Word classes and related issues

Lajos Marosán

The following essay is part of a larger enterprise which purports to overview topics in linguistics that are not usually — or not sufficiently — covered in standard courses. In this essay I will be making extensive use of the material offered by the special, thematic issue of *Lingua* on word classes (see the references for details).

0 The problem

According to long-standing tradition, words belong to different classes, i.e., each word can be labelled either as NOUN, ADJECTIVE or VERB etc. Word classification makes two basic assumptions:

(1) a. There is a formally recognizable unit of sentence analysis to which the general term WORD may be applied which may be greater than or — as a limiting condition — identical to, the morpheme.

b. The words share certain grammatical — morphological, syntactic and, perhaps, semantic — characteristics on the basis of which it is possible to put them into classes. By classification a certain economy of grammatical description can be achieved.

Whether a particular item belongs with one class or another — according to tradition — depends on three factors: (i) form, (ii) syntactic position and (iii) type of meaning. In the following paragraphs I would like to discuss and comment on these points. As the reader will see none of these approaches yields the desired result which is to put each word into one class. Further, the different aspects of classification often clash: a word which is morphosyntactically identified as a noun, looks a verb from the semantic aspect. For instance, *discussion* is a noun since it occurs in certain syntactic positions, but as it expresses ACTIVITY it has the same semantic property as some verbs. As the reader will also see, no solution has been offered either by the author himself or the literature quoted in this paper. A solution along the lines reviewed and discussed on the following pages cannot be expected. What seems to be needed is an approach which ignores the linguistic facts in the sense that extralinguistic theory serves as basis for hypotheses. For example, Brøndal's theory (1948) of the parts of speech is a promising start basing a theory of word classes on logical grounds but, unfortunately, he does not go far enough. In this paper I do not wish to discuss Brøndal's theory; nor do I intend to suggest a new theory.[1]

This essay contains seven sections; in the first I discuss the problem of wordhood, then the traditional approaches to the problem: the phonological, morphological, syntactic and the semantic principle. I also address the question of universalism in relation to word classes and a notorious issue: multiple membership.

1 The word

The notion PARTS OF SPEECH is as old as the study of language itself. The term itself, *parts of speech* is the English equivalent of the Latin translation (*partes orationis*) of the Greek *μέρη λόγου* or *τὰ μέρη τοῦ λόγου.* The problem of word identity—which serves as the basis of the procedure of setting up the system of the parts of speech, i.e., word classes — did not occur to the ancient philosophers discussing different topics concerning language, and it does not seem to be a plausible problem for students of well-documented languages with considerable academic backgrounds. Even in these languages there are elements whose status as words are not absolutely clear, mainly function words, even if we apply the five tests of word identification which are widely known for students of linguistics:

(2) a. POTENTIAL PAUSE: speakers stop at word boundaries rather than in the middle of a word
  b. INDIVISIBILITY: a word is said to be indivisible, i.e., if new elements are added to the phrase or sentence, these are inserted between words already present in the structure
  c. MINIMAL FREE FORM: Bloomfield's widely known claim that a word can occur on its own right in the structure
  d. PHONETIC BOUNDARY: a word constitutes a phonological or phonetic unit
  e. SEMANTIC UNIT: a word is supposed to be a semantic unit; obviously *dog* is one. But if the definite article or a preposition count as

[1] Work on "Brøndal's theory of the parts of speech" is in progress in which I wish to give a (sketchy) outline of a possible new theory.

semantic units since they are commonly recognized as words in different grammars and dictionaries, so does a locative suffix in Hungarian—in fact any other suffix—such as *ban/ben* or *hoz/hez/höz.*

In Hungarian and Finnish stress can identify words (criterion (2d)): the main stress falls on the first syllable of a word, thus the portion of speech between two primary stresses might be analysed as a word. This characteristic might only delimit so called content words, therefore, this method is not always helpful. Hungarian articles never receive stress, so in connected speech they are never identified as words by stress alone; still, we would like to see them as separate words.

(3) ˈPuskin**a**ˈköltő
'Pushkin the poet'

The definite article obviously passes test (2b) but fails (2a): even in deliberate pronunciation it makes up a unit with the following noun. Or it gets associated with the previous word if we take the criterion seriously that stress can delimit words, as shown by the example above. Whether the article is a semantic unit (criterion (2e)) in the same way as a content word, such as *dog,* is a question of the analyst's taste, and as to criterion (2c): if the item in question passed this test, there would be no need for the other four.

There are also reports (Crystal 1987 : 91) that in Welsh words are stressed on the penultimate syllable ˈ*cartref* 'home' ∼ *car*ˈ*trefi* 'homes'; in Turkish words can be identified by vowel harmony: vowels within a word are either palatal or velar. (The case is complicated by the neutral vowel and compound words whose constituents do not harmonize with each other.)

Moreover, a definition of word which can be applied universally has to face the apparent typological differences between languages. In languages which are traditionally characterized as polysynthetic, such as Eskimo, Lakota and some Australian aboriginal languages, the word seems to be a highly complex unit. For instance, the form *palyamunurringkutjamunurtu* 's/he definitely did not become bad' is reported to be a word of an Australian language (Crystal 1987 : 91). In contrast to this form in, say, English or Hungarian a word is a morphologically much less structured unit in the sense that fewer syntactic relations can be expressed in it simultaneously.

What counts as a word, however, also poses serious difficulties in Bantu languages. Van Wyk (1967 : 230f) reports that the hesitation concerning how much language makes up a Northern Sotho word is reflected in the three orthographical traditions known as (i) disjunctivism, (ii) conjunctivism and (iii) semi-conjunctivism. More specifically, only some elements — items that could as well be considered as prefixes — cause the difficulty, the large bulk of the vocabulary is much easier to handle. Van Wyk's example, in (4), illustrates the orthographical and analytical problem with a sentence which translates as 'we shall skin it with his knife':

(4) a. re tlo e bua ka thipa ya gagwe
b. retloebua kathipa yagagwe
c. retloebua ka thipa ya gagwe

The segmentation of this sentence into words follows the three abovementioned approaches to Bantu words: (4a), (4b) and (4c), respectively. In other words, if the elements *re, tlo, e, ka* and *ya* and similar items are considered independent words, each should be assigned to a particular class and then the obvious question arises to which class; if not, they can be considered as morphological features of certain words (see the discussion in section 3). In the latter case grammatical characterization will suffice.

Yokuts is an American Indian language spoken in south central California, of which the most well-researched dialect is Yawelmani. Newman (1967 : 182f) suggests phonological criteria alone (see test (2d) above) help identify the word fairly clearly. With some exceptions it is either stressed on the penult; or on the antepenult if (i) the three-syllable plural noun of an ablauted stem has a long vowel in the first syllable and (ii) if the word is a verb with the durative present *-xo'* suffix. Also each word must begin with a single consonant followed by a single vowel and end in a single consonant or a single vowel (CV, CVC). A two-consonant cluster can only occur in word medial position. As can be seen, considerable knowledge is necessary to be able to define what a word is in this language; this also means that the five criteria in (2) are not necessarily field work discovery procedures.

