the right of the hierarchy in a given language implies the existence of all of those classes to its left. Any language which has a class of adverbs, then, must have all three of the classes of verb, noun, and adjective, and—more to the point—any language that has adjectives must also have verbs and nouns.[8]

In general, implicational hierarchies such as that in (13) are dealt with in terms of markedness—in other words, elements on the right end of the scale are considered to be more marked than elements to their left (Greenberg 1963) in the sense that marked distinctions are expected to be more readily neutralized than unmarked distinctions (Trubetskoy 1969). Thus, the typological distribution of the class of adjectives indicates that, as a lexical class, adjectives are more marked than verbs. In a certain sense, Croft (1991: 130ff) acknowledges this and offers some ancillary explanation for the fact that adjectives are in many ways an "intermediate" class between verbs and nouns, although there is nothing inherent in his approach to account for the markedness of the adjectival class or the typological asymmetries in neutralization patterns in and of themselves. Similarly, Hengeveld (1992a, 1992b), who puts forward the hierarchy in (13), offers little in the way of motivation for it. In the following section, it will be my goal to offer just such a motivation by outlining a definition of "adjective" which both takes into account their typical semantic and syntactic properties and provides a natural explanation of their relative markedness.

2.   Semantics, syntax, and the lexicon

As we saw in the preceding section, while there is no consensus as to the precise definitions of the three major parts of speech, there is a certain consensus as to some of the prototypical semantic properties of these lexical classes and what their unmarked syntactic roles might be, at least for nouns and verbs. Adjectives, however, are a more difficult problem. While nouns and verbs are relatively consistent as classes in terms of their semantic and syntactic properties, adjectives can not be easily characterized in such a way as not to include words that are, either intra- or cross-linguistically, members of other lexical classes. Prototypical nouns are expressions of semantic names or kinds (Wierzbicka 1988) and are universally unmarked syntactic actants of verbs; verbs are analyzable as the expressions of semantic predicates and are always unmarked syntactic predicates. Not only are both of these characterizations—the semantic and the syntactic—accurate and highly consistent within and across languages, but so is their combination. In other words, what is an unmarked syntactic actant in a given language is prototypically the expression of a semantic name and what is an unmarked syntactic predicate is prototypically a semantic predicate. Thus, the prototypical semantic class and the unmarked syntactic role that we have defined for nouns and verbs above represent (in the spirit of Croft 1991—see (9) above) the unmarked mapping between semantic class and syntactic role for these two parts of speech:

(14) Unmarked mapping between semantic class and syntactic role

	Noun	**Verb**
Semantic class	name	semantic predicate
Syntactic role	syntactic actant	syntactic predicate

As we will see below, these correspondences—formalized in Sections 2.1 and 2.2—will allow us to formulate a definition of the third major class, adjectives, in a way which shows it to correspond to a more marked mapping between semantic and syntactic structure, thereby accounting for the attested cross-linguistic variation in the lexical class of adjective (Section 2.3).

2.1. Iconicity at the semantics ⇔ syntax interface

The correspondences between the prototypical semantic class and the unmarked syntactic role of nouns and verbs illustrated in (14) can be represented graphically borrowing some formalisms from Meaning-Text Theory (Žolkovskij & Mel'čuk 1967; Mel'čuk 1988). The diagram in (15) shows the unmarked mapping between a semantic and a syntactic structure, in this case the (simplified) semantic and syntactic representations of the sentence *The boy runs*:

(15) Semantic ⇔ syntactic mapping for *The boy runs*
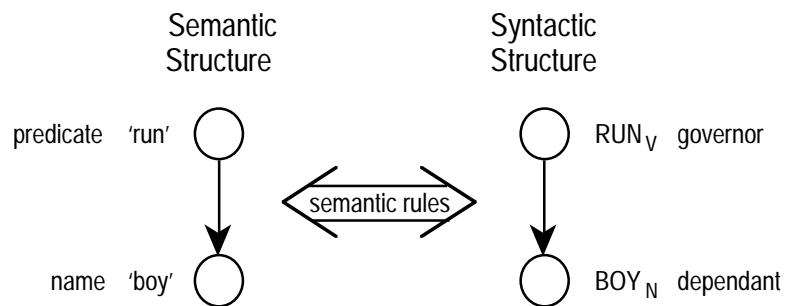
The left half of (15) represents the semantic relation holding between the predicate 'run' and its argument, the semantic name 'boy' and is the graphic equivalent of the more familiar expression 'run'('boy'). The right half of (15) is a syntactic dependency tree showing the hierarchical relation between the lexeme RUN and its single syntactic actant BOY.[9] In both representations, the direction of the single arrow indicates the direction of the dependency, running from the governor and pointing towards the dependant. The broad, double arrow in the diagram represents the correspondence or mapping between the two levels of representation. This correspondence is made by rules for translating between semantic structure and (deep) syntactic structure via the selection and combination of items from the lexicon (lexicalization and syntacticization, respectively). In the syntactic structure shown in (15), RUN is a syntactic predicate (the top node in a dependency tree or the head of the highest lexical constituent in a phrase-structure tree) and BOY is its syntactic actant; thus, each of them occupy their respective unmarked syntactic roles—RUN, a verb, as governor of the expression of its semantic argument and BOY, a noun, as dependant of the expression of its semantic predicate.

This syntactic configuration faithfully mirrors the underlying semantic structure and so can be said to be iconic, in that the syntactic tree in (15) contains a predicate–actant relation which is

a direct expression of the underlying predicate–argument relation in the semantic representation—that is, the semantic predicate 'run' is realized as a syntactic predicate RUN, and its argument or semantic actant 'boy' is realized as the syntactic actant of RUN, making the syntactic structure a direct or isomorphic representation of the semantic structure. Structures that express semantic predicate–argument relations but depart from this pattern in some way can be said to be cognitively complex (section 1.4.1) in that they are not direct reflections of the underlying semantic representation. This allows us to characterize the markedness of a given semantic ⇔ syntactic structural mapping in terms of the Principle of Weak Iconicity:

(16) *The Principle of Weak Iconicity*

> In the unmarked case, syntactic structure will be isomorphic with, or a direct reflection of, its underlying semantic structure.

