DE GRUYTER

Jiří Krámský

THE WORD AS A LINGUISTIC UNIT

THE WORD AS A LINGUISTIC UNIT

JANUA LINGUARUM

STUDIA MEMORIAE
NICOLAI VAN WIJK DEDICATA

edenda curat

C. H. VAN SCHOONEVELD

INDIANA UNIVERSITY

SERIES MINOR

NR. 75

1969

MOUTON

THE HAGUE · PARIS

THE WORD
AS A
LINGUISTIC UNIT

by

JIŘÍ KRÁMSKÝ

1969

MOUTON

THE HAGUE · PARIS

TABLE OF CONTENTS

ABBREVIATIONS

ArOr — Archiv orientální
ČMF — Časopis pro moderní filologii
IJAL — International Journal of American Linguistics
TCLC — Travaux du Cercle Linguistique de Copenhague
TCLP — Travaux du Cercle Linguistique de Prague
TLP — Travaux Linguistiques de Prague

I

INTRODUCTION

One of the questions discussed at the International Linguistic
Symposium at Erfurt in 1959 concerned the problem of the analysis
of a given utterance (*Redefolge*) into its elements. The analyza-
bility of the utterance is generally taken for granted, but there is
little agreement in what to analyze and what methods of com-
parison to use. The analysis may be done at different levels and a
certain hierarchy can be established. According to Ellis this
hierarchy has the following form: an utterance or a text can be
dissected into syntactically independent units (sentences), these
again into syntactic constructions (syntagms). The syntagms
exist at different levels of complexity and the simplest syntagms
can be analyzed into words and words into morphemes (or morphs),
the units of meaning. It is mostly assumed that the analysis should
be done on the basis of a known meaning. Thus Buyssens points
to French sentences *je l'apprends* and *je la prends* which can be
differentiated only if they are understood; it is often necessary
to understand in order even to take in the preceding sentences.
Some linguists, however, do not exclude an analysis that lacks
the participation of meaning. Here we should like only to point
out the following trivial facts: when we perceive a coherent
utterance in an unknown language we seldom manage to dissect
it into smaller sections than those separated from each other by
a pause or marked off by an accent. On the other hand, when we
perceive an utterance in a language familiar to us, we can dissect
it into certain units. It is usually not morphemes but words which
are dissected most easily from the flow of speech. That the word
is a natural and indisputable unit of language can best be shown
in the written expression of the utterance. Another question is,

however, whether every group of graphemes separated from the preceding and following group of graphemes by a gap is really a word. Usually, it is.

Hansjakob Seiler[1] is of the opinion that all troubles caused by the problem of the word originate in the fact that the word is regarded as a unit. According to Seiler, the word cannot be regarded as a unit in the sense in which the morpheme is; "... the word is not a unit but a constituent of a sentence or clause". Seiler further writes: "The word is a constituent which contracts relations within the frame on the level of the sentence, this level being one of relations, not of units." We agree with the view that the word is not the unit in the sense the morpheme is. However, there is a certain overlapping between the morpheme and the word. It is well known that morphemes may be either bound forms or free forms. When they are bound forms, they are parts of words. When they are free forms, they are identical with words. It is the meaning (in a broader sense of the word) which morphemes (both bound and free) have in common with words. As a component of a sentence the word is, of course, the bearer of relations, but this does not preclude regarding it as a linguistic unit. In our opinion it is quite logical to postulate the existence of a higher unit than the morpheme (in its bound form) possessing meaning (in a broader sense) and being 'free'. In linguistics the concept of the word as a linguistic unit is quite common. However, if we are to define the word, this most common concept of linguistics appears to defy, so far most successfully, any definition. In linguistics there are a lot of such concepts which resist correct definition and, as a matter of fact, it can be said that linguistics is continually grappling with definitions. We have only to note how many definitions there are of the phoneme, syllable, sentence, etc.

It is therefore understandable that there are pessimistic views as to the possibility of defining anything in linguistics at all.

[1] Hansjakob Seiler, "On Defining the Word", *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., 1962* (The Hague, 1964), 767.

Thus Vl. Skalička maintains[2] that linguistic concepts have no generally approved definitions. He writes:[3] "We commonly use for different languages such terms as WORD, SENTENCE, CASE, TENSE, MOOD, and yet we are aware of the fact that we use them in a rather different sense. The term WORD undoubtedly means something else in Czech, German, French. We are content with an approximate identity." And further he writes, very realistically:[4] "It is the task of linguistics to look for such a solution that would make possible to give a true picture of the concepts in their change and in their various forms. We must consider the most basic features of the languages and find out what units become asserted in it."

According to Skalička, particular linguistic elements contain different components which render impossible a brief, unambiguous definition. Nevertheless, even if it is true that we often use some terms in a sense not quite established, we do not view the possibility of the definability of linguistic terms as pessimistically as Skalička does. We can assume the existence of a hierarchy of marks which permits us to assign the word a fixed place in the system of language. The problem of the word must be viewed, above all, in terms of typical words, though we shall not avoid marginal problems. It is just these marginal problems that make an unambiguous definition of the word so difficult.

There are some linguists who exclude the concept of the word from linguistics. Thus, for instance, Ch. Bally[5] rejects the concept of the word because in establishing it the standpoints of vocabulary, grammar (language system), phonology, and orthography are allegedly intermixed. Instead of the word he introduces on one hand the concept of SEMEME, i.e., the mark for a lexical concept (e.g., *lup-*), and on the other hand the concept of a syntactical molecule which includes both the sememe and the

---

[2]   Vladimír Skalička, "Komplexnost jazykových jednotek" [Complexity of Language Units], *Universitas Carolina, Philologica*, Vol. 3, No. 1, 15-25.

[3]   *Op. cit.*, 16.

[4]   *Op. cit.*, 18.

[5]   Ch. Bally, *Linguistique générale et linguistique française* (Paris, 1932), 288 ff.

grammatical mark (e.g., *lup-us*).[6] Some other linguists do not reject the concept of the word, but restrict its use. They concede, for instance, the independence of the word in Indo-European languages, but in other languages (e.g., Semitic, Bantu, American Indian) they regard it as not so clear-cut. A. Sechehay[7] and A.H. Gardiner[8] consider the word—in the sense of the opposition LANGUE—PAROLE—as the unit of LANGUE and the sentence as the unit of PAROLE.

Let us revert to the relation between the morpheme and the word. P.L. Garvin[9] regards the word as the "properly defined distributional framework necessary for a precise determination of the distribution of morphemes". Garvin writes: "If the definition of such a frame unit can be based on criteria of a different sort than the definition of any other larger morphemic unit, such a morphemic unit can properly be called a word. Conversely, if no such definition of a larger morphemic unit is possible, it can be said that the language in question has no words in our sense." Garvin further defines the word in Kutenai, Ponape, and Turkish as a sequence of morphemes and morpheme clusters in a fixed order. The application of this definition to French and English has, however, given negative results "due to the absence of consistently ordered long chains of morphemes in these languages". Consequently Garvin concludes that there are many languages that defy the application of the criterion of the fixed order of morphemes. Since in languages such as French, Czech, and Polish the word boundary is formed by the phonological feature of accent, Garvin has tried to define the word by means of this criterion. However, again without success, as there have remained a series of multi-word contours containing proclitics and enclitics not bounded by stress and yet traditionally considered separate words.

[6] Cf. Karel Svoboda, "Pojem slova" [The Concept of the Word], *Listy filologické*, 67 (1940), 173-178.

[7] A. Sechehaye, *Essai sur la structure logique de la phrase* (Paris, 1926), 14.

[8] A.H. Gardiner, *The Theory of Speech and Language* (1932); cf. V. Mathesius, *Slovo a slovesnost*, 1 (1935), 42 ff.

[9] Paul L. Garvin, *On Linguistic Method* (The Hague, 1964), 32.

Equally unsuccessful have been the attempts to define the word in French, English, and Vietnamese by criteria other than phonology or fixed order or by a combination of the two. Garvin concludes:[10] "I have come to the conclusion that a proper definition of the word, that is, a consistent one without unaccountable residue, will emerge only from a complete distributional reanalysis of the morphemic structure, using a properly defined temporary framework such as the utterance."

From these words we may draw a conclusion for our work too, viz., that it is necessary to look, not for such criteria as are applicable to one or a few languages, but for criteria as general as possible, criteria that would embrace the greatest possible number of languages. The resulting definition of the word should, in our opinion, be as broad as possible. On the other hand we are fully aware of the fact that such a definition as would be applicable to all languages is hardly possible. In different languages there are sound complexes that are words from some angles but not from others. The tendency to absolutize the word is asserting itself and it would be wrong to try to suppress that tendency. And, as the criteria of the word so far known have been applicable predominantly to one or a few languages, we shall make it the aim of our study to verify thoroughly the existing criteria of the word, and, possibly, to look for new criteria capable of the most general application, so that the resulting definition of the word will be capable of embracing the maximum number of languages.

---

[10] *Op. cit.*, 33.

THE PLACE OF THE WORD IN THE LANGUAGE SYSTEM

An important task which we assume must precede the proper analysis of word criteria is to discover the place occupied by the word in the system of language. The solution to this problem will provide us with a solid starting point for our analysis of the criteria of the word.

The place of the word in the system of language is a problem of the relation of the word to the units of different language plans. This relation has been very pregnantly expressed by B. Trnka in his paper "Peut-on poser une définition universellement valable des domaines respectifs de la syntaxe?"[1] Trnka discusses the relation between the plan of the word and the plan of the sentence. The sentence and the word are, according to Trnka, on different levels, as the sentence is neither the result of a mechanical arrangement of words nor equivalent merely to the sum of the words of which it consists. The plan of the sentence and the plan of the word are not in a reciprocal relation but in the relation of the function to its realization, that is, in a relation similar to that of the word and phonemes realizing it. Elsewhere[2] Trnka formulates this problem as follows:

Though both plans, the phonological and the morphological one, cooperate with the syntactic plan toward a common aim of language utterance that is toward the intersubjective understanding, each of them works with quite different means. The basic elements of the morpho-

---

[1] *Actes du sixième congrès international des linguistes, Paris, 19 au 24 Juillet 1948* (Paris, 1948), 19-30.
[2] B. Trnka, *Rozbor nynější spisovné angličtiny*, II: *Morfologie slovních druhů (částí řeči) a tvoření slov* [Analysis of Present-Day Literary English, II] (Prague, 1954, mimeographed), 5.

logical plan are, that is to say, not phonemes but words, units endowed with meaning, whose relation to phonemes or to sequences of phonemes (for instance [teibl, taim, buk]) respectively is that of the sign to its realization. The word cannot be realized as a mere grouping of pho-nemes but as a unit realized by phonemes capable of being displaced and capable of semantic opposition. In relation to the sentence as an independent sign of extralinguistic reality to which the speaker takes a stand the word is only the realization of a higher unit — the sentence. Consequently, the word has in the entire language system a double aspect: it is partly (partial) sign, partly realizer. Thus morphology has a central position in language: on one hand it deals with words as partial signs realized by phonemes, on the other hand with words as realizers of the sentence.