Hoijer recommends a pragmatic (see criterion (2a)) rather than strictly speaking linguistic definition of the Navaho word: a word is "a form that can stand alone as a sentence and which is, in the deliberate pronunciation of a longer sentence that contains it, separated from other words by pauses" (1967 : 88). As to the first part of Hoijer's statement: there must be elements which the author would like to see as words but will never occur as sentences alone due to semantic or pragmatic reasons and (as to the second part of the quotation) there might also be words which would never occur separately even in deliberate pronunciation (cf. the Hungarian article) and still Hoijer would like to see them as words. These critical points, however, concern the method itself, and do not wish to invalidate Hoijer's claims about Navaho word classes.

As Kratochvíl reports the status of the Chinese word as a unit and its relation to the morpheme, on the one hand, and syntactic constructions, on the other, has been the topic of scholarly discussions for several decades. Thus, he formulates an ad hoc definition which, as we have seen above, only covers a section of the vocabulary. "In the following discussion the term 'word' will be used for any MSC [i.e., Modern Standard Chinese — LM] form which may occur as a minimal syntactic immediate constituent of a sentence, that is a constituent which a given sentence may be reduced to and still remain a grammatical sentence in the given context, or a constituent which may be eliminated in the process of reduction" (1967 : 132). Modern Standard Chinese offers illustration for a situation in which phonological criteria can only partially define words: the majority of the morphemes are associated with a tone. There are, however, atonic (=toneless) morphemes which can be seen as negative indicators: these can never be words only affixes attached to bound or free forms. As we will see later Chinese is intriguing because each one of the criteria ever devised to delimit parts of speech yields only partial results so the conjunction of these tests are necessary.

Carnochan explains that the criteria for wordhood in Igbo are not at all clear. The following quotation might shed some light on the state of the art as to what counts as a word in Igbo. The author writes: "On what grounds is the Igbo response *isi ḿ* 'my head', for example, analysed as one word or two? If as one, the other forms, *isi yá* 'his head' and so on, will also be one, and the noun will have to be considered as a unit with morphological endings in the category of person. If as two, then the unit will have two items, noun and pronoun. Since other items may come between noun and pronoun, they are established separately" (1967 : 3).

To conclude this section: it seems that the criteria of delimiting words as constituents vary from language to language; as we will presently see so do methods defining word classes in languages.

2 The phonological principle

There is a tradition according to which word classes can be defined on the basis of form, i.e., words can be classified into separate form classes according to what endings they have and/or what affixation and/or markers they receive. Words may possess some recurring phonological or morphological characteristic(s) that could naturally divide them into classes. In this way in languages which have rich morphology analysis can separate off different classes, such as verbs, adjectives, adverbs etc. For instance, in Russian the class of adjectives can be determined morphologically on the basis of the typical variable ending they show when they have the role of premodifier in noun phrases, for instance, *краткий/ая/ое/ие*. Also, nouns can be subclassified into gender classes according to the final phoneme of the word in question, e.g., *дом* 'house' is masculine, since it ends in a consonant but *плечо* 'shoulder' is neuter due to the final /o/.

As Kratochvíl reports in Chinese the phenomenon called *érhuà* by the Chinese linguistic terminology has often been utilized for word classification (1967 : 135f). This means the retroflectization of the final vowel of the last syllable of a given word. This is a typical feature of (some) nouns — in the European terminology — (like, e.g., *lǚr* 'donkey', *jīr* 'chicken') and (some) adverbs. In the case of adverbs, however, the *érhuà* is accompanied by the morphological feature of reduplication (e.g., *mànmānrde* 'slowly') so the two classes can be easily separated. Exceptionally, some other items also have this feature but they are easy to locate (e.g., *wánr* 'to have fun', *tăngtăngr* 'to lie down for a while'). The point is that retroflectization is an important relevant feature but not at all reliable as a criterion since there are other words lacking this feature, which could be analysed as nouns on other grounds.[2]

According to Thalbitzer (1911) in West Greenlandic Eskimo only nouns can end in /p/ since *-p* is the suffix of the relative case; no stem ends in /p/.

(5) arnaq 'woman'
arnap 'of the woman'

Therefore, this morphophonemic feature defines nouns unambiguously unlike the *érhuà* in Chinese. Also, both what we would call nouns and verbs but not adjectives and other parts of speech can receive the same endings, for instance:[3]

[2] Beside retroflectization there are other phonemic features which are of some interest for word classification, such as stress patterning of polysyllabic words and the position of words in the stress patterns of sentences. None of them, however, yield more result than the *érhuà*.

[3] The examples are taken from Schultz-Lorentzen (1945).

(6) uvdloq 'day'
uvdlut 'days'
nerivoq 'eating/he eats'
neripput 'eating-pl/they eat'

This latter situation is true of Hungarian to a (more) limited extent:

(7) [$_N$ház]unk 'our house'
[$_V$fáz]unk 'we freeze'

Also, Nádasdy & Siptár (1994 : 155ff) report that adjectives in Hungarian are mostly so-called lowering stems, which means that the linking vowel between the (consonantal) stem and (some) case endings is low. For instance:

(8) A szárnyas$_N$+o+k szárnyas$_A$+a+k.
the winged+linking V+pl winged+linking V+pl
'The winged ones have wings.'

As the not too ingenious example above shows, the same stem can be a noun (*szárnyasok*) requiring a mid vowel before the plural marker while the word as an adjective must be suffixed with a low vowel. There is also a great number of nouns and pronouns which have this feature. The class of lowering stems is closed so new items, such as foreign names, are always suffixed with a mid vowel, therefore the "regular" noun stems require a mid vowel. There are no lowering verb stems.

In Northern Sotho (van Wyk 1967 : 246) interjectives and ideophones have phonemic structures which are different from the phonemic make up of the members of the other word classes. Words[4] may normally end in vowels or the velar nasal syllabic consonant; for instance, *mo-tho* 'person', *no*-[ŋ̩] 'vulture'. Many ideophones end in other than the syllabic nasal: *tserr* 'representative of heat'. Polysyllabic words always have length on the penultimate syllable in prepausal or sentence final position: *go-bol*[ɛː]*la* 'to talk'. Ideophones and interjections typically have either short penultimate syllables or long final syllables, as shown in (9):

(9) isokohlo 'ouch'
phalakgatha 'representative of sudden flight'
thwii 'representative of straightness'
hubee 'representative of redness'

[4] Cf. the discussion of word identity in Northern Sotho above.

It follows, then, that on these grounds the words of the language can be divided into those with normal and those characteristic of paranormal phonemic features.

To conclude this section, we may state that phonological features can be exploited to different extents for setting up word classes: in some cases certain features can define a class or subclasses, as in Eskimo and Russian and can also be used as discovery procedures while in other languages, such as Hungarian, though some phonological feature delimits a class, this information can only be discovered after a careful examination of the language.

3 The morphological principle

Morphological features can be useful in distinguishing word classes. Though in Hungarian some inflectional endings can be attached to nouns, adjectives and verbs, case endings can clearly differentiate nouns from verbs. Also, the phonologically identical suffixes (see the example in (7)) receive different interpretations: possessive personal suffixes on nouns and adjectives, and verbal inflectional suffixes.

Daniels claims (1967) that in Japanese words are naturally ordered into parts of speech due to their inflection, i.e., paradigm. He thinks that the real problem arises when one wishes to classify uninflected items of the vocabulary.