Thus, by Weak Iconicity, the unmarked syntactic role for the expression of a semantic predicate (which is the semantic governor of its argument) is that of syntactic governor of the expression of its argument; conversely, the unmarked role for the expression of a semantic name is as the syntactic dependant of the expression of its semantic predicate. In English (and probably universally), the lexicon distinguishes two types of words, nouns and verbs, each of which is (prototypically) an expression of one of these two semantic categories (semantic predicates or semantic names) and each of which appears unmarked in the corresponding iconic syntactic role.

2.2. Lexicalization and syntactic structure

In the representation in (15), lexical items in the syntactic tree are shown bearing specification for membership in a particular lexical class—in this case "N" (noun) for the dependant BOY and "V" (verb) for its syntactic governor RUN. This raises the question of where the classes "N" and "V" come from—that is, what sort of knowledge about the lexemes RUN and BOY these labels are intended to represent, and to which component of a grammar this knowledge must be assigned. The commonly-held answer to these questions is that the classes "noun" and "verb" are proper to the lexicon (*i.e.* they are lexical classes). While the specific nature of the lexicon has been a fascinating and frequently contentious issue for both formal (*e.g.* Saussure 1916; Mel'čuk 1995; Pustejovsky 1996; Koenig 1999) and cognitive-functional (*e.g.* Fauconnier 1985; Lakoff 1987; Taylor 1989) approaches to linguistics, the detailed structure of lexical knowledge and the exact nature of lexical entries need not concern us here. For the moment it is enough for us to think of the lexicon as a *lexical inventory*—that is, as an organized and categorized inventory of lexical items available to the speaker of a language to express the contents of semantic representations. This inventory would combine information about the phonological shape of words with information about their semantic structure and their syntactic uses. Sentence-building, then, can be thought of as consisting in part of a matching of elements of semantic structure with appropriate entries in the lexicon, as shown in (17):

(17)   Lexicon

Semantic   A → Ⓥ   Syntactic   Morphological
Structure   Ⓝ   Structure   Structure

○ 'run'   RUN_V ○

                    ⟨ semantic rules ⟩   ⟨ syntactic rules ⟩   the boy run+s

○ 'boy'   BOY_N ○
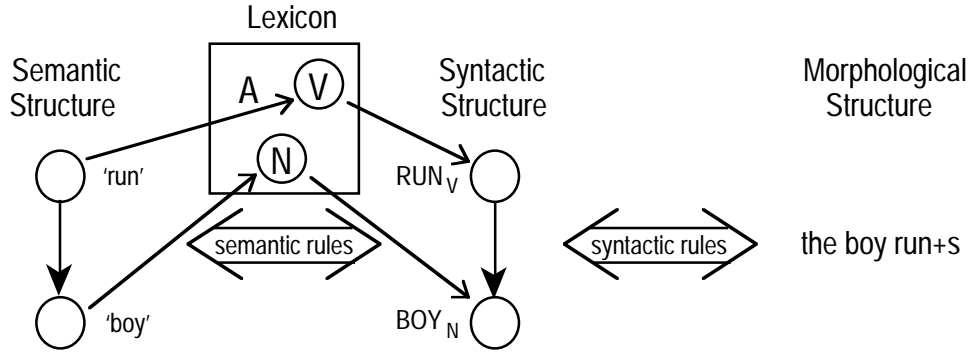
In the first set of correspondences (semantic ⇔ syntactic structure) in (17), the semantic predi-
cate 'run' and its argument, the name 'boy', are matched to the lexical entries for RUN and
BOY, respectively. These lexemes are drawn from various regions of the lexicon (represented by
the capital letters "N", "V", and "A") and so are specified as "verb" or "noun" depending on
which region they belong to. The next set of correspondences in (17), those between syntactic
and morphological structure, shows how lexical class information is used in conjunction with
syntactic configuration to determine the correct morphological strings. In this particular case, the
rules that establish the correspondences between syntactic and morphological structure in Eng-
lish require SV(O) word-order and that a verb agree with its subject in person and num-
ber—hence, the affixation of the third-person singular -*s* to "run" in the morphological repre-
sentation (again, the determiner here is ignored).

    The same set of syntactic rules that accounts for the ordinary inflection of words in their un-
marked syntactic roles accounts for the morphological features of their extended uses as well.
Consider, for instance, (18), which represents the English expression *running boy*:

(18)   Semantic ⇔ syntactic ⇔ morphological mapping for *running boy*

Semantic   Syntactic   Morphological
Structure   Structure   Structure

'run' ○   ○ BOY_N

        ⟨ semantic rules ⟩   ⟨ syntactic rules ⟩   runn+ing boy

'boy' ○   ○ RUN_V
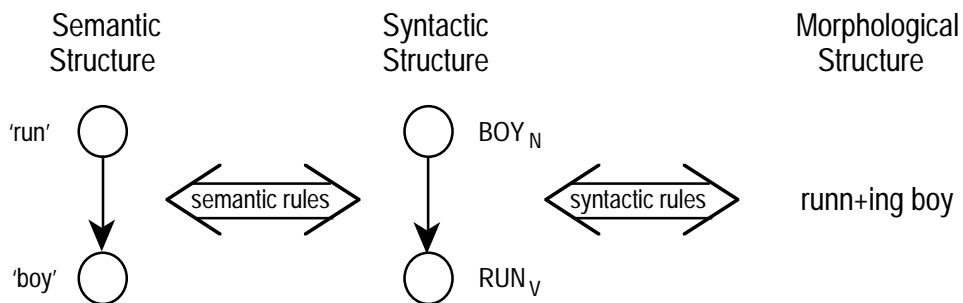
The left side of (18) illustrates the semantic relation between the predicate 'run' and its argument
'boy'; note that the formal semantic relationship between the two is identical to that shown in
(15) and (17), while the syntactic relation between them is reversed (on the reversal of depend-
ency relations between corresponding elements at different levels of representation, see Mel'čuk
1988: 105–149). Thus, in the phrase *running boy*, the semantic predicate 'run' is not realized as
the governor of the syntactic structure, as it is in (15) and (17): instead, the syntactic governor of
the configuration is the expression of the semantic argument 'boy'. Thus, RUN is realized in the
syntactic role of modifier, which can be defined for our purposes as in (19):

(19) *modifier*—the expression of a semantic predicate that is syntactically a dependant of the expression of its semantic argument.