However, we have a few critical remarks on this formulation. Trnka regards words as basic elements of the morphological plan, but he does not mention morphemes, the relation between mor-phemes and words, and the place of both in the morphological plan. We are of the opinion that it is the morpheme as the small-est semantic unit (bound to the word) which must be regarded as the basic element of the morphological plan. A morpheme can at the same time be a word, and a word can be uni- or polymorphemic. The difference between the position of the morpheme and that of the word in the system of language is that the morpheme belongs solely and exclusively to the morphological plan whereas the word belongs equally both to the morphological and to the syntactical and lexical plan.[3] Of course, when we regard as a morphological unit the word as a realizer of the sentence, the boundaries between morphology and syntax must needs be very vague and disputable, if boundaries and, thus, the difference between morphology and syntax can in this case be spoken of at all. The criterion of the displaceability of the word in the sentence will be discussed later.

Having adopted the morpheme as a unit of the morphological plan, we must determine the function the word has in the morpho-

[3]   Cf. Ivan Poldauf, "Tvoření slov" [On the Formation of Words], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 146: "Words formed by morphemes are units of both the lexical and the morphological plan, even if they represent units of a higher order than are morphemes."

logical plan. It is necessary to make a sharp distinction between the function of the word in the morphological plan and the function of the word in the syntactical plan. It is the morphological relations within the word, that is to say, the relations between particular morphemes (as far as multi-morphemic words are concerned), which are of concern, whereas in the syntactical plan the relations are those *between* words as realizers of the sentence.[4] Words as syntactical units are in mutual syntactical relation, whereas words in the morphological plan can be taken into consideration solely in terms of their relation to the paradigmatic axis.[5] Only this double aspect of the word can throw light on the question of whether forms of one and the same word represent different words or only variants of one word. This problem is pointed to by B. Trnka,[6] when he contrasts the word as an element of the syntagmatic axis and the word as an element of the paradigmatic axis. From the standpoint of the syntagmatic axis each of the word forms is a special word which is differentiated through its opposition to other words of the sentence, whereas from the standpoint of the paradigmatic axis all word forms are variants of the same word.[7] This differentiation by the two axes is not without importance, as, for instance, in Chinese and in Bantu languages

[4] Likewise, the position of the word in the system of language and, at the same time, the difference between morphology and syntax is characterized by P.L. Garvin, who writes (*op. cit.*, 65) that morphology is a study of the inner structure of words and syntax a study of the outer functioning of words. The division into morphology and syntax presupposes, according to Garvin, that words are thoroughly defined as linguistic units.
[5] This, of course, does not mean that we put an impenetrable wall between morphology and syntax. In fact, syntax presupposes morphology and morphological relations are realized in syntax. If we, however, wish to delimit morphology somehow, we must, in the first stage of our investigation, center our attention on the relations within morphology, within the framework of the morphological plan, only in the second stage investigating the relations between morphology and syntax.
[6] Cf. the paper cited in fn. 1 on p. 12, 28-29.
[7] Edward Sapir (*Language*, New York, 1921, 26) speaks of radical and grammatical elements. In the same work (p. 25) he writes: "The single word may or may not be the simplest significant element we have to deal with. The English words *sing, sings, singing, singer* each conveys a perfectly definite

the word form can be given outwardly by tone characteristics of another word of the sentence. The dependence of the word form upon another word of the sentence can be met with even in Celtic and other languages. Thus, for instance, in Manx[8] the words beginning with *c* have variants with initial *c, ch, g*, according to the influence of preceding words, such as, e.g., *carrey* 'friend' — *e charrey* 'his friend' — *nyn garrey* 'our friend', the words beginning with *b* have variants with initial *b, v, m*, for instance, *bea veayn* 'long life' — *e vea* 'his life' — *nyn mea* 'our, your, their, life', etc. The well-known 'sandhi' phenomena also belong in this category.

It is beyond dispute that the word is, above all, a unit of the lexical plan. Whereas in the lexical plan the word is the basic and the sole unit, in the morphological plan the basic unit is the morpheme and in the syntactical plan it is the sentence. The problem is whether the word as a part of the word stock, as the bearer of lexical meaning, possesses grammatical form. According to A. Isačenko[9] the word does not become 'form' until it is used in a syntactical whole.[10]   However, this problem is of minor impor-

---

and intelligible idea, though the idea is disconnected and is therefore functionally of no practical value.   We recognize immediately that these words are of two sorts.   The first word, *sing*, is an indivisible phonetic entity conveying the notion of a certain specific activity.   The other words all involve the same fundamental notion but, owing to the addition of other phonetic elements, this notion is given a particular twist that modifies or more closely defines it. They represent, in a sense, compounded concepts that have flowered from the fundamental one.   We may, therefore, analyze the words *sings, singing*, and *singer* as binary expressions involving a fundamental concept, a concept of subject matter *(sing)*, and a further concept of more abstract order — one of person, number, time, condition, function, or of several of these combined."
If Sapir regards the expressions *sings, singing, singer* as binary expressions which have a common basic concept, it is the same as regarding them as variants of the word *sing*.
[8]   Cf. John Kelly, *A Practical Grammar of the Ancient Gaelic, or Language of the Isle of Man, Usually Called Manx* (Douglas, Isle of Man, 1870).
[9]   A. Isačenko, "Slovo a věta" [The Word and the Sentence], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 89.
[10]   On the other hand, K. Hausenblas, E. Pauliny, and B. Havránek expressed disagreement with this assertion (cf. *op. cit.* in fn. 9, 117).

tance for this chapter. More relevant is another assertion of Isačenko, that is to say, that in the lexical system there exist only roots and not words. According to B. Havránek the units of the lexical plan are words, not roots of words; words form an infinite succession which can be prolonged.

If we evaluate the place of the word in particular language plans, we come to an important difference in the place of the word and of other language units in the language system. This difference lies in the fact that the phoneme, morpheme, and sentence have their fixed place in the language system, whereas the word belongs both to the morphological and to the syntactical and lexical plans. The word is a bridge between morphology and syntax, making the transition from morphology to syntax gradual and imperceptible. Extreme cases are those of the identification of the word with the morpheme on one hand and with the sentence on the other hand.

The place of these basic units in the language system can be represented in the following way:
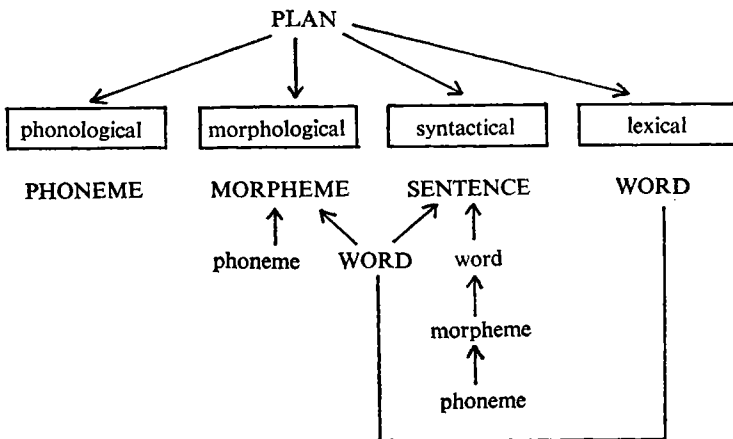
III

CRITERIA OF THE WORD

As we have already mentioned, there are linguists who are pessi-
mists about the possibility of defining the word. They either do
not admit of any definition, or they admit a definition for each
particular language but no universal definition valid for all langu-
ages. According to J. Lohmann[1] the question of a definition of
the word is similar to the old and notorious question of 'defining'
the sentence. To require a definition in the ordinary sense (that is,
the delimitation of the concept according to outer signs) is, accord-
ing to Lohmann, in both cases nonsensical, and in the case of the
word even more nonsensical than in the case of the sentence. This
is because everybody knows, at least in practice, what the sentence
is, even if he is not able to give its proper characterization. How-
ever, this cannot be said of the word, unless we acknowledge as
decisive that division of words which happens to have been
established in the respective language (cf., e.g., the English *post
office*, the German *Postamt*, the French *l'an*, and the Rumanian
anul). C.E. Bazell[2] gives six criteria for the determination of a
sound complex as a word: (1) insertion, (2) substitution, (3) se-
quence, (4) independence, (5) phonemic structure, (6) non-phone-
mic structure; but these criteria are not, according to Bazell,
applicable to all languages to the same extent. As examples he
quotes Czech and Hungarian. Both these languages include the
initial accent in their phonological system of the word, but Czech

---

[1] Cf. *Archiv für vergleichende Phonetik*, 1943, 130.
[2] C.E. Bazell, "Historical Sources of Structural Units", *Misceldnea homenaje
a André Martinet: Estructuralismo e historia*, I (Editada por Diego Catalán,
Tenerife, Canarias, Universidad de la Laguna, 1957), 19-29. Cf. also a review
by Robert A. Fowkes, *Language*, Vol. 35, 78-90.

allows few possibilities for inserting morphemes within the word, whereas Hungarian, on the other hand, allows relatively many. Bazell's conception seems to be acceptable. The difficulties in the application of any definition of the word to all languages seem to confirm the assumption that a universal definition of the word evidently needs a complementary qualification according to the typological appearance of particular languages. It is understandable that differences in the structure of particular languages bring about differences in the structure of the language units too, especially of higher units such as the word and the sentence. Similarly, J. Vachek[3] emphasizes the importance of the typological shape of a language for "proper, special criteria of the independence of the word" which are added to the universally valid criteria of the independence of the word in the sense of Bazell's thesis. Vachek supports this thesis by pointing to the opposition of the word and the collocation, which in Czech is more explicit and clear-cut.

According to W.L. Graff[4] the difficulties in defining the word have their source in the following basic errors: (1) inadequate importance is ascribed to a phonetic or semantic feature at the expense of a complex, semantic-phonetic combination; (2) the relation of the word to the sentence and *vice versa* is wrongly appreciated; (3) the character of the word is often identified with its quantitative extension, or, at least, the character and the quantity of the word are not strictly separated; (4) the facts relating to the evolution of language are not strictly separated from those relating to the static system. As further analysis shows, Graff has given an essentially true picture of the basic shortcomings of the definitions of the word given so far. However, in order to be able to give a really universal definition of the word, we must first of all check those criteria for the independence of the word which are currently used in the identification of the word.

[3] In a lecture on the position of the word in the system of Czech and English (1960).
[4] W.L. Graff, "The Word and the Sentence", *Language*, 5 (1929), 163.

1. THE SEMANTIC CRITERION

According to this criterion the word is determined semantically by
its meaning.   Unfortunately, the criterion of meaning immediately
gives rise to considerable difficulties in determining the morpheme.
There are, therefore, many linguists who reject the semantic
criterion.   According to A. Martinet[5], purely semantic[6] or purely
phonological criteria are fallacious.   To support the assertion that
meaning cannot be a universal criterion of the word it is often
pleaded that, if the semantic criterion is used, the so-called 'syn-
semantics', that is, words which by themselves cannot realize the
sentence (cf. the Czech prepositions *v*, *s*, *k*, the English *my*, *your*,
the, *a(n)*, etc.) or can do so in very rare cases only[7], would have

[5]   Cf. the paper quoted in fn. 1 on p. 12 .