Matthews (1967 : 174) divides the Latin vocabulary into two large classes on the basis of inflection. The first comprises five morphological classes:

(10) a. uninflected words or particles
  b. nouns (inflected for number and case)
  c. adjectives (inflected for — at least — number, case, gender)
  d. participles (inflected for number, case, gender, type)
  e. verbs with three subtypes:
    i. finites (inflected for mood, tense, number, person, voice)
    ii. infinitives
    iii. imperatives (inflected for mood, voice, number)

The second contains different sorts of pronouns: (a) pronominal noun forms, such as *te,* which are also said to be inflected for person, number and case; (b) pronominal adjectives, such as *hic* and adverbs, such as *uehementer,* inflected for grade alone. In what way(s) this second group differs from the first and why the distinction is necessary is not revealed.

Hockett (1958 : 221ff) reports that in Nootka, an Amerindian language, there is a bipartite system. Words fall into two groups: inflected and uninflected; all inflected stems have the same range of inflectional possibilities, in other words, they can take the same affixes. So the same stem can occur as a noun when uninflected and as a verb when inflected (the words in (11) receive another translation: 'he goes', 'he is a man', 'he is large' and 'he does it at night').

(11) wała·k- 'going/a trip'
qo·ʔas- 'a man/person'
ʔi·h- 'a large thing'
ʔathija- 'the night time'

There is no report about how the uninflected members of the lexicon relate to the inflected part; in other words, Hockett does not inform his readers how the two word classes relate to syntactic positions or functions, i.e., how functions traditionally identified with adverbs, pre-, or postpositions and adjectives in well-documented languages are expressed in Nootka. Hockett also remarks that the Nootka system disproves the claim that there is a noun/verb distinction universally. The Nootka system does not prove or disprove anything unless we would like to see the categories NOUN and VERB as morphologically identifiable entitities. Cf. similar bipartite languages below.

We have no reason to doubt the validity of Hockett's analysis; still, the question arises whether his classification is not a result of the application of too few criteria, which might happen when one examines languages which are not easily accessible. Crystal warns that "too few criteria produces the alternative danger of underclassification — major classes with a very uncertain and miscellaneous constitution, lacking any readily perceivable homogeneity" (1967 : 30). The following remark, though related to Hockett's analysis of Nootka, is of a more general nature. Languages which have some inflectional system also have invariable items (for instance, English, Hungarian, Navaho, Bilin, Yokuts, Japanese); therefore, words naturally divide into two groups: inflected and uninflected. As will be shown, the class membership of uninflected words can be determined by what is referred to in the literature as the 'syntactic principle'. Apart from this fact, very little or no relevant grammatical information can be said about uninflected words so this division does not contribute to the formulation of a description.

Nootka is not a particularly well-known language in Central-Europe; we might, however, look at English in the same way for the sake of the discussion. Classifying invariable words together in English would yield a class whose members could be *must, then, the, sheep, ought* etc. This classification does not seem particularly insightful. Nor does it in Hungarian, Yokuts, Japanese or Russian as far as I can tell. The same applies to the inflected portion of the vocabulary: *slowly, tables, looked* and *themselves* could be classed together.

Digression 1: Universalism and parts of speech

Hockett used his findings in Nootka to arrive at a conclusion that is relevant to languages in general. When one surveys the literature, researchers seem to claim that the concept of word class (a.k.a. part of speech, form class, lexical category) and the categories (noun, verb, adjective etc.) associated with it are universal. Robins contends that "The words as isolated in the language may be shown to share numbers of grammatical characteristics, such that an economy of formal description can be achieved by grouping them into classes, about which certain statements can be made which are applicable to all members of a class and can be incorporated into the rest of the grammar. *Such an assumption is apparently borne out in every known language and may be presumed to be a universal feature of language design*"[5] (1967 : 211).

However, scholars' opinions differ as to whether it is possible to formulate definitions which are valid for all languages, or these are only empty categories with some intuitive content and languages differ as to what counts as a noun, adjective, adverb etc. Sapir remarks: "... no logical scheme of the parts of speech ... is of the slightest interest to the linguist. Each language has its own scheme. Everything depends on the formal demarcations which it recognizes" (1949 : 119).

Matthews (1967) is of the opinion that class names should have some abstract definitions regardless of the language; if these definitions are satisfied in the language to a reasonable extent, then these labels are not just empty terms.

The system of the parts of speech has been seen as closely related to the philosophical categories of reality, therefore linguists have not been interested in (or capable of) separating linguistic investigations from the logical analysis of reality. Hence the traditional statements that, for instance, the category NOUN typically denotes objects whereas verbs refer

[5] Emphasis added.

to different types of acts, while adjectives identify qualities etc. Furthermore, there is disagreement as to how many classes should be recognized. When the Greeks "invented" the parts of speech, they recognized eight word classes: nouns, pronouns, verbs, adverbs, participles, conjunctions, prepositions and the article. The Romans added a new class, the interjection.[6] Later nouns and adjectives were separated and numerals recognized but these basic categories have been in the air for centuries. Sometimes a new category is created to satisfy the needs of the language. For instance, there are languages in which there are no prepositions but postpositions (e.g., Hungarian, Finnish). In these languages grammarians might find it expedient to add this to the already existing inventory. Also, grammarians reinterpret these categories and consider one as a subclass of the other; for instance, verbs and adjectives are usually put into one class in Chinese and some other languages, while in Hungarian and Finnish adjectives and nouns belong to a supercategory identified as SUBSTANTIVES. In English some grammars consider the pronouns as a separate category, while others[7] regard them as a noun subclass. Such decisions are more or less motivated by evidence drawn from the language under discussion or the reader is referred to general principles or authorities on the subject, depending on the grammarian and the purpose of the book. A traditional school grammar invites its readers to accept the statements as facts while an academic grammar makes attempts at motivating and/or explaining its statements.

To conclude the section discussing the universal nature of word classes, let me introduce the reader to two opinions which doubt the existence and importance of the parts of speech in language analysis. Though the author of these pages basically disagrees with what follows, it is extremely difficult to put it to serious criticism since these statements are not accompanied by any actual analysis which did not make use of the concept of parts of speech; intuitions can be accepted or refuted but not criticized.

In the introductory paragraphs of his essay on Modern Standard Chinese Kratochvíl writes that "It is not presumed here that the system of word classes is necessarily part of the grammar of a given language, although the hitherto experience of general linguistics leads to the conclusion that all known languages contain features commonly associated with the concept of word classes" (1967 : 130).

[6] As there is no article in Latin, their grammars never had this category.

[7] E.g., Huddleston 1984.

In his research article on Northern Sotho van Wyk suggests that it would be wrong to insist that all languages have parts of speech, unless this has been proved by linguistic analysis (1967 : 234ff). It follows, then, that since there are no universal classes, no definition of word classes is possible, thus the parts of speech of every language—the author continues (if such a thing really exists in the language under discussion)—must be named and defined individually.

If one "checks" a language for word classes, one will certainly find them — at least, I cannot imagine another result of a serious investigation along these lines. Van Wyk seems to say the same as Sapir, he just sounds more fatalistic since he does and does not believe in word classes simultaneously.