Given that RUN is specified in the lexicon as a verb and that verbs in English are not unmarked syntactic modifiers, the rules for making morphological strings in English require that verbs appearing in modifier position make use of further morphosyntactic measures—in this case, the suffix *-ing* as shown on the right of (18). Thus, verbs can be said to be marked in this role by dint of increased structural complexity.

2.3. Adjectives, markedness, and iconicity

If verbs are marked in the role of modifier (that is, they require further morphosyntactic measures to be realized in syntactic structure as the dependant of one of their arguments), adjectives have the opposite characteristic: syntactically, they are unmarked dependants of the expressions of their semantic arguments. The contrast between the unmarked configurations of the two lexical classes can be seen by comparing (15) above with (20), which shows the mapping between semantic and syntactic structure for the English modifier construction *big boy* :

(20)   Semantic ⟺ syntactic mapping for *big boy*
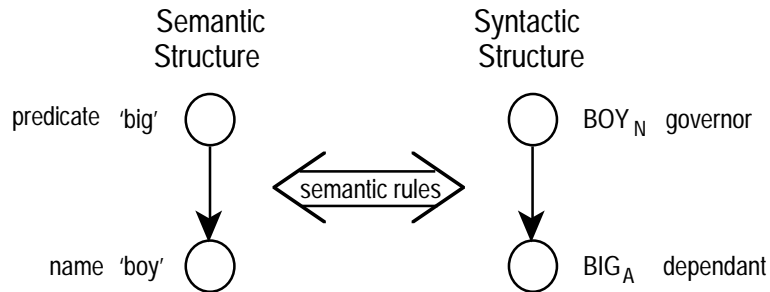
Semantically, adjectives resemble verbs in that they are the expressions of semantic predicates (*e.g.* Lakoff 1965; Chafe 1970; Bhat 1994) and typically take semantic names as their arguments, as shown on the left of (20). Here, as in (18), the direction of the semantic dependency between the predicate 'big' and its argument 'boy' is reversed in the syntactic structure, where the noun BOY is the syntactic governor of its modifier BIG. The difference between (18) and (20) lies in the fact that the expression of the predicate 'big' is, in English, an adjective, whereas the expression of 'run' is a verb; these two parts of speech are treated differently by the rules for forming morphological structures, as illustrated in (21):

(21)   Lexicalization of adjectival versus verbal modifiers
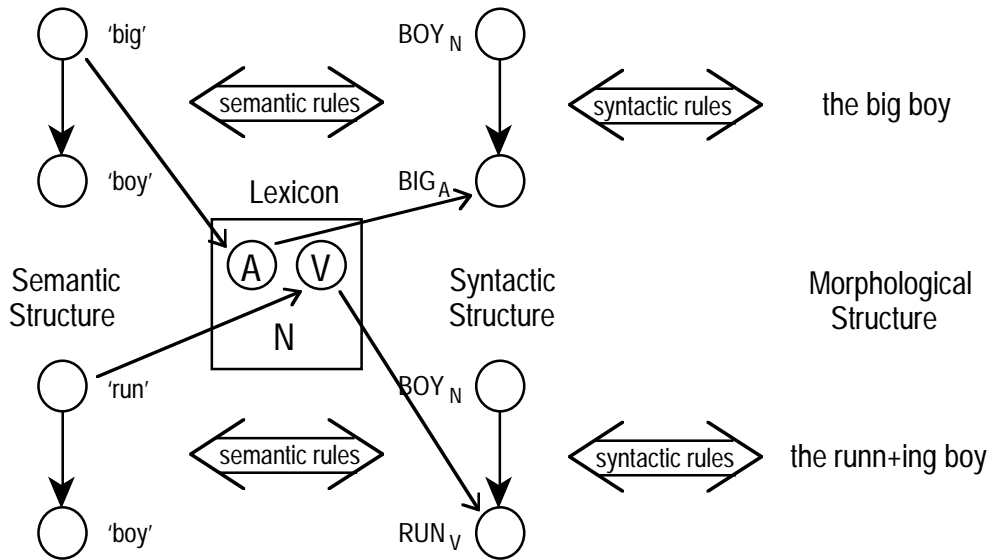
Although BIG, like RUN, is the expression of a semantic predicate, its appearance as a syntactic dependant of BOY (the expression of its semantic argument) is not marked because in the lexicon it is classified as an adjective—and this syntactic configuration is unmarked in English for this part of speech. By the same token, the use of BIG as a syntactic predicate in an expression like *the boy is big* is also a marked configuration requiring the use of further morphosyntactic measures—specifically, the copula *to be*. It is this pattern of contextual markedness that allows us to identify the unmarked syntactic role of adjectives both within and across languages: in all languages that have adjectives, adjectives are unmarked modifiers of nouns, whereas in only some languages are adjectives also unmarked syntactic predicates.

 The fact that adjectives are unmarked syntactic modifiers on the one hand and that they are semantic predicates on the other allows us to define them as follows:

   (22) *adjective*—a lexical item expressing a semantic predicate that can be an unmarked
        modifier of a noun.