[6]   A purely semantic criterion of this type is, for instance, Sapir's following
definition (*Language*, 35): "... the word is one of the smallest, completely
satisfying bits of isolated 'meaning' into which the sentence resolves itself",
though in the preceding account he also pays attention to the word form.
Thus, on page 35 he writes: "The word is merely a form, a definitely molded
entity that takes in as much or as little of the conceptual material of the whole
thought as the genius of the language cares to allow.   Thus it is that while
single radical elements and grammatical elements, the carriers of isolated
concepts, are comparable as we pass from language to language, the finished
words are not."

[7]   The distinction between autosemantics and synsemantics, which comes
from A. Marty, is similar to that made by E. Benveniste ("Les niveaux de
l'analyse linguistique", *Proceedings of the 9th International Congress of Lin-
guists*, The Hague, 1964, 266-293) between autonomous words functioning as
components of sentences and synonymous words *(mots synnomes)* entering
into the sentence only in connection with other words; for instance, the French
le (la ...), *ce (cette...)*, *mon (ton...)*, or *de*, *à*, *dans*, *chez*, but not all prepo-
sitions: compare, for instance, *je travaille avec* or *je pars sans*.   This distinction
between autonomous and synonymous words does not, however, fully agree
with Marty's distinction.   Under synsemantics may, for instance, be classed
the auxiliary verbs which Benveniste classes with autonomous words just
because they are verbs and enter directly into the constitution of the sentence.
      We should further mention that K. Hausenblas in his paper "Über die
Bedeutung sprachlicher Einheiten und Texte", (*TLP*, 2 (1966), 61), adds to
Marty's two word groups a third one which he names 'quasisemantics', which
are components of the so-called (semantically) motivated expressions which
possess a motive feature such as, for instance, the first component of the

to be excluded. Ernst Otto[8] ascribes to the word a double meaning: *Begriffsbedeutung* and *Beziehungsbedeutung*. Even prepositions and conjunctions possess *Begriffsbedeutung* as well as *Beziehungsbedeutung*. On the other hand, A Sechehaye[9] asserts that prepositions are not genuine words but proclitic particles which have no other meaning than to express relations. Ernst Otto regards this opinion as unrealistic. Equally incorrect, according to Otto, is the statement of O. Naes[10] that the conjunction *und* does not point to any objective condition or state. However, Otto admits (cf. p. 10) that a particle can lack *Begriffsbedeutung* (for instance, *de* in *de la table*, *ist* in *ist grün*, *der* in *der Tisch*, etc.).

With regard to synsemantics Anton Marty[11] gives the following definition of the word:

> Unter "Wort" aber verstehe ich jedes Sprachmittel, das als besonderes Glied des Organismus der Rede empfunden und als besondere semantische Einheit behandelt wird. Ich sage nicht: eine besondere semantische Einheit *ist*. Denn wie schon früher bemerkt wurde, kann man dies selbst von den logisch begründeten Synsemantika nur in einem weiteren Sinne, im strengen Verstande aber bloss von den Autosemantika sagen.

Marty does not give up the semantic criterion but this cautious formulation leaves the door open to synsemantics too. However, Marty's definition lacks a more clear-cut stress of the independence of the word in contrast to the morpheme.

B. Trnka[12] criticizes Bloomfield's definition according to which the word is the smallest expression of meaning ('utterance' or 'act

---

compound *Jungfernapfel*: this component has not only a distinctive function but also a quasisemantic function which under certain circumstances may fully assert itself.

[8] Ernst Otto, *Sprache und Sprachbedeutung, Abhandlungen der deutschen Akademie der Wissenschaften in Prag, Philosophisch-historische Klasse*, 7 (Prague, 1943), 8.

[9] A. Sechehaye, *Essai sur la structure logique de la phrase* (Paris, 1926), 83.

[10] O. Naes, "Versuch einer allgemeinen Syntax der Aussagen", *Norsk Tidsskrift for Sprogvidenskap*, XI (Oslo, 1939), 157.

[11] A. Marty, *Satz und Wort* (Reichenberg, 1925), 40.

[12] Cf. *op. cit.* in fn. 2 on p. 12, 6.

of speech') simply because this definition excludes synsemantics from the class of words. Against this, Trnka defines the word as "a unit realized by phonemes, capable of being displaced and capable of semantic opposition." Thus this definition implies two criteria, that of displaceability and that of semantic opposition.

J. Vachek[13] uses as his starting point the purely formal definition of V. Mathesius[14] according to which "the word is the smallest section of sound flow which in no way is bound to other such sections", but he adds the semantic criterion. Thus, according to Vachek, "the word is a section of utterance which is in relation to some component of extra-linguistic reality and which can, as an indivisible whole, more or less change its position toward other such sections of utterance". As at present a close tie between the sound form and the semantic contents is generally considered the very essence of language, this connection being also emphasized by Vachek[15], we regard the semantic criterion as part of a universal definition of the word. However, we must be aware that meaning is added to the word from outside, as it is possible to have first a sound complex without meaning, the meaning being added to it later. On the other hand, in a linguistic description we have to do with units that already have some meaning; our analysis cannot, consequently, exclude the meaning. The word has a communicative function, which as its basic function cannot be neglected.

---

[13]   In the unpublished lecture "On the Position of the Word in the System of Czech and English", Prague, 1960.

[14]   V. Mathesius, "O potenciálnosti jevů jazykových", *Věstník Král. čes. spol. nauk, tř. hist.*, II (1911), 1-24 (cf. the English translation of this paper: "On the Potentiality of the Phenomena of Language", *A Prague School Reader in Linguistics*, Indiana University, Bloomington, 1964, 1-32).

[15]   In his work *Jazykovědný pohled na sovětské zkoušky srozumitelnosti telefonního přenosu* [Linguistic View of the Soviet Tests of Intelligibility of the Telephone Transmission] (mimeographed, Prague, 1957), 22, J. Vachek illustrates the linking of two factors, the acoustic form and the semantic content, with a very instructive example: the sound of the English word *err* may be regarded as a linguistic fact only when it refers to the concept 'make mistakes, be incorrect, sin'; the entirely identical sound we often hear from people lazily stretching their limbs lacks such a semantic correlate and therefore cannot be denoted as a language fact.

Even when, theoretically, meaning is not inherent in the word, the form being primary, it is not possible to separate the form and the meaning in connection with words.

As to synsemantics, we cannot accept the assertion that they are words lacking any meaning. They possess what we call 'grammatical meaning' or, according to Otto, *Beziehungsbedeutung*. And, as this 'meaning' is essentially different from the meaning of the other words, 'autosemantics', it is necessary to include this fact in some way in the universal definition of the word. Another important problem, the relation between meaning and form, will be discussed in a separate chapter.

2. THE CRITERIA OF SEPARABILITY, REPLACEABILITY, AND DISPLACEABILITY

These criteria are so closely connected that it is better to treat them simultaneously. Of course, this does not mean that one cannot exist without the other. Thus, Vladimír Hořejší[16] mentions the criterion of the separability of a word from neighboring words by other words and the replaceability of a word by other words. He does not mention the semantic criterion. Hořejší is quite right in not regarding displaceability[17] as a reliable criterion, as displaceability is limited by the combinatorial capacity of the words. Thus, for instance, in the connection of a preposition with a noun (cf. in Czech *v domě* 'in the house') the preposition *v* can be replaced by another preposition (*na domě* 'on the house', *o domě* 'about the house'), separated from the noun by another word (*v mém domě* 'in *my* house'), but it cannot be displaced (sequences like *domě na* or *domě o*, etc., are not possible).[18] Further examples

[16] Vl. Hořejší, "Postavení morfologie v mluvnici a její obsah" [The Position of Morphology in Grammar and Its Content], *Časopis pro moderní filologii*, 39 (1957), 75 ff.
[17] V. Mathesius had already mentioned the criterion of displaceability in his article "O potenciálnosti jevů jazykových" (cf. fn. 14 on p. 21: words in a sentence are more or less displaceable but syllables in words are not).
[18] Cf. also J. Vachek, "České předložky a struktura češtiny" [Czech Prepositions and the Structure of Czech], *Naše řeč*, XIX, Nos 6-10, reprint, p. 5.

can be given from other languages.   Thus, in French the criterion
of displaceability is not valid in the sequence *les braves gens* 'good
people', because when we displace the words to form *les gens braves*
we obtain another meaning: 'brave people'; similarly, *une certaine
chose* 'a certain thing' as against *une chose certaine* 'certain thing,
a safe thing', etc.   The displacement of words without any change
or distortion of the meaning of the given word or word group is
not allowed by languages in which some of the morphological
functions of the words are expressed by the word order and not
by the ending as for instance, in English the position of the nomi-
native and accusative and of the dative and accusative of nouns,
the position of the attributive adjective, the position of the depend-
ent possessive pronoun).   The flexional character of the language
and the freedom of its word order make a greater degree of
displaceability possible (as for instance, in the majority of Slavic
languages).

The criterion of separability is formulated by J. Vachek in the
following way:[19] "In so far as a certain sound section has the
validity of a coherent word whole, it is not possible for any part

---

The indisplaceability of the preposition would be evidence in favor of the
preposition being a part of the word to which it belongs.   Another reason
why prepositions in Czech are regarded as parts of words is that the preposition
attracts the word accent.   However, Professor Vachek convincingly shows on
the basis of the example /g řeči/ 'to the speech' in contrast to /křeče/ 'spasms,
cramps' that, as far as assimilation is concerned, the prepositional expression
behaves as a group of two words.   That is to say, the voiced or unvoiced
character of realizations in sound of the Czech phoneme /ř/ is determined by
their position in the sentence and by the sound environment, which means that
the voiced variant of the Phoneme /ř/ occurs at the beginning of a word and
between voiced phonemes, whereas its unvoiced variant occurs at the absolute
end of the word and in the neighborhood of an unvoiced phoneme.   According
to Vachek, only the segmentation /g/řeči/ is possible, as the voiced /ř/ stands
either at the beginning or in the middle of a word following a voiced conso-
nantal phoneme.   In the latter case we should have to regard this sound
segment as a single word, which is not possible because it would presuppose
the existence of the voiced phoneme /g/ in Czech.   However, it is well known
that in Czech /g/ is only a voiced variant of the phoneme /k/ the voicing of
which must be interpreted as an assimilation with the following voiced pho-
neme.
[19]   Cf. *op. cit.* in fn. 18 on p. 22, 8.