⋆ ⋆ ⋆

Hoijer (1967) claims Navaho offers a case which is easy to discuss: words basically divide into two main classes: inflected and uninflected. Inflected items can be clearly classified into nouns and verbs on the basis of the type of inflection they receive. Other parts of speech can be defined syntactically, which also means that the "uninflected" class contains items whose only distinctive grammatical property is the lack of some marker. From a methodological point of view this means that at least two types of criteria can be applied in setting up Navaho word classes: morphological and syntactic. In a later section I will discuss the problem of how consistent the approach—as to criteria applied in the definitions—should or should not be.

A similar situation can be witnessed in Yokuts (Newman 1967): only nouns and verbs are inflected, the other members and, consequently, other word classes are invariable. Nouns are inflected for case, verbs for either tense, aspect or subordination. Both categories can take several "layers of thematization," i.e., nouns and verbs can combine with derivational suffixes, but only one, obligatory inflectional ending, which is not an unusual phenomenon in languages.

Bilin, a Hamitic language spoken in Eritrea, Ethiopia, is another language which only possesses two classes: nouns and verbs, which are morphologically identifiable. Palmer (1967) reports that although the morphology of Bilin is extremely complex—the total number of forms of a single lexeme can be over ten thousand, nouns can be associated with the inflectional categories of number-gender (treated as one category: sg. masc., sg. fem., pl. common) and case; verbs, on the other hand, are inflected for number-gender, person, aspect and tense. An interesting feature of the case endings that they usually appear on the last element of the noun phrase, which is not necessarily the head. Therefore, verbs can also carry case inflection, not only postpositive adjectives. (Adjectives are a subclass of nouns.) When the verb follows the noun, it occurs in the relative form (=relative clause); in such cases verbs can assume a specific inflection which makes them "ready" to adopt further case inflections. Genitive nouns (=nouns inflected in the genitive) typically appear in post head position; in such cases the noun is doubly inflected: for the genitive and for the required case. Triple inflection is not impossible either.

Another interesting claim of Palmer's article is that adverbs and prepositions constitute two subclasses of the noun. A preposition is analysed as a noun by virtue of the fact that it always occurs together with a genitive noun, which latter performs the function of premodifier. (Why does he dub this class prepositions, then?)[8] In Palmer's view we might get structures parallel to real NPs as in (12):

(12) demonstrative + gen + … + verb
nén *yədǻw* gən
'they are *all around me*'

nén *yəfán* gən
'these are *my brothers*'

This is one of the cases when the distribution of two different items overlap: I do not think that this case justifies Palmer's claim.[9]

The author does not mention whether there are uninflected elements of the vocabulary, which would not be a surprising fact.

The situation in Modern Standard Chinese is far more complicated. Kratochvíl claims that there are two types of affixes: word-formative and grammatical. The former typically combine with bound forms, though there are examples which show that words can also take these affixes. The latter can only constitute words with free forms (1967 : 136–148).

[8] In Navaho postpositions receive the same inflections as nouns; still, Hoijer (1967) treats them as separate categories because they perform different syntactic functions. This might be true of Bilin nouns and prepositions.

[9] In Hungarian a similar phenomenon can be witnessed: *körülött-em* 'around-poss._1sg. → around me', *yə-dǻw* 'poss._1sg-all around'.

Investigation shows that WORD-FORMATIVE AFFIXES, except for a few easily recognizable cases,[10] are noun suffixes. These suffixes have different degrees of functional burdening. Some only occur in a handful of words, such as *-jiang* in *shíjiang* 'stone mason' (*shí* is a bound morpheme meaning 'stone'); others, such as *-r* (see above) or *-z* in several hundred. For instance, *háiz* 'child' from the bound morpheme *hái* meaning 'child'. Before anyone would jump to the conclusion that this feature, however limited it is since it only defines nouns, can be exploited for defining a word class, it must be mentioned that such affixes only identify approximately one-third of the words we would like to analyse as nouns since there are monomorphemic and polymorphemic words, too, not containing word-formative suffixes.

The presence of GRAMMATICAL AFFIXES is not only governed by the free morphemes they combine with but by the larger constructions in which they occur so theoretically they are direct word class indicators since they are part of the substitution frames on the basis of which form classes are established.[11] I write "theoretically" because the value of grammatical affixes for the purpose is greatly diminished by the fact that they are not obligatory in uniform cicumstances. This phenomenon is closely related to the question of definiteness/indefiniteness of MSC syntax, not yet fully explored. Still, MSC grammars utilize these grammatical affixes as class indicators in the potential rather than the actual sense. Kratochvíl refers the reader to a Chinese grammar:[12] "... words which ... may use the particles *le, zhe, guo* etc. ... to denote time-aspect changes", i.e., these affixes are used to set up subclasses of the verb. The only problem with statements like this is that, though they are intuitively correct, they have never been substantiated by sufficient data. Another type of morphological phenomenon which could be of great value as a class indicator is reduplication, i.e., a specific type of arrangement of (a) root morpheme(s) in a word. Reduplication is also connected to the system of definiteness but the only thing that is clear at the moment is the observation that most reduplicated words are definite. The most commonly known cases, however, are the ABAB and the AABB patterns of dimorphemic words belonging to the class of verbs, like in (13):

[10] For instance, *-ran* in *dāngran* 'of course'.

[11] Cf. the syntactic principle.

[12] *Zhànnǐ hànyǔ jiàohué yǔfǎ xìtǒng.* Page 16.

(13) kăolü →	kăolükăolü
'to think about something'	'to give something a thought'
ānjing →	ān'anjìngjìng
'to be quiet'	'to be very quiet'

The reduplication patterns above are often quoted as one of the main criteria according to which subclasses of verbs can be established.

To conclude this section on Chinese, it can be stated that morphology offers valuable as well as interesting clues for establishing word classes in MSC but, unlike in the languages discussed above, none of the features themselves are sufficient to define if only one class.

In Northern Sotho the fate of a classification based on a careful morphological analysis is also sealed. Van Wyk proposes some approaches[13] but only one of these looks serious. For instance, he suggests that words may be grouped according to whether they are morphologically simple or complex; this is almost the same as the inflected/uninflected dichotomy (though compound words are also complex).[14] Also, in his opinion a class should be set up for every morpheme. Just one critical point: the division of a vocabulary into parts of speech concerns words and not morphemes.[15]

In the proposal that is worth examining van Wyk claims that a classification could be based on MORPHOLOGICAL STRUCTURES, i.e., all words which occur, for instance, in the structures

subjectival concord + -a- + __ + -a[16]

and

class prefix + __[17]

[13] Van Wyk's strategy that no a priori approach to linguistics is possible forces him to evaluate logically possible methods and/or solutions most — or in a few cases all — of which result in impossible conclusions. What eventually happens is that he accepts a more or less intuitive (i.e., aprioristic) classification. His discussion, however, touches on relevant points.

[14] *Mo-sadi* 'woman' contains only one root, while *bo-sego-gare* 'midnight' two. On the other hand, the word that means 'truly', *ruri* is invariable.

[15] In this section he works with the notion of WORD but how much this "word" conforms to that of the definition in the opening paragraphs of his essay I cannot judge.

[16] E.g., *ke-a-tseb-a* 'I know', *ba-a-tsheg-a* 'they are laughing'.