An important feature of this definition is that it offers a straightforward explanation of the relative markedness of the adjectival class: the contextually unmarked syntactic role of an adjective (modification) is a cognitively marked syntactic relation. Unlike reference and predication, modification as a syntactic role represents a reversal of the direction of the dependency that obligatorily holds between semantic predicate and semantic argument, whereas the direction of the dependency in a syntactic structure like RUN ○→○ BOY from (15) is a direct reflection of the direction of the dependency in the semantic structure. Thus, the syntactic representations in (18) and (20) can be said to be marked in terms of cognitive complexity with respect to the structure shown in (15), meaning that the unmarked syntactic role for adjectives is a marked or non-isomorphic one.

 Those languages like English that distinguish a lexical class of adjective designate a specialized subset of semantic predicates for this marked role—that is, for the purposes of modification. These words are considered, either on the basis of their contentive features or by dint of diachronic processes, to occur most naturally as syntactic modifiers and so are allowed by the rules

of the grammar to take on this role. How this type of differentiation takes place is, of course, an open question and is likely a combination of cognitive considerations, the effects of grammaticalization processes, features of language acquisition, and discourse factors. As interesting as some of these issues might be, it is not the purpose of this investigation to account for the idiosyncrasies of the class of adjectives in a given language. The domain of definitions, after all, is the domain of regularities, and it is the role of a criterial definition to capture those generalizations that hold universally for that which it defines.

In this case, our proposal is that by definition adjectives are the expressions of semantic predicates and that they are unmarked modifiers of nouns. The former characteristic is a function of the meanings expressed by individual words and the latter of how the expressions of these meanings are treated in the syntax. In some languages, this treatment will be largely by rule (*e.g.* property-concept words are adjectives) while in others it will be by convention (*e.g.* 'good' and 'red' are expressed as adjectives but 'nice' and 'black' are not). Similarly (although to a lesser extent), languages may vary as to the treatment of what are, at least in translation, largely synonymous concepts. In English and Lushootseed, for instance, $cold_A$ and $t̓əs_V$ 'cold' are the expressions of semantic predicates, but in Upper Necaxa Totonac *lon·* 'cold' is a noun like the abstract noun $cold_N$ in English.[10] Again, how such meanings, peripheral to the semantic prototypes of both the nominal and the verbal/adjectival categories, will be realized lexically in a given language can not be predicted by our definitions, but definitions based on semantic prototypes do allow for this variation and make some predictions about where and when it might be expected. These two ways in which languages can differ in their treatment of certain types of meanings—particularly those which fall into the semantic domain of property concepts—are at the root of a great deal of typological variation in lexical class systems in the world's languages.

3.   Conclusion: The typology of parts of speech

One of the most important aspects of the approach to defining parts of speech outlined in the previous section is the idea that lexical classes are neither strictly semantic nor strictly syntactic. Instead, both the semantic properties of words and syntactic information about their unmarked distribution are made use of in the lexicon to subdivide words into various classes. By positing parts of speech systems as a part of the knowledge speakers have about the lexical inventory of their language, it becomes possible to delineate a constrained set of organizational criteria—some syntactic, others semantic—which speakers can use to divide and subdivide words into sets for the purposes of building syntactic structures. Cross-linguistic variation in these systems can then be characterized in terms of variation in the criteria used to shape the lexicon which fall into two separate categories—semantic and syntactic—and furnish us with the basis for criterial definitions of lexical classes. Languages that have a typical Indo-European style parts of speech system have three major lexical classes which can only be differentiated using both the semantic and the syntactic parameters. However, not all languages make the tripartite distinction between nouns, verbs, and adjectives in their lexical inventories, suggesting the possibility that these languages may not make use of both semantics and syntax in the organization of their lexica.

In terms of cross-linguistic variation in the three major parts of speech, there are four common inventory-types proposed in the literature (*e.g.* Schachter 1985; Croft 1991; Bhat 1994):[11]

(23) *full NAV inventory*—the lexicon distinguishes between three major word classes—noun, verb, adjective (*e.g.* English, Russian).

N[AV] inventory—the lexicon distinguishes only nouns and verbs, conflating property-concept words with verbs (proposed—Cora, Salish).

[NA]V inventory—the lexicon distinguishes only nouns and verbs, conflating property-concept words with nouns (proposed—Quechua, Totonac, Hausa).

[NAV] inventory—the lexicon conflates all three major lexical classes, making no distinctions at all (proposed—Tongan, Mundari).

As noted above, the asymmetry in these patterns—that is, the fact that when only one distinction is neutralized it is that between adjectives and some other part of speech—argues for the markedness of the adjectival class relative to the class of nouns and verbs. As we have seen, this markedness can be accounted for by Weak Iconicity, which requires that the unmarked syntactic role of the expression of a semantic predicate (the governor of its semantic argument) be that of a syntactic governor rather than a modifier (a syntactic dependant).

A four-member system of the type shown in (23) can be easily derived by reformulating our parameters in terms of whether the organization of a particular lexical inventory is driven by one or both of semantics or syntax:

(24) semantics: a language is said to be semantically driven if its lexicon distinguishes lexemes expressing semantic predicates from those expressing semantic names.

syntax: a language is said to be syntactically driven if its lexicon distinguishes lexemes that can be unmarked syntactic actants from those which can be unmarked modifiers.

Languages organize their lexica around either or both of these factors. This can be expressed as a feature system which derives the inventory-types in (25):

(25)   Types of lexical inventory

Semantically driven

		+	−
Syntactically	+	full NAV inventory	[NA]V
driven	−	N[AV]	[NAV]

In the case of full NAV inventory languages, then, the lexicon can be said to be organized on the basis of both semantic and syntactic factors, whereas in the case of an N[AV] inventory conflating verbs and adjectives, the lexicon is sensitive to purely semantic considerations. As a result, all predicates belong to the same lexical class and function as unmarked syntactic governors of their arguments, meeting the criterial definition of verbs.[12] This seems to be the case in Bemba (see (12) above) and the Salishan family of languages. Conversely, [NA]V languages would organize their lexica around the syntactic distinction between unmarked governor (of the expres-

21

sions of its semantic argument) and unmarked dependant (of the expression of its semantic predicate). However, finding examples of this type of language proves to be more problematic.