of it to separate from its other parts.''   Vladimír Hořejší is aware
of the difficulties created by the criterion of the separability of the
word from neighboring words by other words, or even by the
criterion of the replacement of the word by other words.   He
shows that there is a disagreement between these criteria and the
written form of language and draws attention to frequent variations
in the spelling of various languages, as for instance the writing of
prepositional expressions as one word and adverbial expressions
as separate words.   Thus, he quotes cases where one word is
expressed in writing by a group of words, for instance, in Czech
na jaře 'in spring', *na podzim* 'in autumn', which, according to this
criteria, should be written as one word, because, allegedly, it is not
possible in this case to separate the preposition *na* from the noun
by another word.   However, in our opinion, if in addition to
do jara 'until spring' we can say, e.g., *do příštiho jara* 'until next
spring', which means that we are separating the noun from the
preposition by inserting an adjective, and if it is possible in other
connections also to separate the preposition *na* from the noun,
as in *na stole* 'on the table' — *na mém stole* 'on my table', then
by analogy it is possible to view the sequences *na jaře, na podzim*
also as connections of two words.   J. Vachek points to a pheno-
menon which even today is regarded as an offence against the
rules of Czech accentuation, that is to say, where the accent is
wrongly laid on the noun and not on the unisyllabic preposition.
According to Roman Jakobson,[20] in the shift of the accent is
manifested "the germinal stage of the transformation of the fixed
accent into the movable one".   Another case is where a group
of words is written as one, such as, according to Hořejší, the
Czech adverbial expressions *sbohem* 'good-bye', *zčásti* 'partly',
dohromady 'together', as well and German and Hungarian verbs
with separable prefixes.   In our opinion, even in this case the
criterion of separability is not satisfactory.   It is true that in Czech
we can say *z velké části* 'for a great part', which may seem to be
evidence in favor of the independence of the two parts, *z* and *části*,

---

[20]   Roman Jakobson, *Základy českého verše* [Fundamentals of Czech Verse]
(Prague, 1926).

but we believe that *z části* cannot be identified with *zčásti*, simply because *z* in the latter example cannot be regarded synchronically as a preposition in cases where the complex is understood as an adverb.   In other words, *zčásti* does not equal *z části*, as *zčásti* is a word no section of which is separable by other sections, whereas *z části* is analysable into the preposition *z* plus the genitive of the noun *část*.   A similar evaluation can be given to the words *zjara* 'early in spring', *zvečera* 'early in the evening', etc.   In the case of German verbs with separable prefixes it is, in our opinion, necessary to assume the existence of distant morphemes *(morphème brisé)*.

In English and in Swedish the genitive *s* is, according to Hořejší, an independent word, because it can be separated from its basis by another word, for instance, *the king's*, as compared with *the king Arthur's*. This conception is also upheld by American linguists, evidently in consequence of their defining words as "minimum free forms", as pointed out by Togeby.   According to this definition, the French *de*, *à*, *en* would not be independent words.   Let us add that, even if the ending of the group genitive in English complies with the criterion of separability, it does not comply with the criterion of isolatedness to be discussed in the next chapter.   The consequential application of these criteria even leads to absurdities like regarding the adverbial ending *-mente* in Spanish and, similarly, the endings of the future and conditional in Portuguese, as independent words because between them and the verbal basis pronouns can be inserted; for instance, in Portuguese instead of *vos darei*, 'I shall give you' one may say *dar-vos-ei*, so that in this case *-ei* would be an independent word.   However, both cases are at variance with the criterion of isolatedness.

A. Martinet[21] regards as most reliable the criterion of inseparability but adds, in order to make it more precise, that it is possible, without impairing the unity of the word, to insert between the two signs *(signes)* of which a word consists another sign which exists solely in that type of combination (e.g. *dormont-dormiont*).

---

[21]   Cf. *op. cit.* in fn. 1 on p. 12, 27.

Compound words must, according to this criterion, be regarded in a series of cases as simple words, not as combinations of words. Let us complement Hořejší's criterion of separability by adding a condition according to which the separating element must be an independent word into which it itself could again be separated.[22]

The criterion of separability is also contained in A. Zaretski's definition[23] according to which

A word is a part of speech which can be separated from the surrounding parts and cannot itself be divided. If it can be divided, it is more than a word; if it cannot be separated, it is less than a word. "Dividing" means transposing the parts or interposing other words between them.

This definition is thoroughly analyzed by Uriel Weinreich,[24] who argues that Zaretski's definition is practically useless and theoretically unsuitable for the structure of Yiddish. Thus, for instance, he argues that in the sentence *ikh hob ge/zen* 'I saw' we can insert between *ge* and *zen* the sequence *hert ober nit ge...*, so that we get *ikh hob ge/hert ober nit ge/zen* 'I heard but did not see'. This line of argument does not seem convincing. As *ge-* is a prefix which occurs in this form regularly with other verbs too, it is evident that we can add to it, e.g., the morpheme *-hert* and other words or morphemes, the last of which, again the prefix *ge-*, would be added to another morpheme, in our case to *-zen*. In our opinion we should not go to such absurd lengths in the application of the criterion of separability. We could do that in any language with a sound sequence; for example, in English, if we have the sequence *I have s/een*, we can insert between *s-* and *-een* the sequence *-ought and s-*, so that we obtain *I have s/ought and s/een*. And this type of insertion can be further extended to theoretically infinite lengths. But what is the sense of this playing with word sequences? The

[22] Cf. *op. cit.* in fn. 16 on p. 22, 76. Hořejší quotes as examples the words *nenésti* 'not to carry' and *nepřinésti* 'not to bring'. If we compare them, the complex *nenésti* seems to represent two words. This is not the case, however, as the complex *při* is not an independent word.
[23] Ayzik Zaretski, *Yidishe gramatik* (Vilna, 1929), 13.
[24] Uriel Weinreich, "Stress and Word Structure in Yiddish", *The Field of Yiddish*, I (New York, 1954).

incorrectness of such a procedure is hold up to the light when we remember the earlier mentioned formulation of Vachek to the effect that, insofar as a certain sound section has the validity of a single word when it is a whole, it is not possible for any of its parts to separate from its other parts. The other cases mentioned by Weinreich, that is, *shlof/n, shlof-inke-n, shlof-de-n,* are in their form analogous with the Czech forms *spát* 'sleep', *spávat* 'to be in the habit of sleeping'. Here we are concerned with the insertion of morphemes which modify the meaning of the word and, consequently, create new words. A more detailed discussion will be devoted to this problem in the chapter dealing with word cohesion.

Another theory of the word discussed by Weinreich is the theory of E. Falkovitsh.[25] According to this theory, word is a grammatical-semantic unit. Falkovitsh's criteria are mixed. On the semantic level he is generally right in distinguishing between affixes (which have 'a formal meaning') and roots (which have 'a material meaning'). The word is the root or the root with its affixes. However, Falkovitsh also gives formal tests of separability and displaceability (p. 27): "That which can be separated or transposed in the phrase is not a prefix or a suffix and it is to be written separately." When he comes to deal with the problem of a subject pronoun which is inseparable from the preceding word, Falkovitsh determines the word-dividing line by two means: (a) by the possibility of transposition, (b) by the possibility of commutation; *bet er* 'so he asks' contains the word-dividing line, because we also have *bet zi* 'so she asks'. The possibility of transposition is, however, excluded, as the position of both words is bound by the word order, and so is commutation, because it can concern even affixes (e.g. *di shen-er-e meydlekh* 'the prettier girls' — *di shen-st-e meydlekh* 'the prettiest girls'). Falkovitsh's theory of the word contains many acceptable elements. Its merit is above all in the fact that it regards the word as a grammatical-semantic unit. We do not share Weinreich's opinion that in Falkovitsh's theory roots and affixes are distinguished on unreliable semantic grounds.

[25] E. Falkovitsh, *Gramatik far dervaksene,* I (Moscow, 1930), esp. 16-57.

It may be admitted that Falkovitsh's differentiation of formal and material meaning is a rough one,[26] as there are no sharp boundaries between formal and material meaning, yet it is so far the best WORKING differentiation available for solving the difficult problem of linking together word form and word meaning.

3. THE CRITERION OF ISOLATEDNESS

This criterion can be found in Bloomfield, who defines 'minimal free forms' as "smallest items which are spoken by themselves, in isolation".[27]   The same view is shared by E. Polivanov:[28]

Bei der Lösung der Frage, was als Wort aufzufassen ist, halte ich es für möglich, das allgemeine Kriterium der "Isolierungsfähigkeit des Wortes" anzuwenden: das Wort is ein potenzielles Satzminimum, d.h. ein solcher Redeabschnitt, den man isolieren und als einzigen Satzbestandteil aussprechen kann (z.B. im Gespräch bei Teilwiederholungen von Gesagtem, Fragen und Antworten).

Consequently, Polivanov regards the sequence *je te le dis, je te l'ai dit, je ne dis pas, je ne te le dis pas* as particular words composed of elements which cannot exist in isolation.   The weak point of this definition has been disclosed by Knud Togeby[29] through another quotation from Bloomfield:[30] "Are English forms like

---

[26]   Josef Filipec (cf. his monograph *Česká synonyma z hlediska stylistiky a lexikologie* [Czech Synonyms from the Standpoint of Stylistics and Lexicology], with a German summary, Prague, 1961, 183) points to the fact that the meaning of some words, especially of conjunctions and prepositions, also reflects *relations* to a certain reality.   O. S. Akhmanova (*Očerki po obščej i russkoj leksikologii*, Moscow, 1957, 67ff.) shows that the meaning of these words is given as a lexical, not as a grammatical relation.   Filipec defines the meaning of the words as follows: "It is a linguistically specified and formed reflection of a reality or of its relations with our consciousness (possibly obtained through the medium of our consciousness) which is in a given language relation with the word form, with meanings of other word forms and with extralinguistic reality."

[27]   Cf. Harris's concept of 'minimum utterance' (*IJAL*, 1947, 54).

[28]   E. Polivanov, "Zur Frage der Betonungsfunktionen", *TCLP*, 6 (1936), 79.

[29]   Knud Togeby, "Qu'est-ce qu'un mot?", *TCLC*, V (1949), 105.

[30]   Cf. his *Language* (1933), 179.

the, a, is, and ever spoken alone? One can imagine a dialogue: *Is? — No, was.*" Let us add a nice example in Russian quoted by A. I. Smirnickij:[31] *S sacharom ili bez? — Bez!* ('With or without sugar? — Without!').

We can fully agree with Togeby when he writes that there are no limits to what we can imagine. Therefore, Bloch and Trager[32] propose the following definition of the word: "... any fraction that can be spoken alone with meaning in normal speech." If by "meaning" the authors also understand grammatical meaning or grammatical relation, then we could even include synsemantics in the definition of the word.

4. THE PHONETIC CRITERION — THE PROBLEM OF THE ACOUSTIC IDENTITY OF THE WORD

The phonetic criterion is represented by the problem of the acoustic indentity of the word. It is one of the most crucial problems of linguistics. The fact that it is not one of the most discussed problems in linguistic literature may be due to its complexity and to the difficulty of tackling it in a simpler way. For we usually have to simplify complicated situations in order to be able to analyze them. In fact, we often cannot do without such simplifications; they help us to determine structural relations between various language phenomena which we are unable to extricate by any other means. However, in order not to distort reality we must not lose sight of the fact that we have introduced into our study of language a simplification of the linguistic concepts. The danger of forgetting that we use simplified concepts is more or less latent in nearly all linguistic concepts, even in such as are most useful for linguistic research, e.g., in the concept of the phoneme, as will be shown later.

[31] A.I. Smirnickij, "K voprosu o slove" [Problema 'otdel'nosti slova'], *Voprosy teorii i istoriji jazyka v svete trudov J.V. Stalina po jazykoznaniju* (Moscow, 1952), 196.
[32] *Outline of Linguistic Analysis* (1942), 54.