[17] E.g., *ma-semo* 'fields', *ma-sengw-ana* 'small fields', *ma-semo-ng* 'in the fields', *ga-bo-hloko* 'painfully', *go-ma-tla-fal-a* 'to become strong(er)'.

should be classified into the same word class, respectively. The idea is similar to paradigm classes mentioned above (of Japanese and Latin), but the application of this principle to Northern Sotho yields controversial results since "in principle every substructure and every possible combination of morphemes is equally valid"[18] (1967 : 238). Therefore, the difficulty with this criterion is to decide which structural feature(s) should be chosen as basis for the classification. The condition of exploiting such morphological structures is to recognize the items surrounding the stem as morphological features, i.e., affixes, and not syntactic function words.

Van Wyk's verdict can end this section: "The conclusion is unavoidable, then, that purely morphological classification cannot be successful either with a restricted or a comprehensive purpose. If there are morphological features which are more pertinent to a classification into word classes, then these will have to be identified on other than purely morphological grounds" (*op.cit.* : 239). This last claim makes sense only if van Wyk wants to say that classes are set up on the basis of, let us say, syntactic position and then can be characterized by morphological features to various extents.

⋆ ⋆ ⋆

There are, however, languages with little morphology, in which it is practically impossible to find out anything about the class membership of a particular lexical item on the basis of its form. Therefore, a possible strategy would be to dispose of the category "part of speech" or "word class" in such languages. This is not the strategy I agree with, especially, because even a classless grammar has to account for why elements occur in the particular order they do and why an alternative arrangement is (or may be) ungrammatical. If we discard the notion of word class, we relegate the analysis of language into the realm of probability: the occurrence of each item will be unpredictably different from the other, no explanation or, at least, discussion will be possible of why *házunk* 'our house' can be preceded in any context by, let's say, *a* 'the' but *fázunk* 'we are cold' cannot. In other words, the notion of "word class" is part of the grammatical meta-language which helps us formulate statements about grammatical relations and restrictions exercised by groups of items rather than individual words even if the class membership of words is unpredictable as it is in a number of

[18] This statement is only sensible if it means that the same morphemes combine with elements which, intuitively, we would like to belong to separate classes.

languages, for instance, in English or Chinese; still a classification can shed some light on the workings of the language.

Van Wyk also addresses some questions of theoretical importance concerning the methods of classification. He examines the usual aspects — phonological, morphological, syntactic and semantic — and finds that inherently none of these approaches will yield word classes and he concludes that whether this or that approach is useful will eventually depend on the purpose of the analysis. An interesting feature of his article is that the author does find a phonological principle which inherently determine classes but rejects these findings saying that these classes do not help a description of this language. So his agnostic approach advertised in the beginning paragraphs of his essay — no parts of speech exist aprioristically unless found in the data, any approach is relevant depending on what direction the analysis takes — is refuted by the author himself in a later section: there is no possible analysis which would profit from a word class system based on phonological features, i.e., toneme patterns.[19] Indeed, because the notion of "parts of speech" is intimately related to a certain approach to language which, however strange this may sound, has not changed since the first attempts were made in language analysis. Scholars have always been interested in the workings of language, had the belief in the orderly nature of language, been convinced that language is accessible for investigation and studied language trying to base hypotheses on and work out methods within language. These constant features, then, delimit the scope and also the purpose of the investigation. Thus van Wyk's statement that classification depends on the "purpose of investigation" is meaningless since it has always been limited in the sense outlined above. No wonder, then, that van Wyk finds his own classification based on toneme patterns useless since it does not contribute to the purposes closely associated with the idea of word class; in other words, there appear to be no rules for which this kind of classification would have any relevance.[20]

[19] He sets up 7 classes: 1–2: HL/HL; 3–4: LL/LL; 5: LH; 6–7: LHL/LHL. Groups 1 and 4 do not while groups 2 and 3 do change this pattern in some contexts. Group 5 never changes while 6 becomes LH and 7 HL in certain contexts. (H=high; L=low.) See pp. 246f.

[20] In sections 7, 8 and 9 van Wyk does suggest an insightful classification of Northern Sotho words in spite of his negativistic remarks in the earlier sections of his essay though the way these classes relate to the problems discussed in the earlier sections is unclear.

In a language in which groups of words do not show some appreciable similarity in make-up it would be possible to set up as many classes as there are words of the language so each class would contain only one member; in such a case each class could carry its phonetic form as class name. For instance, the class to which the word *dog* belongs is labelled DOG. This might certainly be very appealing to the logician but does not lead anywhere as linguistic analysis, especially, because there is no grammatical, semantic or any basis for the classification, so this is not classification at all. In such a case the classes can combine in any order with no regularity. There might, however, be another approach to the above proposal. Individual items are examined against a background of a theory or some hypothesis; in other words, investigation is usually top down rather than bottom up. It might not be without any use, after all, to try what things look like if we scrutinize individual words and see whether they can be classified together with other words or they should remain the only members in their own classes. For instance, *house* and *say* are unique in morphology: in no other form do the [s] and the [ei] of the stem change to [z] and [e], respectively, before an affix. On these grounds alone, *house* and *say* could be assigned to a one-member class. Further, *wife* ~ *wives, half* ~ *halves*, i.e., words which are referred to as "voicing plurals" are also unique in this respect but they seem to be assignable into one class on the basis of this characteristic. A word-centred approach draws our attention to the differences rather than to the similarities and in this way helps to reconsider the age-old distinctions.

4 The syntactic principle

An analysis which does not recognize classes based on formal features mentioned above can rely on the notions of FUNCTION and/or POSITION in the sentence. In actual fact, the notion of "class" can be converted into statements of where particular items occur relative to each other, that is, it is possible to make use of the notion of word class stealthily without actually adopting the notion as a term in the metalanguage of the analysis. In such a case a distributional class would mean word class. In actual fact, function and form class seem to mean the same for Bloomfield "… our list of parts of speech will depend upon which functions we take to be the most important" (1933 : 269).

Distributional or substitution frames have been made wide use of to establish word classes on the basis of syntactic position. For instance, Kratochvíl (1967 : 131) suggests setting up three substitution frames based on the simple MSC sentence *wǒ yě qù* 'I also to go away' i.e., 'I'll go too':

X *yě qù, wǒ* Y *qù* and *wǒ yě* Z. The class X would consist of items, such as those in (14):

(14) *nǐ* 'you'
wáng bóshi 'Dr Wang'
gāngcái lái de nà ge rén 'the man who has just come'

I quote Kratochvíl's examples as illustrations for the method and I do not wish to discuss intimate questions of Chinese philology. We could examine English examples to illustrate the point. The structures that can potentially occur in the empty space in (15) make up a syntactic class.

(15) the ___ table
the [big] table
the [very big] table
the [kitchen] table
the [very big kitchen] table

It is obvious that the sequence of words *very big kitchen* cannot qualify for word status and even if this phrase suggests that *big* and *kitchen* are the same type of element, other structures will reveal that this is not so.[21]

The point is that these substitution frames define syntactic classes which are not necessarily identical to word classes; it requires further analysis to establish word classes, which may involve the examination of, for example, morphological features. Kratochvíl calls these FORM CLASSES, Daniels ROLE CLASSES, Matthews PATTERN CLASSES.