For a language to qualify as an [NA]V language, there are two possible patterns of markedness in the distribution of the putative conflated noun-adjective lexical class. The first, by far the more commonly referred to in the literature, is represented by Quechua, which shows the markedness pattern illustrated in (26):

Quechua

(26) (a) rikašaka…   alkalde+ta
         see:PST:1SG   mayor+ACC
         'I saw the mayor'

     (b) čay   alkalde   runa
         D   mayor   man
         'that man who is mayor'

     (c) rikašaka…   hatun+ta
         see:PST:1SG   big+ACC
         'I saw the big one'

     (d) čay   hatun   runa
         D   big   man
         'that big man'

<div align="right">(Schachter 1985: 17)</div>

In (26a) and (c), the words *alkalde* 'mayor' and *hatun* 'big' both surface unmarked in the role of actant, while (26b) and (d) show the same words acting as unmarked modifiers. According to Schachter, these examples indicate that words that fall into the classes of noun and adjective in English meet the definitional criteria of both in Quechua, filling the syntactic roles of actant and modifier without further measures.

Closer examination of this data, however, reveals two fundamental flaws in this line of reasoning. The first has to do with the assertion that the use of the noun *alkalde* 'mayor' as a "modifier" in sentences like (26b) is an unmarked one. On the semantic level, noun–noun attributive constructions like these are substantially different from the true modification structures illustrated in (20) in that they necessarily involve the presence of a semantic predicate which does not surface overtly in the syntactic structure. Consider the semantic representations of the English noun–noun attributive constructions in (27):
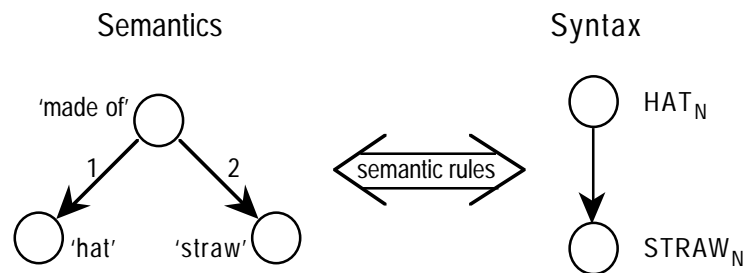
(27)   Semantic structure of noun–noun attributive constructions
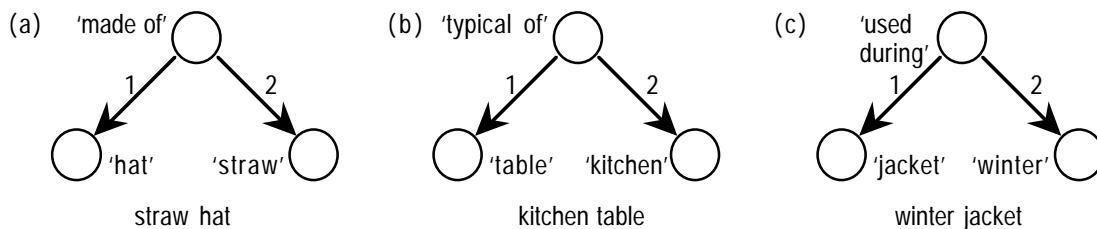
In each of these three examples, we see that the semantic representation of the phrase involves an elided two-place predicate, and in each of these cases the elided predicate is slightly different.[13] The phrase *straw hat* in (27a), for example, has the semantic structure 'made of'('hat','straw'), the nominal attribute specifying a material. In (27b), on the other hand, the nominal attribute specifies a type of table, and in (27c) it denotes the season in which the jacket is worn. Other common types of predication associated with noun–noun attributive constructions in English are instrumentality (*steak knife* 'used for'('knife','steak')), origin (*California wine* 'made in'('wine','California')), and part-whole relations (*computer screen* 'component of'('screen','computer')). The interpretation of such expressions—that is, which particular predicative relation linking the two nouns is reconstructed by the hearer—is typically pragmatically-driven (wines can not be made out of California, knives are not typically components of steaks, etc.) and tends to be phraseologized (*kitchen tables* must be high enough to sit at while eating, a windbreaker worn in January is not a *winter jacket*). When this is not the case, attributives are often open to more than a single interpretation (*e.g. alligator shoes*).

Noun–noun attributive constructions, then, are clearly marked in terms of cognitive complexity and, hence, non-iconic in that the underlying semantic predicate does not appear in syntactic representation. Consider the diagram of *straw hat* in (28):

(28)   Semantic ⇔ syntactic mapping for *straw hat*

The essential point to draw from this example is that the mapping between the semantics and the syntax of noun–noun attributive constructions is a fundamentally different one from that illustrated for adjective–noun constructions in (20). Frequently, although not always, the syntactic distinction between attributive and modifier is neutralized morphosyntactically in the grammar of a given language. Careful semantic analysis of the situation, however, reveals the presence of the underlying semantic predicate specifying the relationship between the two semantic names, neither of which is a semantic argument of the other. Although not realized lexically, the elided semantic predicate finds its expression in the syntactic subordination of the attributive noun, which shows multiple signs of decategorization including loss of referential meaning, loss of inflection, and the potential for multiple readings. Thus, by two of the three criteria for markedness outlined in (7) above (cognitive complexity and contextual markedness), attributive constructions are marked with respect to modificative constructions and illustrate the distinctiveness, rather than the neutralization, of the classes of adjective and noun.