Another reason why the problem of the acoustic identity of the word has not so far received due attention from linguists is to be found in the historical development of linguistics. It goes without saying that this problem could not have aroused the interest of historical comparative linguistics of the 18th and 19th centuries. It bordered, however, on the province of the linguists who followed the Saussurian distinction of "langue-parole". Structural linguistics following any trend analyzes living speech, especially its acoustic aspect, trying to bring a functional order into a mass of language phenomena and thus elucidating the structure of the language. In order to do this it was necessary to replace the concept of the 'sound' with its innumerable nuances by a simplified linguistic unit which would be the bearer of a certain function common to a certain number of individual sounds. The creation of an 'ideal' sound, the phoneme, somewhat diverted the attention of linguists from real, individual sounds. This is one of the reasons why our problem has until quite recently received little attention, even from structuralists. If it is of more importance today, it is because of the theory of information, which deals with the amount of information of a speech event which is transferred from the speaker to the hearer. Any treatment of the problem of linguistic communication must necessarily include the problem of the acoustic identity of the word.

Let us begin our discussion with the simple statement that words are composed of phonemes which are in opposition to each other. These oppositions differ in terms of their relevance; some have a greater, some a smaller degree of relevance. The quality of these oppositions depends on the arrangement of the sounds in the word and on the quality of these sounds. It is important to know to what extent words resemble each other in the arrangement of the sounds they are composed of and whether they possess more or less distinctive features by which they differ from each other. If two words differ, for instance, only in terms of the voiced- or voicelessness of the final consonant, they are, when out of context, not identifiable for Czechs,[33] whereas for the English this is a very

[33] A characteristic feature of the Czech language is the neutralization of the

relevant feature.   The distinctive feature of vowel quantity would be irrelevant if an opposition did not exist between, for instance, the English words *leave* and *live* or the Czech words *kát* /ka:t/ *(se)* 'to repent' and *kat* /kat/ 'executioner'.

Thus, we distinguish between SYSTEMIC RELEVANCE (which implies that a certain opposition exists in the system of the language) and LEXICAL RELEVANCE (which depends on the utilization or exploitation of phonemes).   Last but not least, the distinctive power of the phonemes themselves is of importance.   In this respect, languages possessing a great number of phonemes are worse off than languages possessing a smaller number of phonemes.

The main criterion for the differentiation of the meaning of words is the existence of a word pair differentiated solely in terms of one phoneme (cf. *bed — bad*).   If such a pair of words exists in the language, then the sounds by which words are differentiated are phonemes possessing full relevance.   Thus, in the words *pit — bit — pat — pig* all phonemes are phonemes possessing full relevance.   If we change a sound in a word without the meaning of the word being changed (cf., *clinch, clench* with the same meaning), then both sounds are phonemes with incomplete relevance.   Thus, the same phoneme can be a phoneme with full relevance (i.e., distinguishing the meaning of two words when replaced by another phoneme) in one word, whereas in another word it can have the function of a phoneme with incomplete relevance (i.e., if it is replaced by another phoneme, it does not distinguish two words).

The problem of the acoustic identity of the word is part and parcel of the problem of the existence of the word as a linguistic unit.   We are accustomed to apprehend words as linguistic units of a firm formal structure.   This conception seems to be confirmed by the written form of the word, which changes only seldom.   But the crucial question is that of the structure of a SPOKEN word.

---

voice opposition of final consonant phonemes, the voiced being replaced by the unvoiced member of the opposition.   Cf. the pronunciation of the words *let* and *led* in Czech: *let* [let] 'the flight' — *led* [let] 'the ice', and of the graphically identical words in English: *let* [let] — *led* [led].

Can it be guaranteed that the same speaker will always articulate all the sounds of the same word and the word itself as a whole in the same way? Certainly not, and the more speakers there are, the less chance there will be of this happening. What then can guarantee the acoustic meaningful identity of the word? According to W.L. Graff,[34] neither a certain quantity nor a certain quality can constitute a substantial basis for the acoustic identity of the word. Both the number and the kind of sounds required by the word are determined by external conditions and circumstances. Graff is aware of the varying communicative value of the particular sounds of which words are composed. Thus he writes that if we replaced /t/ in *cut* /kʌt/ by /θ/ or /k/ by /t/, the acoustic distortion of the word would undoubtedly be felt, but would be tolerated much more than the replacement of /k/ in *cut* /kʌt/ by /g/ or the replacement of /t/ by /d/ or of /ʌ/ by /æ/ in the same word. However, Graff does not really tackle this highly interesting problem of acoustic identity of the word; he is merely stating that, in spite of the above-mentioned restrictions, the phonetic word symbol is limited to a large extent in its extension and quality, because it is surrounded by other words in the system.

If we examine this problem from the point of view of communication we see that it is also a problem of the redundancy of sounds. It is not possible, of course, to omit any sound of the word, as not all sounds have the same communicative value, even when they are phonemes. If, for instance, in Persian we can use the reduced form *čār* instead of the word *čahār* 'four', it must be due to the redundancy of the sequence *ah*. We could quote from the phonological development of languages a lot of examples of a similar sound reduction.[35] According to Vladimír Skalička,[36]

---

[34] "The Word and the Sentence", *Language*, 5, 1929, 164 ff.
[35] However, we must distinguish between the sound reduction which implies a definite, lasting change of the word and becomes a norm, and a reduction which takes place in individual cases only and is the result either of rapid speech or of negligent pronunciation or of other factors connected with individual speakers.
[36] "O příčinách jazykových změn" [On Causes of Linguistic Changes], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 72.

"it is a characteristic of a great number of sound changes that they mean a reduction, a lessening of the number of sounds in the word...". The important question is, however, whether the sounds that may be omitted without the meaning of the word being changed are phonemes. There can be no doubt that only those sounds possess a full functional relevance which cannot be dropped out of the word without their communicative function being substantially impaired. In the word *čahār* the phonemic function of the sounds *a* and *h* must have evidently been weakened to such a degree that it was possible for both sounds to be dropped without the meaning of the word being changed or the word becoming unintelligible.

In the system of the language these redundant sounds can be phonemes only if they occur in other words from which they cannot be dropped and in which they have their full functional relevance. This subsystem is connected to the system of the language as a whole. It is bound to it by various laws which it cannot infringe or violate. It is therefore necessary to reconsider the concept of the phoneme, since only those phonemes in the system of the language which belong to its sound inventory have full functional relevance. Interchange of two phonemes does not always mean word differentiation. Often it implies merely a distortion of the phonic structure of the word without any change of meaning.

Let us remember that as early as 1948, at the VIIth International Congress of Linguists in Paris, B. Pottier pointed out the distinction between CONSTITUTIVE PHONEMES (e.g., *s* in the Spanish words *cosa, bolsillo, son, pesadas*) and SIGNIFICATIVE PHONEMES (the second *s* in *cosas, pesadas* signifies the plural; the *s* in *tienes* signifies the second person singular of the verb; these phonemes neither contribute to the basic meaning of the word in question, nor do they disturb it). Pottier's distinction rightly indicates the basic difference in the relevance of both kinds of phonemes. However, as shown before, it is necessary to proceed still further into the differentiation of the relevance of phonemes. If, in Persian, the word *čār* had existed beside the word *čahār* in a different meaning, the reduction of sounds in the word *čahār* could hardly have taken

place. However, it is not impossible to encounter such a case in a language which has a tendency towards the formation of homonyms.

From all that has been said above we can conclude that the distortion of the word has certain limits. Beyond those limits the identity of the word disappears; this means that either a new word or a distorted sound complex appears. If it is only a question of an isolated word, the matter is quite a simple one, because the limit to which the word may suffer distortion without any change of meaning depends on the sound which replaces a certain sound of the word; i.e., it depends on whether the former sound does or does not change the phonic structure of the word to such an extent that a word of another meaning does or does not originate. Whether a new word of a different meaning originates from the replacement of a sound depends on the distribution of the sounds over the particular sound positions of a certain word-type.

How different a value the sound changes (sound differences) may have is shown by Roman Jakobson,[37] who quotes as examples the differences between /big/ and /bi:g/ with an emphatically prolonged vowel, which is a conventional, coded linguistic feature, and the difference between the short and long vowel in Czech /vi/ 'you' and /vi:/ 'he knows'. In the first case (/big/ — /bi:g/) the difference i/i: is not phonological but only emotional, whereas in the second case (/vi/ — /vi:/) the difference i/i: is phonological. This can be supported by another argument to the effect that the distinction i/i: does not occur in the neighborhood of $b — g$, and therefore it does not violate the acoustic identity of the word, whereas the same difference in the neighborhood of $l — v$ violates the acoustic identity of the word, which is phonemic in the pair *live — leave*. Thus we see that the phonetic distinction between sounds plays a secondary role in the corruption of the word structure.

The most important thing is to decide to what extent the phonemic oppositions between the sounds of a particular position in

[37] "Linguistics and Poetics", Roman Jakobson's Concluding Statement at the Conference on Style, Indiana University, April 1958, 8.

a particular environment in a particular sound pattern are exploited by the system of the language. If every language exploited all its sounds to one hundred per cent in a functional respect, i.e., if for the purpose of word differentiation it exploited all possible combinations of sounds in the different sound positions of the word pattern, then any corruption of the phonic form of the word would necessarily lead to a change of meaning. In this case, redundancy would be zero. When the corruption of the word does not mean a change of meaning but only a certain obscuring of it, it is important to state whether the corrupted word still has any link with the original word in respect of meaning or whether it is a distorted word with which the original word cannot be identified. However, no objective general criterion can be established here according to which it would be possible to state when the word is still comprehensible, as this depends to a considerable degree on the individual comprehension of the word by individual hearers. There are people who can identify a word which is distorted to a high degree, whereas to other people the same word is incomprehensible even when the distortion is small.

The distortion of the phonic structure of the word can be much greater when it is in context than when it is isolated, without the distorted word losing its identity. In this case, the degree of redundancy is considerably greater. The measure of the permitted distortion depends on the mutual relation of words in context, as words are co-determined by the context. Sometimes a word in context may suffer only a slight distortion, but it becomes unintelligible not only by itself (in isolation), but also in relation to the other words of the context. As this word is of key importance for the context of the sentence, its distortion has an influence upon the meaning of the whole sentence. Consequently, such a word is not redundant in the given connection. The extreme case is when one or more words are omitted from a sentence without the meaning of the sentence being changed. Such a case is the omission of the conjunction *that* or of the relative pronoun in the accusative in English.

The more conventional and more frequent certain word connect-

ions are, the greater is the redundancy of sounds and words.   It is
the communicative relevance of words and their arrangement in a
message (e.g., a telegram) which matters.   Thus, in the sentence
I wish you good night the word *good* has the smallest semantic
communicative relevance, because it is very closely associated with
the word *night* as well as with the words *morning, afternoon, evening*.
The word *good* forms word groups with those words, and, as it is
hardly conceivable that word groups could exist in which the word
good in connection with the preceding sequence *I wish you* could
be replaced by a word with an opposite meaning, the word *good*
can suffer corruption to a great extent without losing anything of
its communicative value.   Therefore, it is possible, e.g. in Czech,
to use reduced forms such as *'brý den* (*dobrý den* 'good day') or
'brýtro (*dobré jitro* 'good morning'), etc.   Consequently, *I wish you*
may be dropped, as is usually done.