Newman sets up four syntactic classes which are clearly not identical to word classes: predicaters, substantives, modifiers and relaters and examines what types of word can realize these classes. The result is not at all surprising: except for predicaters, which can only be verbs, all the other classes can be realized by more than one word class.

Van Wyk also agrees that words can be classified according to their ability to enter into word-group or sentence structures; following Reichling he terms this ability of words VALENCE and suggests that there are different types of valences, such as sentence, combinatory, nuclear vs. attributive valences (1967 : 239–244). This is basically the same idea as the substitution

[21] Van Wyk's "morphological structures" can be used as illustrations in this section as well if we consider the items preceding and following the stem as independent words and syntactic units.

frames of an earlier paragraph, only the former is a theoretical side while the latter the methodological side of the same coin.

In this section I would like to follow Bloomfield's venerable instructions according to which a form class "is determined ... by the structure & constituents of the form, by the inclusion of a special constituent (a marker), or by the identity of the form itself" (1933 : 268) and show what classes can be set up in English on a purely formal basis. English seems to steer a course between two extremes: some words are special in the sense that they can only be assigned to one particular class, others can belong to two or even three classes. The following classification is adopted from Hockett (1958).

(16) i. **N:** This class contains all the words that can combine with determiners, can occur as heads in attributive constructions and the resulting endocentric phrase can function as subject, object or predicate. (For instance, *strength, food, boy, danger, elephant.*) Hockett does not mention the inflectional properties of nouns.

ii. **A:** This class comprises words which inflect for degree (*nice* : *nicer* : *nicest*) or participate in equivalent inflectional phrases (*beautiful* : *more beautiful* : *most beautiful*). These words also combine with the suffix *-ly* but do not inflect in the same way as the members of class N do. (For instance, *long, beautiful, likely, short.*) In this case syntactic positions are missing.

iii. **NA:** These are items following both the N and the A patterns, such as *American, sweet, white, savage* etc. Though Hockett could have mentioned that $American_{\mathrm{A}}$ does not possess the same inflectional potential as e.g., *sweet*:

sweet + er/est but **more/most American*
sweet + ly **Americanly*

iv. **V:** These words are characterized by inflection (*-ing, -ed, -en* etc.). Uninflected ones like *must, ought* belong here because of the same syntactic patterning. These words have a characteristic syntactic position occurring before class N words.

v. **NV:** Stems which show both the noun and the verb pattern, such as *walk, love, eye.*

vi. **AV:** Items occurring either in positions where class A items do, or as verbs, such as *dry, thin, idle.*

vii. **NAV:** Some words may show the distributional and/or inflectional characteristics of three classes, such as *fancy, black, damp.*

viii. **Particles:** a ragbag of items which do not belong to any of the above. These words, however, may turn out to possess totally different syntactic features.

The analysis above shows that even though the classes were set up on syntactic grounds, inflectional (morphological) features are also utilized. If this did not happen, the analyst would be forced to put adjectives and nouns in the same class, which is fairly counterintuitive in English.

Digression 2: Multiple membership

Another important lesson to be learned from Hockett's classification is that there are items which possess the characteristics of two or even three[22] parts of speech, such as *walk* or *dry* or *damp.* Traditional grammars state that, for instance, *walk* can be used as a noun and as a verb. What is *walk* then? A noun or a verb? Or a third type of entity that can be "used" or "acts" or "has the force" now as/of a verb, now as/of a noun? There are different answers to these questions.

One traditional answer[23] is that *walk* belongs to two classes: N and V, therefore it is a multiple membership item; one of these, however, is more basic therefore the other word can be derived from it. The "basic" or "proper" role the word has is often taken to be the earliest one it is believed to have had, or considered to be conceptually more basic. For example, *bottle* as a noun is easier to regard, at least intuitively, as more basic than the verb *bottle* in the sense that the action *bottle* requires the existence of an entity which can hold liquids in a specific shape. In actual fact, the criteria on the basis of which grammarians define this membership priority is never explicated.[24] Crystal (1967) suggests a statistical method: the more frequent case should be considered basic. This suggestion would only yield appreciable results if there were a marked difference between the occurrences tested; if, let's say, the noun *bottle* would be more frequent than verb *bottle.* But how would Crystal solve the issue if the two were

[22] Kratochvíl quotes Dobson (1962 : xxi–xxii) on Early Archaic Chinese that all plerematic (i.e., content) words are operationally ambiguous.

[23] See for example Curme (1935).

[24] There is, however, a type of conversion in which priority is more than convincing: the head noun of the NP *a Shakespeare* occurring in the type of sentence illustrated by *This playwright is not a Shakespeare* is traditionally considered a common noun but its existence as such is logically dependent on the existence of *Shakespeare* as a proper noun.

more or less in balance? Also, Crystal's suggestion raises the issue what the relationship is between statistics and grammatical rules. In my opinion, we deal not with a statistical problem requiring calculations but with a theoretical question; furthermore, statistical considerations are incompatible with the concept of grammar the notion of the parts of speech is a metatheory of.

There is, however, a fairly clear case in which priority can be established. Some proper nouns reclassify as common nouns as in (17):

(17) $\text{Einstein}_{\text{PN}}$ was a great physicist
Her new boyfriend is not an $\text{Einstein}_{\text{CN}}$

In the first sentence the family name *Einstein* refers to a famous twentieth-century scientist thus the word is a proper noun while in the second clause the same sequence of phonemes is a common noun because the word refers to certain characteristics (e.g., intelligence) associated with the bearer of the name. In all these cases, the proper noun is more basic in the sense that it makes the common noun possible. However, the establishment of priority need not be connected to historical or pragmatic facts since we have a descriptive and not a historical linguistic problem. In an earlier paragraph I referred the reader to the pair $\textit{bottle}_{\text{N}}/\textit{bottle}_{\text{V}}$ and suggested that intuitively the object seems to be more basic than the action. In Hungarian there is a similar pair: $\textit{kapa}_{\text{N}}/\textit{kapál}_{\text{V}}$ 'hoe/to hoe'. We could argue in the same way. Historical facts, however, contradict intuition and logic: the verb was taken over from a Slavonic language and the noun was created analogically from the already existing Hungarian verb.

The same problem is sometimes referred to as CONVERSION, i.e., a word is transferred from one class to another. The other theoretically possible solution is to treat these words as belonging to two (or three) different parts of speech. The drawback of this analysis is that it ignores the obvious relation which holds between, for instance, $\textit{bottle}_{\text{N}}$ and $\textit{bottle}_{\text{V}}$ and has nothing to say about it. This latter approach, Daniels[25] reports, is applied in Bloch's Japanese grammar:[26] "Words that are alike in form but differ in syntactic function (i.e., belong to different word classes) are different; thus the adverbs *keredomo* 'nevertheless' and *to* 'so saying', which appear at the beginning of clauses, are different respectively from the particles *keredomo* 'although' and *to* 'thus' ..." (1967 : 76).

[25] Daniels calls multiple membership CLASS CLEAVAGE.

[26] No bibliographical details are given.