A similar line of reasoning applies to the assertion that the use of the property-concept word *hatun* 'big' as an actant in sentences like (26d) is an unmarked one. As observed in all my sources on Quechua (Gradeja & Vela 1976; Cole 1985; Cerrón-Palomino 1987), such constructions are elliptical in that they are ungrammatical outside of a discourse context in which the identity of the object to which the property 'big' is attributed is recoverable (the same is true of

parallel structures in Spanish, Russian, Upper Necaxa Totonac, and almost certainly in all other languages in which they are possible). In this sense, the identity of the object referred to is semantically present and must be included in any complete semantic representation of the meaning of the sentence. The fact that the identity of this object is elided from the surface form of the sentence makes the adjectival-actant construction (like the attributive with its elided predicate) non-iconic and, by Weak Iconicity, marked. Thus, languages of the Quechua type do not qualify as [NA]V languages because of the inherent markedness or non-iconicity involved in using the expressions of semantic predicates as actants in elliptical expressions.

The other possibility for realizing an [NA]V lexical inventory is the opposite of the Quechua type: instead of showing a pattern wherein the putative conflated [NA] class consists of words that are unmarked both as actants and modifiers, languages of the second type would require that the [NA] class be *marked* in the syntactic role of modifier and unmarked in the role of syntactic actant. As we saw in the case of Quechua property-concept words, however, the requirement that property-concept words be unmarked actants precludes the use of adjectives in elliptical constructions as actants—in other words, property-concept words appearing as syntactic actants would have to have readings as expressions of abstract properties. To my knowledge, the only language in the literature that appears to meet this requirement is the Chadic language Hausa.[14] In Hausa, most property concepts are expressed by words, called *quality nouns*, which appear to be syntactically identical to more ordinary nouns in that they are marked as modifiers (see the examples in (11) above) and they function as actants in sentences like (29), in which they have abstract nominal readings:

> Hausa
> (29) k̓aöfi+n+sà   ya…   ba…   mù   ma..ma..ki…
>   strength+LNK+3MASC:PO   3MASC:CMP give 1PL amazement
>   'his strength amazed us' (lit. 'gave us amazement')
>
> (Ma Newman 1990: 9)

As promising as this appears, however, there are two difficulties with analyzing Hausa as an [NA]V language. The first is that, in fact, Hausa does have ten or twelve true adjectives which function as unmarked modifiers of nouns and have elliptical readings in actantial position:

> Hausa
> (30) (a) kà..wo… àlkyabbà…   bak̓a…
>   bring   burnoose_FEM   black_FEM
>   'bring the black burnoose'
>
>   (b) kà..wo…   bak̓a+r
>   bring   black_FEM+LNK
>   'bring the black one'
>
> (Cowan & Schuh 1976: 313–14)

Even though these true adjectives form a closed class in Hausa, the existence of a class of unmarked modifiers is essential for any treatment of Hausa quality nouns as marked in this same syntactic role. In the absence of a class of unmarked modifiers, it would be impossible to make the claim that quality nouns like those in (11) are marked modifiers: as noted in section 1.4.1,

markedness depends on the existence of contrast (see also note 8 above). Thus, the claim that Hausa is an [NA]V language is weakened by the existence of a small class of adjectives whose presence is needed for a proper characterization of the organization of this language's lexical inventory.

The second, and perhaps more convincing, difficulty with analyzing Hausa as an [NA]V language arises from the inherent non-iconicity of quality nouns themselves: given that a property-concept is a semantic predicate, the fact that its realization in the syntax is a non-predicate makes this realization non-isomorphic and, therefore, marked by dint of cognitive complexity.[15] In semantic representation, the definitive property of predicates is that they take arguments and prototypical syntactic predicates, as iconic expressions of semantic predicates, should also take syntactic actants. When quality nouns appear in actantial position, however, they frequently appear without an actant, as with *ma̲ma̲ki̲* 'amazement' in (29), or take the expression of their semantic actant as a possessor, as *k'aɓfinsà* 'his strength' (*cf. kuàinsà* 'his money')—in other words, quality nouns are syntactically non-predicative, although they seem to be the expressions of semantic predicates. Thus, the semantic structure of prototypical adjectives (property-concept words) prevents them from functioning as unmarked actants of verbs, just as the semantic structure of prototypical nouns prevents them from functioning as unmarked modifiers. Ultimately, what this reveals to us is that the typological system proposed in (23) above, despite its currency in the literature, is not entirely accurate: while there are languages that conflate adjectives and verbs, there does not seem to be even the possibility of languages that conflate adjectives and nouns. The fault, however, does not lie with the basic organizing principles and definitions that we used to make up the typology, but rather in the way in which our semantic and syntactic criteria have been combined in the feature system in (25) above: what we have uncovered is not an error in defining the organizing principles of lexical inventories, but a constraint on the way these principles are organized. Rather than having equal weight and therefore having equal opportunity to be the sole driving factor behind a parts of speech system, the semantic and syntactic criteria we have invoked here seem to be ranked, with semantics taking precedence over syntax. In terms of the three major class distinctions, the lexical inventories of the world's languages make the first (and sometimes only) division in their lexicon based on the semantic distinction between predicates and names. Three-class languages then make a further, syntactically-based, distinction setting apart those words in the class of semantic predicates that are treated as unmarked modifiers by the rules of the grammar. Because it is a secondary distinction—and because it results in words whose unmarked syntactic function is non-iconic—this criterion is cross-linguistically marked and, hence, frequently neutralized. Thus, our definitions not only accurately characterize the membership of lexical classes within and across languages, they allow us to account for the attested cross-linguistic variation in these classes and to properly constrain this variation. The analysis offered here afford us a unique insight into the fundamental organizing principles of the lexicon and show us that lexical classes, in particular adjectives, must be defined in semantic and syntactic terms, as artefacts of the semantics ⇔ syntactic interface.

Notes

[1] Astonishingly little typological work has been done on the class of adverbs, and in the absence of the necessary data I will have nothing further to say about them in this paper, although I think that adverbs can be dealt with quite comfortably in the same manner as the other three lexical classes discussed here.