    The considerable degree to which the utterance is clarified by
the situation may be well exemplified by those cases in which one
word stands for a whole sentence.   We have not in mind one-word
sentences of the type *Fire*!   Although these tell us something about
fire, their content is quite vague.   We are thinking rather of those
cases where one word represents quite a concrete sentence.   Thus,
in German the word *Mahlzeit* represents the sentence *Ich wünsche
Ihnen gute Mahlzeit* or the word *Tag* represents the sentence *Ich
wünsche Ihnen guten Tag*, etc.   The concrete situation is so unambi-
guous that a single word can stand for a whole sentence.   Of cour-
se, we are aware of the fact that such expressions have a more or
less expressive coloring.   Moreover, they may be taken as a whole
which we do not analyze.

    A special category of such one-word sentences is that of com-
mands which also represent concrete sentences and are not analyz-
ed.   It is doubtful whether we can regard as sentences inscriptions,
headlines, headings, titles of books, chapters, etc.   In this paper,
however, we are not concerned with this problem, for our interest
lies only in the informational value of these expressions.

    In dialects the phonic structure of the words of the literary
language is often 'distorted'.   It is, of course, a distortion only

from the point of view of the person who knows the literary language but does not know the dialect. From the point of view of the person who knows the dialect well or for whom the dialect is his or her mother tongue, no distortion of the word can be spoken of, as it is simply a word which differs acoustically from the corresponding word of the literary language. After all, if the phonic structure of words were not unstable, no linguistic changes would take place, and, consequently, no dialects would arise.

To sum up, the acoustic identity of the word is dependent upon the functional relevance of particular phonemes of the word in their interrelations not only within the framework of the same word, but also — where words in context are concerned — within the framework of the semantic relations between the particular words of the sentence. The closer the semantic relations between the words of the sentence, the greater the permissible acoustic distortion of the words without their acoustic identity being endangered.

It is precisely because a language does not exploit all the possible sound combinations in its words that redundancy is possible. Moreover, it is a well-known fact that redundancy is necessary to facilitate comprehension. If there were no redundancy in the language, the slightest distortion of the phonic structure of the word would lead to a semantic distortion of the word, and, moreover, the attention of the hearer would have to be continuously fixed on the utterance of the speaker without there being any possibility of 'enjoying' the 'rest' which is available when redundant components are present in the utterance. Consequently, redundancy is not an unnecessary or useless phenomenon in language; on the contrary, it is an organic part of language, without which the communicative function of language could not be effectively realized.

5. THE COHESION OF THE WORD

It is necessary to consider why some of the criteria of the word which we have so far discussed fail in some languages. We are of the opinion that one of the main reasons is a different degree of

cohesion or closeness of the components of the word in different languages.   By cohesion we mean a connection of the elements of the word which is such that it makes it impossible for any other element to enter between them; this is the highest degree of cohesion.   According to a varying degree of obstruction, according to the possibility or impossibility of inserting some other element between the elements of the word, we distinguish different degrees of word cohesion.   In this respect languages behave differently.

The importance of the cohesion of word elements has been emphasized by J. Vachek, who, in his lecture on the position of the word in the systems of Czech and English (in 1960), criticized F. Mikuś[38] for underestimating the difference in the degree of cohesion of word elements in various languages.   According to Mikuś, there is no essential difference between the components of the English verb form *I sing* and the French *je chante*, on the one hand, and between the components of the Latin form *cant-ō* ('I sing'), on the other hand.   In each of these two categories the only essential fact, according to him, is that it is merely the relation between a determinandum (*sing, chante, cant-*) and a determinatum *(I, je, -ō)* which is involved.   Further, according to Vachek, the cohesion in Latin, Russian, and Czech is so great that both components of the word cannot be separated by any other component existing independently apart from the given word.   On the other hand, the cohesion of word components in English and French is considerably smaller, as they do not offer resistance to a mutual separation of this kind (cf., *I very often sing, je le lui chante*).   However, it is necessary to give precision to Vachek's statement; the cohesion of the word in Czech is admittedly not so great, as in Czech the components of the word can be separated by inserting another component, not, of course, a component which exists independently apart from the given word, but a dependent component, a grammar morpheme.   Thus, in Czech we have *zpívám* 'I sing' — *zpívávám* 'I am accustomed to sing', etc.   By inserting a grammar morpheme between two components of a word

[38]   Cf. F. Mikuś, "En marge du Sixième Congrès International des Linguistes", *Miscelánea homenaje a André Martinet*, I (Tenrife, 1957).

in Czech we do not obtain two words which are separated by the inserted morpheme but a new word whose meaning is related to that of the original word. We should be inclined to say that the word is more solid in English than in Czech. Though the English word can be expanded by prefixes and suffixes, like the Czech word, the English word is inwardly compact; no element can be inserted inside a simple English word.[39]

A degree of cohesion similar to that in Czech may also be found in Modern Persian. In this language causatives are formed by expanding the stem of verbs by the morpheme *-an-*, e.g., *fahmīdan* 'understand' — causative *fahmānīdan*. In Turkish, quite a number of infixes can be inserted between the root and the suffix, e.g., *evde* 'in the house', *ev-ler-de* 'in (the) houses', *ev-ler-im-de* 'in my houses', *ev-ler-imiz-de* 'in our houses', etc. From Portuguese we can again quote the earlier mentioned example of *darei* as compared with *dar-vos-ei*. A smaller cohesion of word elements is evidenced also by instances in a language when two word forms express the same morphological function, both synthetically and analytically. This is the case in Turkish, where *açtım* 'I was hungry' exists side by side with *aç idim*, or where *yorgundum* exists side by side with *yorgun idim* 'I was tired'. The prepositions *için* 'for' and *ile* 'with' can function independently or as parts of words; in the latter case they are subject to vowel harmony and are modified into *çün, çin* and *le, la*, respectively. Examples: *Yusuf efendi bahçesiyçin çok güzel çiçekler aldı* 'Mr. Yusuf bought for his garden very nice flowers'; *Bu çay benim için midir?* 'Is this tea for me?'; *Bunlar ile* (or *bunlarla*) *geldim* 'With those I came'. From the written language a looser or closer cohesion of different elements of the word can also be inferred.

Different degrees of word cohesion in Japanese are presumed by J.V. Neustupný:[40]

[39] As for the problem of compounds, E. Kruisinga (*Handbook of Present-Day English*, II, 3, 1932, 4) points to "various degrees of closeness, so that attempts strictly to divide compounds and syntactic groups would be contrary to the character of the phenomenon". This is in agreement with our concept of the cohesion of the components of the word.
[40] J.V. Neustupný, "Accent in Japanese and Russian", *Ar. Or.*, 27 (1959), 131.

It is true that the cohesion of the lexical (plus conjugational) morpheme and the morpheme of declension is not so close as that of the verb lexical morpheme and the conjugational morpheme in Japanese or as the cohesion of the Latin, Czech or Russian declension.   However, we think that the opposition "word — a group of words" is gradual in Japanese.   It is strongest with conjugation, but still the declension and perhaps also such groups as *hanāshite kudasai* or *tàbete iru* are one word.

We can conclude that the cohesion or closeness of word elements can vary not only in different languages but also within one language.   Thus, in English the suffix of the group genitive seems to be looser than the other affixes.

In our view, the varying cohesion of word elements is one of the basic factors influencing not only the phonological but also the morphological and syntactical structure of language and is, consequently, one of the important typological factors as well.

IV

FORM AND MEANING

The problem of the word is to a considerable degree a problem
of mutual relation between form and meaning. Let us consider
two extreme cases: form without meaning and meaning without
form. The first extreme, form without meaning (in the sense in
which we currently understand the concept of 'meaning'), can
exist as a quite formal means of designating some grammatical
relation. Whenever we refer to 'meaning' in this chapter, we shall
not be thinking of the 'grammatical' meaning (for instance, the
'meaning' of prepositions, conjunctions, pronouns), because the
'grammatical meaning' is a concept quite different from that of
meaning in its usual sense, that is, the designation of things,
qualities, and activities. It would be better to speak of a 'gram-
matical relation' than of a 'grammatical meaning'. Only in this
sense is it possible to speak of form without meaning. On the other
hand, it is difficult to speak of meaning without form. If we are
inclined to accept the opinion that thinking is not possible without
language,[1] then even pure meaning is not possible. Meaning is

---

[1] There are various opinions as to whether thinking without language is
possible. We will not discuss this problem in detail but only quote here the
opinion of Robert M. W. Dixon (*What Is Language? A New Approach to
Linguistic Description*, London, 1965, 163 ff.) with which we agree in principle:
"*An animal or person does not have to be familiar with language patterns in
order to think: but for anyone who is a language user thinking is bound to be
tied up with language ... we may be able to recognise some non-language patterns
as 'logical' merely on the basis of the action involved, and without thinking of
language patterns that could describe the logical action. But it does seem that,
as our thinking about a certain pattern becomes more directed and detailed, the
language patterns are far likelier to enter in. Only a limited degree of directed
pattern-correlation is possible in terms of non-language patterns; language
patterns enable the direction to be taken much further than would otherwise be*

always connected with some form. If we isolate the words express-ing only a grammatical relation, we see that the relation between form and meaning is close and that words mutually differ or agree in both a formal and a semantic respect; two words can differ in both form and meaning; or in only their form, the meaning being the same; or in their meaning, the form being identical. From one to another of these three possibilities there is, of course, a large degree of transition, as is illustrated by the following diagram:
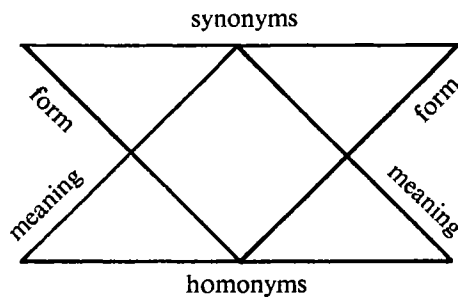
Two extremes, synonyms and homonyms, designate the upper and lower limits; between them there is a transition. The meanings of two words merge in synonyms, and similarly the forms of both words merge in homonyms.

The relation between form (that is, the acoustic form of the word) and meaning is also the subject of the earlier mentioned study by W.L. Graff. He discusses the question of when the phonetic difference results in word differentiation. And he comes to the conclusion that a purely phonetic difference is not the cause of word differentiation if the two sound combinations are felt to be

---

the case, so that all directed thought must be language-oriented; the reverse does not hold: some language thought is not at all directed". With regard to the impossibility of a scientific examination of non-language thought Dixon writes: "... non-language thought will normally occur fairly spontaneously, and as soon as we become aware that it is non-language thought and direct our attention to examine it, language patterns will have crept in, so that we are only really able to analyse them."