In Hungarian the type of conversion that is widely-known about English (i.e., zero derivation) does not exist; there is a great number of derivational suffixes which assist in changing the class membership of words. There are, however, a handful of words which reclassify athematically, e.g., those in (18):

(18) *zár* 'a/to lock'
nyom 'trail/to press'
tér 'space/to have room'[27]

Uralic languages are said to have had a category what is known as NOMEN-VERBUM in the literature. Therefore, these words are precious fossiles. They might combine with certain suffixes which do not disambiguate these forms out of context, as shown in (19):

(19) *zárunk* 'our lock/we lock'
nyomunk 'our trail/we press'

In other cases, however, the ending is the same but the vowel between the stem and the suffix may betray the class membership of the word, like in (20):

(20) *zárom*$_{\mathrm{V}}$ 'my lock' vs. *záram*$_{\mathrm{N}}$ 'I lock', but

Or the form takes part in another vowel alternation, illustrated by *tehén*+$\emptyset$ 'cow-nom.' $\sim$ *tehen*+*et* 'cow-acc.'. Thus the last item of (18) has the form *terünk* 'our place' as a noun, while *térünk* 'we have room' as a verb.

In English, just as in several other languages, many words show multiple class membership, such as *round, milk* etc., while there are many others which belong only to one class, such as *life, parallel* etc. It seems that the potential for convertibility stops at one point; for instance, some nouns[28] can be "used" (or reclassified, converted) as verbs and the other way round but numerous nouns resist conversion. Why should this be so? Convertibility potential is not necessarily a strictly speaking grammatical but rather a pragmatic matter. There might be concepts which do not need expression; for instance, following the traditional analysis it is possible to say that the noun *earth* is converted to the verb *to earth* to describe a technical detail in installing electrical appliances or an important feature of such equipment.

[27] This word occurs as a verb only in certain dialects.

[28] Conversion is also possible from one subclass to another of the same part of speech, such as *Hoover* $\rightarrow$ *a hoover.*

We might coin, on this analogy, the verb *to sky*—I have not seen this verb before—but what is the act it is supposed to describe? Either it does not exist (yet), or if it does, it is expressed in a different way.

5 The semantic principle

There is another possible and traditionally accepted approach to word classes: semantics. This section will be discussing the semantic elements, too, appearing in the definitions; a number of other issues will also be addressed.

Analysis might try to define certain common characteristics between words which can be shown to have the same semantic features. Researchers are not very optimistic — to say the least — concerning the chances of a semantic classification. Newman claims that "... semantics has not offered any systematic procedure for defining meaning classes which would apply to all the references in a language, and the pursuit of meaning in parts of speech has proved to be a snare and a delusion. Caution dictates that semantic criteria should be excluded from grammatical classifications" (1967 : 192).

The most serious criticisms have been levelled against the semantic component of the definitions. As we will presently see, the definitions do not apply consistent viewpoints since definitions contain morphological, syntactic and semantic elements. Also, the criticisms themselves can be criticized, too; they are sometimes just as shallow and empty as the definitions themselves. Some authors suggest that the main problem with the definitions is that they are not mutually exclusive and this feature comes under attack. However, some grammarians contend that definitions along these lines are simply impossible because word classes cannot be mutually exclusive. It seems that, in contrast to critics commenting on traditional definitons of parts of speech, Bloomfield (1933) does not find anything wrong with word classes overlapping and including each other instead of showing neat boundaries and being mutually exclusive. He must have thought of overlapping items and not overlapping criteria.

Gleason (1961 : 92ff) goes as far as saying that a meaning-based class should be distinguished from grammar-based class to avoid the endless confusion resulting from their discrepancy. His discussion of the topic, however, suggests that only classes based on common structural properties (i.e., morphological and syntactic) are useful for grammatical analysis.

The remarks on the issue of a possible semantic analysis of parts of speech sound negativistic, though a purely formal analysis also may also lead to discrepancies. Newman writes: "I know of no language in which analyses based on the two criteria [morphological and syntactic] would result in isomorphic or parallel classifications which could then be economically treated as a single system" (1967 : 193). He is of the opinion that the two approaches can be treated as two separate hierarchies which interact. The author finds it more consistent to keep these two separate rather than to combine them into a system. Syntactic and morphological criteria are entirely different: the former are defined in terms of the relations between words and word sequences, i.e., relations external to the words, while the latter are strictly defined by processes internal to the word.

Matthews thinks that parts of speech definitions should not be expected to be "... homogeneous in the sense that should be all notional or that they should be ... 'morphological' (as, for example, the fourfold classification in Varro) or all 'syntactic' ..." (1967 : 156). For instance, there is a class containing elements such as *decem, tres* which is defined as NUMERALS for notional reasons while a purely formal definition will do for the PARTICIPLES. He also claims that there is no reason to demand that word classes should exhaust the vocabulary of the language; Matthews considers Charisius' category *pandéktēs* the dustbin for items which do not fit the other classes.

On the other hand, studying the semantic properties of word classes may lead to interesting observations. I do not wish to go into details in this essay; I would only like to mention that, for example, Jespersen (1924), following earlier speculations, claims that nouns denote substances, therefore, they are associated with many distinctive features whereas an adjective only denotes a quality; so it seems to him that a noun is more special than an adjective in that there are fewer entities that can be subsumed under the substantive. An adjective can be used to refer to a much wider range of entities. But what about adjective–noun pairs, such as *tall–tallness*? *Tallness* is an abstract noun and Jespersen suggests that abstract nouns denote qualities, too, just as adjectives. At this point let me remind the reader of Gleason's suggestion to clearly distinguish between meaning-based and grammar-based classes. According to the former, some verbs and nouns expressing action should be classified together, such as *destroy* and *destruction,* and also some verbs, adjectives and nouns, which denote state, for instance, *know, knowledgeable* and *knowledge* or *love, fond* and *love.* What are the borders of such a classification? A definition that says that nouns denote objects while verbs actions is certainly intuitively correct in that not all nouns can express actions. For instance, *table* is one of those nouns. Critics are right when they say that nouns can also denote actions, whereas verbs can never refer to objects; but a criticism along these lines implicitily requires that nouns should be subcategorized according to meaning, such as "nouns of reporting", "nouns denoting objects", "nouns referring to qualities" etc. Such a subcategorization sometimes seems grammatically feasible: there are a handful of nouns that can take appositive clauses as postmodifier, such as *fact, report, idea* etc., while — it seems — a relative clause can always be attached to any noun. This is obviously a semantic difference between nouns. On the other hand, critics perceive the overlapping cases between verbs and nouns, i.e., that both *destroy* and *destruction* refer to an action, but they are not willing to appreciate the cases when this is not so. Thus the problem with the definitions is not that they refer to semantic properties.

In this section I will be examining the definitions Curme (1935) offers in his English grammar. The purpose of the following paragraphs is to see what the definitions of the parts of speech look like — since such definitions and the comments on them have been discussed so far — and to make some relevant observations.

Curme's definition of a noun ("A noun or a substantive is a word used as the name of a living being or lifeless thing: *Mary, cow* ..." (1935 : 1)) refers to the referential property of the class; morphological or syntactic properties are ignored.