[2] *q̓iyaƛ̓əd* 'slug' here is a common noun, although in this example it is also used as the proper name of an anthropomorphized slug with human characteristics. This human–animal ambiguity is typical of traditional Native North American narrative form.

[3] Alternatively, it could be argued that *tu-* is not a past tense marker because it appears on both verbs and nouns—but again this presupposes that words like *q̓iyaƛ̓əd* 'slug' and *sč̓istxʷ* 'husband' are nouns and therefore preclude the application of a true tense marker.

[4] On the noun-verb distinction in Pacific Northwest languages see also van Eijk & Hess (1986), Jelinek & Demers (1994), Beck (1995), and Matthewson & Davis (1995) on Salish, and Jacobsen (1979) on Salishan, Wakashan, and Chimakuan.

[5] Note that I have re-formulated Hengeveld's (1992a: 58) definitions—which in the original are couched in the terms of Functional Grammar (Dik 1978)—to make them more accessible to those unfamiliar with the framework. As pointed out to me by Igor Mel'čuk (p.c.), these definitions are highly reminiscent of the three "ranks" proposed by Jesperson (1928).

[6] In the exposition of Croft's work I will adopt his terminology for heuristic purposes. For the most part, these are not the terms I will make use of in the rest of this discussion.

[7] Tones and vowel length are not indicated in the original source. I have added these here based on the forms given in other sources of Hausa data, for the most part Kraft & Kraft (1973), Cowan & Schuh (1976), and Ma Newman (1990). The interlinear glosses are also mine, based on the same sources.

[8] The hierarchy also makes the claim that there are possibly languages that have only verbs (Tuscarora being Hengeveld's example of a language of this type). In the approach being developed

here, based on markedness of lexical items in given syntactic roles, this situation is a logical impossibility: markedness is based fundamentally on contrast—that is, an opposition between marked and unmarked elements. In order to declare that a set of words is marked in the role of actants (*i.e.* that they are not nouns), it is necessary to be able to contrast the morphosyntactic properties of this set with the properties of another set of words which is not marked in this role. This latter set would therefore qualify as nouns. In the absence of an unmarked set, it is not possible to establish a contrast and we can not speak of markedness to begin with.

[9] The determiner is deliberately left out of this tree for simplicity of presentation.

[10] The issue of whether the grammatical distinction between $cold_N$ and $cold_A$ in English (or $cold_A$ in English and $lon\,{}^{.}_N$ in Totonac) also corresponds to a semantic distinction (semantic name vs. semantic predicate) is a contentious one and there is no space for us to go into it here. It is my own feeling, following Langacker (1987) and Wierzbicka (1986, 1988), that there is a subtle semantic shift involved, so that $cold_N$ is to $cold_A$ as $square_N$ is to $square_A$ and as *circle* is to *round*. Unfortunately, further discussion of this issue is largely peripheral to the purposes of this article.

[11] The existence of the last of these inventories is highly controversial and there have been a number of interesting proposals made for them in the literature. Some of the more widely cited examples are Salishan (*e.g.* Kuipers 1968; Kinkade 1983), Tongan (*e.g.* Broschart 1997), Mundari (Bhat 1994), and Tuscarora (Hengeveld 1992a). While I am skeptical about these claims (see Beck 1999), I will not attempt to deal with [NAV] inventories here.

[12] Of course, this also means that all expressions of predicates in some N[AV] languages may meet the definitional criteria of adjectives (*i.e.* that they are also lexical items expressing semantic predicates which can be unmarked syntactic dependants of lexical items expressing their semantic arguments). This raises the rather trivial issue, familiar from phonology, of what, in the case of the neutralization of the distinction between two elements, to name the resulting conflated category. I have chosen here to make use of the name of the less marked of the two categories (*i.e.* verb) principally because the category would contain meanings such as 'hit' or 'devour' which would otherwise be unattested as adjectives in any other language type.

[13] The numbers associated with the arrows in the diagram are simply a means of indexing a particular argument with one of the semantic roles specified by the predicate; in this sense, it is equivalent to the ordering of arguments in a formalism such as 'made of'('hat','straw') (= *straw hat*), the inverse of 'made of'('straw','hat') (? = *straw made of hats*).

[14] Another possibility is Motu, which Wetzer (1996: 176*ff*) describes in the same terms as Hausa.

[15] An alternative analysis of the Hausa situation, one which I believe will ultimately prove to be the correct assessment, is that quality nouns are not, in fact, the expressions of semantic predicates at all but correspond to abstract nouns (see note 10 above). This would not only explain their syntactic behaviour as marked modifiers and unmarked actants, but also accounts for the fact that in terms of quantification they pattern with nouns rather than verbs:

<u>Hausa</u>

(i)   (a) aikì…   mài   wùya…   sò..sai
              work   ATRB   difficulty   much
              'very hard work'

        (b) bafl   ni   dà   kuài…   sò..sai
              NEG   1SG:CONT   CMT   money   much
              'I don't have a lot of money'

        (c) ba…   yà…   aikì…   dà   yawà…
              NEG   3MASC   work   CMT   quantity
              'he doesn't work much'
              (lit. 'he doesn't work with quantity')

                                (Ma Newman 1990: 295)

Given that this analysis is highly controversial and that it bears only on the situation in Hausa and does not address the typological issue of the possible existence of [NA]V languages, it is not the one I have chosen to pursue here.

to appear. In L. Wanner (ed.), *Advances in Meaning-Text Theory*. Amsterdam: Benjamins.

References

Bates, Dawn, Hess, Thomas M., & Hilbert, Vi. (1994). *Lushootseed dictionary*. Seattle: Washington University Press.

Beck, David. (1995). A conceptual approach to lexical categories in Bella Coola and Lushootseed. In *Papers for the 30th International Conference on Salish and Neighboring Languages*, 1 – 31. Victoria: University of Victoria.

Beck, David. (1999). The typology of parts of speech systems: The markedness of adjectives. Ph.D. dissertation, University of Toronto.