We call the reader's attention to Acta Psychologica, X (1954), Nos 1-2, which is devoted exclusively to the problem of thought and language.

closely related.   If two phonetic forms are not felt to be closely
related, we obtain synonymous but formally different words, e.g.,
heaven and *sky*.   Within the same category, word differentiation
is caused by the semantic difference of a complex of identical
sounds in cases where the linguistic community no longer feels the
difference created when one reference is replaced by another.
Otherwise, they are variants of one and the same word or they are,
in fact, one and the same word.   Thus, the English words *top*
'highest part, summit' and *top* 'a kind of toy' are, according to
Graff, two words, whereas the German *Himmel* 'heaven' and
Himmel 'sky' are one and the same word.   When the reference
and the sounds represent partial differences, different words are
involved in cases where either the semantic or the phonetic differ-
ences, by themselves or combined, are such as would have resulted
in a word differentiation under previous conditions.   For instance,
the German words *schlecht* 'bad' and *schlicht* 'plain' represent two
different words, though phonetically they are very close.   On the
other hand, let us mention the English words *clinch* and *clench*,
which, from the semantic point of view, must be regarded as one
word represented by two phonetic forms.

Further, we must, according to Graff, distinguish between a mere
categorical difference and a semantic-phonetic-morphological
difference.   If only a semantic categorical difference is involved,
the question of a differentiation or relative identification is a very
delicate one, as the feeling for the relative relationship of both
categories can vary in the language in question.   Thus, in English
the adjective is felt to be categorically close to the noun, whereas
in German it is close to the adverb.

Graff's characterization of the relation between the word and the
sentence is also interesting.   He writes (p. 182):

> If we look for the word within the sentence, we must be prepared to
> find perhaps no more than relatively identifiable debris which can only
> be put together and expanded with the help of the system.   This shows
> how utterly impossible it is to define the single word as long as we look
> for it within the sentence...   The single word is a systematic unit and
> the word-within-the-sentence can only mean the word-material used by
> the sentence in its own characteristic way.

An example of the way in which the sentence treats the word-material is provided in German by the verbs with separable prefixes (e.g., *stattfinden* 'take place' — *ich finde statt* 'I take place'), which, being in a sentence, are often divided into two parts which are separated from each other without, however, their systemic unity being disturbed. Verbs with separable prefixes are composites, the components of which, however, do not lose their mutual relation even when separated in the sentence (therefore, we speak of 'distant' morphemes).

On the basis of this phonetic-morphological-semantic standpoint Graff supplies the following definitions: (a) a SIMPLE WORD is a word which is felt by the language community to be indivisible into smaller semantic-phonetic parts or not to result from them synthetically; (b) a MORPHOLOGICAL WORD is a word which is felt by the language community to be the result of a combination of two smaller semantic-phonetic parts of which at least one cannot become the material of a sentence without the aid of a further systematic structure; (c) a COMPLEX WORD is a word which is felt by the language community to be the result of a combination of two simple or morphological words the phonetic elements of which are united according to a systematic pattern and whose every reference-context appropriates a part of the other.

To give a general evaluation of Graff's study we can say that he uses the correct prerequisites as his stating point, fully realizing that analyses of language units have so far failed because undue emphasis has been placed on either the phonetic or the semantic features to the detriment of the semantic-phonetic combination; he further appreciates that the relation between the words and the sentence has been wrongly interpreted, that the character of the word is often identified with its quantitative extension, and that facts relating to the evolution of language have not been strictly separated from those relating to the static system. However, the question is whether Graff has succeeded in removing the obstacles of which he is so well aware. As we have seen, he has failed to find a universal definition of the word on the basis of his complex phonetic-morphological-semantic view. Instead, he presents a

classification of various kinds of words. He is successful in those parts of his work in which he discusses the acoustic aspect of the word. The importance of his work lies in the fact that he has understood and suggested the significance of the phonetic approach to the problem of the word even if this aspect of the problem has remained mostly on the periphery of his study. When he deals with meaning his arguments and conclusions are not very convincing. That is to say, his criterion is often the linguistic feeling, which is a criterion that cannot be considered objective, and even if we accept it with certain reservations we cannot regard it as the basic criterion.

1. THE FUNCTIONAL THEORY OF MEANING AND FORM

The problem of the relation between form and meaning has been treated far more successfully by W. Haas.[2] According to Haas, it is not sufficient to investigate just the distribution; it is also necessary to analyze functional relations. In our opinion, Haas is right to say that the core of the problem is to differentiate form and meaning without disturbing the unity and identity of the language sign. Haas sees the solution of this problem in the so-called functional theory of meaning and form.[3] He uses as his

[2] W. Haas, "On Defining Linguistic Units", *Transactions of the Philological Society*, 1954, 60.
[3] Haas gives a good survey and evaluation of present theories dealing with the relation between meaning and form. The traditional theory is the *theory of correspondence between meaning and form* which presupposes that the linguistic sign emerges from the correspondence — a kind of psycho-physical parallelism — between form and meaning. A rougher and more traditional form of the theory of correspondence represents every language sign simply as an *association* of a physical with a psychological fact (Jespersen). The *substitution version* of this theory (L. Hjelmslev) should guarantee that of the mass of physical and psychological facts ('expressions' and 'contents') only relevant facts are admitted. Substitutions must be carried on along two separated lines: along the line of 'expression' and along the line of 'content'. It is simply not allowed to interchange signs, the substitution of forms for forms and meanings for meanings having to take place separately. Hjelmslev assumed that the identity of the linguistic sign would be asserted if both operations of

starting point J.R. Firth's study,[4] according to which meaning
must be considered as a complex of contextual relations, phonetics,
grammar, lexicology, and semantics each dealing with its own
components in an appropriate context.   To obtain the form of the
sign we analyze it and to obtain its meaning we extend it and
place it in the context of units of a higher level.   Thus, the sign
plays a distinctive role in higher contexts.

Characteristic for the functional theory of meaning and form
are the following facts:

---

the substitution mutually corresponded; the form ('expression') should be
regarded as relevant only when its substitution by another form corresponds
with the substitution of one meaning by another. On the other hand, it is
necessary to regard a meaning ('content') as relevant only if its substitution by
another meaning corresponds with the interchange of forms. Forms and
meanings are then in a mutual relation of 'commutation'. In his evaluation
Haas writes that the application is problematical and extremely uncertain.
We are not operating here with signs but (a) with physical formations and (b)
with disembodied concepts and ideas. Whereas we know how to substitute
physical formations mutually by selecting signs and exchanging them, it is
difficult to imagine how to select meanings and to exchange them if the latter
operation is to be different from the former. We are confronted not with two
entities, expression and content, but with one, viz. the sign. *The distributional
version* ('morpheme components') is contained in Harris's *Methods in Structural
Linguistics* (Chicago, 1951, Chapter 17, § 3). According to Haas, this theory
contains some elements that conform with the theory of sign. The procedure
is as follows: different genders (components *he* or *she* in English) are extracted
by observing the differences in the environment (rather than by means of the
semantic contrast of substitution in the same environment). The word *bull*
and not *cow* occurs in the environment *He is a —, The — raises his head,* whereas
cow and not *bull* occurs in the environment *She is a —, The — raises her head.*
The advantages of the distributional interpretation of meaning are undeniable.
The semantic difference is observed here as a difference in the environment,
in fact as a difference in function. Meaning, though it is in fact defined by
the function of the sign, is presented here as if it were contained within the
limits of its form, as the 'content' of the form. Thus, according to Haas,
sounds and meanings are here forced into a kind of parallelism, an empty
framework of forsaken psycho-physical theories of language. There is no
parallelism between form and meaning and we gain little if we change the
definition of meaning by replacing 'notions in the mind' by distributions and
assuming that the meaning of the sign must be sought in itself, just as its form
must.
[4]   J.R. Firth, "The Technique of Semantics", *Transactions of the Philological
Society*, 1935, 53.

(1) Facts referred to in the description of the semantic value[5] of an element include its form — they are its functions. The meaning of the element represents not the content but the function of its form.

(2) Facts referred to in defining the form or meaning are accessible by means of the same method of observation — they are concrete language facts in a concrete environment.

The conclusion following from this theory is important for us: the word cannot be fully characterized and defined until it is linked up with a broader context. And *vice versa*: the word is that unit of speech which can be taken out of the sentence, but as soon as we remove it from the context and isolate it, it ceases to be that which it was in the context, as it has been taken away from those relations which alone are capable of fully determining it. These relations are very complicated and it is not possible to generalize about their hierarchy in any way, as every utterance, every sentence, in spite of its formal similarity with other sentences of the same kind, is a unique event. We hardly find two sentences which were uttered under identical circumstances and in quite the same situation. Nor is the environment ever the same. Even such utterances as are represented by fixed clichés and proverbs are not always introduced in quite identical situations. Every word has two meanings: ACTUAL and POTENTIAL. The actual meaning is the meaning of the word in a given context and the potential meaning is the meaning of the word as a lexical unit. In a lexical unit all topical meanings are potentially included; however none of these meanings is entirely complete without the contextual situation. It is the contextual situation by which they are fully determined. Consequently, a lexical unit cannot contain meanings in a topical form (that is, in the form in which they occur in the context), but only in a potential form. The actual meaning is not

[5] There are three types of semantic value: (1) *distinctive* values of phonemic features and of phonemes (mere semantic value); (2) *relatively constant* distinctive values of morphemes, words and constructions (proper meaning); (3) *extralingual* and relatively constant values of sentences (sentence meaning or sense). A complete utterance contains in its context situation all three types of semantic value; it contains elements of all three levels.

that given by the word itself, but that given by its relation to all the members of the utterance in which the word in question is contained.

Thus, we should say that the word has the same importance for the level of utterance as for the lexical level. This view, however, is not generally shared. M.E. Seidel[6] sees the difference between the word and the sentence in the fact that the sentence is, above all, a unit of topical speech *(la parole)*, whereas the word is primarily a unit of the language system *(la langue)*. A. Rosetti[7], on the other hand, points to the fact that a word in a sentence expresses something other than that expressed by an isolated word. An isolated word is characterized by semantic inaccuracy and vagueness. Words in isolation have a general meaning and this places them, according to Rosetti, in the category of 'langue', but in a sentence, in an utterance, they are manifested by a great number of different meanings. According to Rosetti, in its general meaning the word is a fact of 'langue', whereas in its specific meaning it is a fact of 'parole'. This is why Rosetti complements and makes more precise the definition of M.E. Seidel by stating that the word is both abstract and concrete entity, according to whether it is isolated or used in a sentence, in the same way as the phoneme is abstract in relation to sound, whereas the manifestation of the phoneme and its realization in 'parole' is concrete.

Rosetti's comparison of the word with the phoneme leads us to state that there is a certain isomorphism between the phoneme and the word, even if this is not entirely consistent. The phoneme as a unit of the acoustic plan has no meaning; it receives its meaning as a realization of units of a higher plan, i.e., of morphemes and words. And since phonemes, as members of the structure of a certain language, are capable of combining with other phonemes according to certain rules of distribution, it can be assumed that phonemes as units of a language structure also have certain potential meanings which, of course, do not appear until they are com-

---

[6] M.E. Seidel, "Wort und Satz", *Rev. des ét. indoeuropéennes*, IV (1947), 50.
[7] A. Rosetti, "Sur la définition du 'mot' ", *Acta Linguistica*, IV (1944), 51. Cf. the same author, *Le mot* (Copenhague-Bucureşti, 1947), 38.

bined into words. Similarly, words, as units of the lexical plan, have certain potential meanings which are not concretized until they become an utterance. Thus it is possible to speak of a kind of isomorphism between the phoneme and the word.