"A pronoun is used instead of a noun" — Curme writes on page 7. In a world of XPs this definition sounds awkward, to say the least, but before we would rashly conclude that Curme speaks nonsense, there are two things to bear in mind: (i) the pronoun class is defined in terms of syntax, in other words, Curme seems to suggest that a pronoun can occur in the positions a noun can, i.e., they have the same distribution. The noun class was defined as a notional category, the pronoun as syntactic. (ii) In current syntax it is customary to identify the terms NOUN, ADJECTIVE etc. as word level categories, i.e., single-member items. In an earlier tradition, which Curme together with Bloomfield seems to follow, a structure which could potentially substitute a single noun was considered a noun.[29] In Bloomfield's terminology, then, an endocentric phrase was identical with its head or centre. In other words, no substantial distinction was made between a noun and a noun phrase.

[29] Curme's position leads to a paradox: the PP in *He is dancing in the kitchen* is said to be an adverb since it carries out the same function; still, it is considered a phrase so in the clause it is an adverb outside it it is a prepositional phrase.

The definition of the verb is partly syntactic and functional: "The verb is that part of speech that predicates, assists in predications ..." and partly discourse functional: "... asks questions or expresses a command: The wind blows, ..." (*op.cit.*: 63). No meaning or substitutability potential is invoked, nor distribution. A considerable amount of goodwill is needed not to identify non-verbs on the basis of "... assists in predications ..."; Curme must have meant to define auxiliaries in this way.

The adjective ("An adjective is a word that modifies a noun or pronoun, i.e., a word that is used with a noun or pronoun to describe or point out the living being or lifeless thing designated by the noun or pronoun: a little boy ..." (*op.cit.*: 42)) receives an implicit syntactic and an explicit functional definition — at least, in my interpretation. "Modification" is a syntactic process, which obviously exerts some influence on the meaning of the modified item; but that seems to be the purpose of putting words together. The semantic properties of adjectives are not invoked in this definition. It would have been nice of Curme to illustrate the case when an adjective modifies a pronoun. Does this also mean that whatever can modify a noun is an adjective?

Basically, the same applies to the definition of the adverb ("An adverb is a word that modifies a verb, an adjective or another adverb" (*op.cit.*: 71)): it is related to adjectives, verbs and other adverbs without an explicit statement of what semantic, referential or formal properties it might have. Again, what does "modification" mean? Since Curme does not deem it important to make this relation explicit, it must be a formal relationship between two items.

"A preposition is a word that indicates a relation between a noun or a pronoun it governs and another word, which may be a verb, an adjective or another noun or pronoun: I live in this house." (*op.cit.*: 87). It seems that a new type of relationship is possible: government. The definition explicitly refers to some syntactic property; but whether the type of relation that holds between the noun and the adjective/verb is semantic or syntactic is not in the definition. The definition, however, does not account for several adverbial structures which Quirk et al. call CONJUNCTS, DISJUNCTS and SUBJUNCTS: some of these are introduced by a preposition but it does not indicate such a relationship since there is none: for instance, *in other words, on the other hand, in actual fact* etc. belong here.

"A conjunction is a word that joins together sentences or parts of a sentence: Sweep the floor and dust that furniture ..." (*op.cit.*: 92). Can "joining" mean the same as "indicating a relation"? If we put a coordinator between two clauses or phrases, we certainly create a relationship between them, which is indicated by the conjunction. This is a syntactic definition.

"An interjection is an outcry to express pain, surprise, anger, pleasure, or some other emotion as Ouch! ..." (*op.cit.* : 105). I do not think interjections express emotions; they are the emotions themselves. *Table* is the name of an object, *ouch* is the pain itself.

Adjectives and adverbs are defined in terms of modification, i.e., syntactically and semantically; prepositions and conjunctions as elements indicating a further not specified relation between items of different complexity, nouns in terms of reference. Verbs and pronouns are defined in syntactic terms, verbs again, along with interjections in discourse terms. An interesting feature of Curme's definitions is that the categories "noun" and "verb" are defined independently whereas the definition of all the other word classes depends on either that of the noun and verb or one or another part of speech. The lesson Curme's and other grammarians' definitions offer us is that there is no definite set of criteria which uncontroversially define all types of words into classes.

⋆ ⋆ ⋆

This short and somewhat sketchy outline has offered ample evidence that the issue of word classes is fairly complex and various questions need to be addressed. Is it possible to formulate definitions on syntactic and morphological grounds which are more or less adequate to describe a number of languages? I think a positive answer is possible. As we have seen semantic approach is only feasible for what are usually referred to as major word classes: adjectives, nouns and verbs. I also believe that there is a correspondence between the semantics and syntax of word classes, that is, once a certain number of items are characterized as nouns, they have certain stable semantic features. In another context, however, their class membership may change, so may their semantic properties but this should be the topic of another essay.

I hope to have shown that word classes are a many-splendoured thing which has not revealed all its secrets yet. This is a typically "traditional" topic not much discussed in the last couple of decades since there have been more fashionable topics around. Still, it is worth examining an age-old problem again in case something occurs which proves to be worth pursuing.

REFERENCES

Bloomfield, Leonard. 1933. Language. New York: Holt, Rinehart and Winston.

Brøndal, Viggo. 1948. Les parties du discours. Copenhagen: Einar Munksgaard. (French translation of the original Danish work, which appeared in 1928.)

Carnochan, J. 1967. Igbo. Lingua 17 : 1–23.

Crystal, David. 1967. English. Lingua 17 : 24–56.

Crystal, David. 1987. The Cambridge Encyclopedia of language. Cambridge: Cambridge University Press.

Curme, George Oliver. 1935. Parts of Speech and Accidence. Boston: Heath.

Daniels, F. J. 1967. Japanese. Lingua 17 : 57–87.

Dobson, W. A. C. H. 1962. Early Archaic Chinese. Toronto.

Gleason, H. A. 1961. An Introduction to Descriptive Linguistics. New York: Holt, Rinehart and Winston.

Hockett, Charles F. 1958. A Course in Modern Linguistics. New York: Macmillan.

Hoijer, Harry. 1967. Navaho. Lingua 17 : 88–102.

Huddleston, Rodney. 1984. Introduction to the Grammar of English. Cambridge: Cambridge University Press.

Jespersen, Otto. 1924. Philosophy of Language. London: Allen & Unwin.

Kratochvíl, Peter. 1967. Modern Standard Chinese. Lingua 17 : 129–152.

Matthews, P. H. 1967. Latin. Lingua 17 : 153–181.

Nádasdy Ádám and Siptár Péter. 1994. A magánhangzók [The vowels]. In: Kiefer Ferenc (ed.). Strukturális magyar nyelvtan 2: Fonológia [A structural grammar of Hungarian]. Budapest: Akadémiai Kiadó. 42–182.

Newman, Stanley. 1967. Yokuts. Lingua 17 : 182–199.

Palmer, F. R. 1967. Bilin. Lingua 17 : 200–209.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. A Comprehensive Grammar of Contemporary English. London: Longman.

Robins, R. H. 1967. Yurok. Lingua 17 : 210–229.

Sapir, Edward. 1949. Language. New York: Harcourt, Brace & Comp.

Schultz-Lorentzen. 1945. A Grammar of the West-Greenlandic Eskimo. København: C. A. Reitzels Forlag.

Thalbitzer, William. 1911. Eskimo. In: Franz Boas (ed.) A Handbook of American Indian Languages. Washington, D. C.: Smithsonian Institute. 971–1069.

van Wyk, E. B. 1967. Northern Sotho. Lingua 17 : 230–261.