Bhat, D.N.S. (1994). *The adjectival category*. Amsterdam: Benjamins.

Broschart, Jurgen. (1997). Why Tongan does it differently: Categorial distinctions in a language without nouns and verbs. *Linguistic Typology* 1, 123 – 65.

Cerrón-Palomino, Rodolfo. (1987). *Lingüística quechua. Biblioteca de la tradición oral andina #8*. Cuzco: Centro de Estudios Rurales Andinos "Bartolomé de las Casas".

Chafe, Wallace L. (1970). *Meaning and the structure of language*. Chicago: University of Chicago Press.

Chomsky, Noam. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.

Cole, Peter. (1985). *Imbabura Quechua*. London: Croom Helm.

Cowan, J. Ronayne, & Schuh, Russell G. (1976). *Spoken Hausa*. Ithaca, New York: Spoken Language Services.

Croft, William. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Dik, Simon C. (1978). *Functional grammar*. Amsterdam: Elsevier.

Dixon, R.M.W. (1982). *Where have all the adjectives gone?* Berlin: Mouton.

Dolinina, Inga, & Beck, David. (1998). The dual unselectivity of adverbial-quantifiers: The case of Bella Coola. In Sheila Embleton (Ed.), *The Twenty-Fourth LACUS Forum*, 79 – 89. Chapel Hill: Linguistic Association of Canada and the United States.

Fauconnier, Gilles. (1985). *Mental spaces*. Cambridge: MIT Press.

Givón, Talmy. (1979). *On understanding grammar*. New York: Academic Press.

Grajeda Challco, Braulio, & Vela Flores, Asís Orlando. (1976). *Gramática quechua: Enciclopedia de gramática quechua integral*. Lima: Instituto Superior de Quechua del Perú.

Haiman, John. (1980). The iconicity of grammar. *Language* 56, 515 – 40.

Hengeveld, Kees. (1992a). Parts of speech. In Michael Fortescue, Peter Harder & Lars Kristofferson (Eds.), *Layered structure and reference in a functional perspective*, 29 – 56. Amsterdam: Benjamins.

Hengeveld, Kees. (1992b). *Non-verbal predication: Theory, typology, diachrony*. Berlin: Mouton de Gruyter.

Hess, Thomas M. (1993). *Lushootseed reader with introductory grammar: Volume I—Four stories from Edward Sam*. Victoria: Tulalip.

Hopper, Paul, & Thompson, Sandra A. (1984). The discourse basis for lexical categories in universal grammar. *Language* 60, 703 – 52.

Jacobsen, William H. (1979). Noun and verb in Nootkan. In Barbara Efrat (Ed.), *The Victorian Conference on Northwestern Languages, Heritage Record* #4, 83 – 153. Victoria: British Columbia Provincial Museum.

Jelinek, Eloise, & Demers, Richard A. (1994). Predicates and pronominal arguments in Straits Salish. *Language* 70, 697 – 736.

Jesperson, Otto. (1924). *The philosophy of grammar*. London: Allen and Unwin.

Kinkade, M. Dale. (1983). Salishan evidence against the universality of "noun" and "verb". *Lingua* 60, 25 – 40.

Koenig, Jean-Pierre. (1999). *Lexical relations*. Stanford: Center for the Study of Language and Information.

Kraft, Charles H., & Kraft, Marguerite G. (1973). *Introductory Hausa*. Berkeley: University of California Press.

Kuipers, Aart. (1968). The categories of verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21, 610 – 26.

Lakoff, George. (1965). *On the nature of syntactic irregularity. Report No. NSF-16, Mathematical Linguistics and Automatic Translation to the National Science Foundation*. Cambridge: Computation Laboratory of Harvard University.

Lyons, John. (1977). *Semantics*. Cambridge: Cambridge University Press.

Ma Newman, Roxana. (1990). *An English–Hausa dictionary*. Yale University Press.

Matthewson, Lisa, & Davis, Henry. (1995). The structure of DP in St̓át̓imcets (Lillooet Salish). In *Papers for the 30th International Conference on Salish and Neighbouring Languages*, 55 – 68. Victoria: University of Victoria.

Mel'čuk, Igor A. (1988). *Dependency syntax: Theory and practice*. Albany: State University of New York Press.

Mel'čuk, Igor A. (1995). The future of the lexicon in linguistic description and the Explanatory Combinatorial Dictionary. In Lee, I.-H. (Ed.), *Linguistics in the morning calm 3*, 181 – 270. Seoul: Hanshin.

Pustejovsky, James. (1996). *The generative lexicon*. Cambridge: MIT Press.

Saussure, Fernand de (1916). *Cours de linguistique générale*. Paris: Payot.

Schachter, Paul. (1985). Parts-of-speech systems. In Timothy Shopen (Ed.), *Language typology and syntactic description, Volume I: Clause structure*, 3 – 61. Cambridge: Cambridge University Press.

Taylor, John R. (1989). *Linguistic categorization: Prototypes in linguistic theory*. Oxford: Clarendon Press.

Thompson, Sandra. (1988). A discourse approach to the cross-linguistic category "adjective". In John A. Hawkins (Ed.), *Explaining language universals*, 167 – 85. Oxford: Blackwell.

Trubetskoy, Nikolai S. (1969) *Principles of phonology*. Berkeley: University of California Press.

van Eijk, Jan P., & Hess, Thomas M. (1986). Noun and verb in Salishan. *Lingua* 69, 319 – 31.

Wierzbicka, Anna. (1986). What's in a noun? (or: How do nouns differ in meaning from adjectives?) *Studies in Language* 10, 353 – 89.

Wierzbicka, Anna. (1988). *The semantics of grammar*. Amsterdam: Benjamins.

Žolkovskij, Aleksandr K., & Mel'čuk, Igor A. (1965). O vozmožnom metode i instrumentax semantičeskogo sinteza. *Naučno-texničeskaja informacija* 6, 23 – 28.