Some linguists make a distinction between basic or primary meaning and secondary meanings. Thus, according to B. Trnka,[8] if we accept the polysemanticity of words we must ask what gives identity to the word in a language system. For example, shall we see in the form *head* as many different words and as many different meanings as there are actual utterances or shall we see only one word? The answer, according to Trnka is, that every word has its basic meaning which in actual utterances is used in different variations of meaning without its basic oppositional meaning fully disappearing. The variations of meaning connected with the phonological formation *head* are linked to the basic meaning and not directly to the extralinguistic reality. Josef Filipec[9] distinguishes primary meanings from secondary meanings which are felt to a certain degree to be derived from the background of primary meanings. The primary meaning gives the secondary meaning a special coloring and indication. We should, however, prefer the terms basic and secondary meaning, as they better express the situation concerned. It is not so much the primariness and secondariness that matters here, but, rather, which of the meanings is BASIC. Primariness and secondariness imply a moment in time, and this is not our concern. However, we do consider it important to realize that the distinction between the primary or basic meaning and the secondary meaning can be made only in a dictionary, on the lexical level, and only from the point of view of the frequency of occurrence. In a given context we cannot speak of the basic and secondary meaning of a word, because the meaning which is used in that context is always the primary, basic meaning in that unique case.

[8] B. Trnka, *Rozbor nynější spisovné angličtiny* [Analysis of Present-Day Literary English], II: *Morfologie slovních druhů* [Morphology of Word Classes] (Prague, 1962, mimeographed), 11.
[9] Josef Filipec, *op. cit.* in fn. 26 on p. 28, 191.

2. HOMONYMY

It is remarkable how languages differ in semantic diversity of
words.   There are languages which show great semantic differenti-
ation and other languages which are only slightly differentiated
semantically.   To the former languages belongs English, for ex-
ample, where nearly every word, particularly verbs and nouns, has
a number of meanings.   We have only to mention the abundance
of meanings of such verbs as *to set, put, do, make*, etc.   A number
of further studies would be needed to clarify the semantic diversity
of particular languages.

Of considerable importance for our problem of the word are the
relations between the meaning and the form of several units.   One
can distinguish between relations of one meaning with more
designations and relations of one designation with more meanings.
The first case is that of synonymy and the second is that of polysemy
and homonymy.   Synonymy presents no problem in our case, as
synonyms are words different in form.   Homonyms, on the other
hand, are formations of identical sound forms but with different
meanings (for example, *night* /nait/ and *knight* /nait/).   In this case
the question arises as to whether one or two words are involved.

Vl. Hořejší solves the problem of homonyms on the basis of the
morphological plan of language:[10] The problem of homonyms is
in fact reduced to the opposition 'the same word/different words'.
The identity of the word is morphologically determined by the fact
that it belongs to a certain class of words and within the framework
of this class of words it expresses each of the categories in a certain
way.   Different words, on the other hand, belong to different
classes of words or within the framework of the same class express
the same categories in a different way with a difference in meaning.
For instance, the word *zámek* in Czech is one and the same word
whether it means 'the castle' or 'the lock', because it behaves
morphologically in the same way.   Two words are, however,

---

[10]   Vl. Hořejší, "Postavení morfologie v mluvnici a její obsah" [The Position
of Morphology in Grammar and Its Content], *ČMF*, 39 (1957), 85.

represented by the form *peníze* 'money' with the forms *penízů* (genitive), *penízům* (dative), etc., as compared with the forms *peněz* (gen.), *penězům* (dat.), etc.   Likewise, the word *nohy* 'feet' with two genitive forms *nohou* and *noh* is homonymous.   Hořejší's explanation of homonyms may be acceptable for Czech, but cannot be generalized for the other languages.   In several languages we meet with homonyms resulting, in consequence of the development of the pronunciation, from the merging of the acoustic form of two or even more words written in a different way, without their written form having been merged (cf., for instance, the English words *rite*, *right*, *write*).   Otto Ducháček in the discussion at the conference on scientific knowledge of contemporary languages[11] pointed out that, since in contemporary language it is not clear in many cases whether we are concerned with two meanings of the same word or two words of different origins which have been formally merged as a result only of the influence of phonetic development, the term 'homonyms' should be reserved for words etymologically different. On the other hand, when it is etymologically a question of a single word, the term 'multi-meaningful word' should be used, even if the differentiation of particular meanings is such that their con-nection is no longer felt.   If, according to Ducháček, this solution seems unacceptable from the point of view of the synchronic conception, then for wholly differentiated meanings of the same word the term 'improper homonyms' and not just 'homonyms' should be used.   Though Ducháček's proposal may make the situation more precise, it does not solve it.   Even if we were not to exclude *a priori* the possibility of a diachronic solution of these problems but were, on the contrary, to recommend the linking of synchronic with diachronic treatment in all cases where it helps to solve problems, we are still of the opinion that in this case it would neither bring any clarification nor be suitable.   In our view, the problem of homonymy is a purely synchronic problem, as it is the acoustic form of the words of contemporary language which is of real concern, and the etymological point of view will not provide any solution to the 'form — meaning' problem in which we are

[11]   Cf. *O vědeckém poznání soudobých jazyků* (Prague, 1958), 208.

actually interested.[12]   Moreover, in languages in which the graphic
form differs from the spoken one, it is in the graphic form that
we meet with complications.   Two solutions are possible: either
we entirely ignore the graphic aspect (spelling), as a result of which
the problem of homonyms becomes simply a problem of words
with different meanings, or we take the graphic aspect into con-
sideration, with the result that we must consider several particulars.
It is chiefly a question of whether the speaker and the hearer
apprehend the spelling of the word.   In a real speech situation
only the acoustic image of the word seems to exist, at least at the
moment when the word is uttered.   As soon as a sufficiently long
text is known from which may be concluded which of the meanings
is involved, the hearer can associate a certain acoustic form with
the written form.   To what degree he realizes this depends on the
semantic clarity of the context.   Consequently, the association
with the written form does not begin until the word is unambi-
guously determined by the context.   However, we do not think
that one's consciousness of the written form matters much.   In
most cases the word is fully determined by the context and this
determination is sufficient for a meaningful communication.   We
admit that there are people who do not even know the written
form of the homonyms or who interchange written forms and

[12]   Otto Ducháček speaks of a tight link between the form and the meaning
of the word (cf. *O vzájemném vlivu tvaru a významu slov* [On Mutual Influence
of the Form and Meaning of Words], Prague, 1953, 175).   Elsewhere (cf.
op. cit. in fn. 11 on p. 51, 171) he writes: "The meaning is like the soul,
whereas the form is the body of the word.   As a matter of fact, the word
consists of two components: 1. the inner (ideal) component which is its meaning
(content, extent, and feeling or sentiment) and 2. the outer (material) component
which is its form (in speech a group of sounds, in writing a group of letters)."
Vl. Skalička (cf. his paper "Komplexnost jazykových jednotek" [Complexity
of Linguistic Units], *Universitas Carolina, Philologica*, Vol. 3, No. 1, p. 19)
writes that word is "... a unit where form and meaning intertwine.   The
meaning of the word, i.e. the denotative meaning, is parallel to the word, but
does not correspond to it — it forms a proper unit, designation.   On the other
hand, the word is determined by a number of formal features, none of which,
however, is a hundred-percent valid: the word is given by tact, but there are
enclitics; the word is given by the phonological constitution, but often it is
not differentiated in this way; the word is given by the possibility of displace-
ment in the sentence, but often it is possible to displace even parts of words."

corresponding meanings without being hampered in speech com-
munication.   Of course, there may be cases where even the context
does not determine the word unambiguously, in which case we are
faced with confusion or puns.

There are reasons for our choice of the first solution.   When one
word is involved, we are dealing only with different meanings of
the same acoustic forms of the word, assuming, of course, that it
is a case of words of the same category.   There can only be
homonymy of grammatical morphemes and homonymy of words
belonging to different word categories.   For example, *right* is an
adjective, adverb, noun, and verb, but it can be identified with the
verb *to write* only in its verbal function.   If /rait/ has, for instance,
the function of an adjective or adverb, it cannot be identified with
the identical /rait/ having the function of a noun or verb.   *Lain*
and *lane* are acoustically identical formations but different words
and, consequently, homonyms, because they belong to different
word categories and in context they are always unambiguously
determined.   The words *night* and *knight* must, on the other hand,
be regarded as identical, as they belong to the same word category.
Knight, of course, may also occur in a verbal function, in which
case it is homonymous in relation to the noun *night*.

On the lexical level, however, even the graphic form of words
must be taken into consideration.

3. THE WORD IN THE WRITTEN LANGUAGE NORM

There is no doubt, especially in a culturally well-developed language
community which has created a written norm of the language, that
there is a very close relation between the written and the spoken
norms of the language.   Professor Vachek (in his lecture given at
the conference of Czechoslovak linguists at Liblice in 1954) speaks
of the correlation and correspondence between both norms.   We
will not discuss here the correspondence between both norms in
particular language plans, as it does not directly relate to our
problem.   After all, the correspondence between the spoken and

written norms of the language is dealt with in a number of excellent studies.[13] For the purpose of our work, that is, for the delimitation of the word as a linguistic unit, the question of the delimitation of word units is, however, of some importance, as it arises in the written norm of various languages.

In the written language norm we usually meet with units which correspond to formations which are generally regarded as words. It is not always so, however, especially with compound words, as in English. English compounds acting as junctions of two or more words are written as one word or with a hyphen or separately. There is no unified plan in this and in different dictionaries different ways of writing these words can be found. Nevertheless, this does not mean that no unified plan could be drawn up or, on the other hand, that the differences have no justification. Compound words possess different degrees of cohesion which are expressed by writing the elements of the compound as one word or with a hyphen or separately. There are some compounds the elements of which are always written as one word, as for instance, *bookseller, bookmark, headlong, airtight, waterproof, bricklayer* (but *taxi-driver*), *quartermaster*, etc. In this case the junction between the elements of the compound is evidently closest. There are, however, many compounds the writing of which varies. It is, for instance, the attributive connection which determines whether the writer regards the whole attributive expression as a whole or as a free connection of two or more words. Thus, V. Fried[14] quotes the following examples: *Covent Garden*, but *Covent-Garden Opera*; *the family was poverty stricken*, but *a poverty-stricken family*; *nineteenth century* but

---

[13] Cf., e.g., A. Artymovič, "Fremdwort und Schrift", *Charisteria Gu. Mathesio ... oblata* (Prague, 1932), 114 ff.; Josef Vachek, "Český pravopis a struktura češtiny" [Spelling and Structure of Czech], *Listy filologické* 1933, 287 ff.; "Zum Problem der geschriebenen Sprache", *TCLP*, 8 (Prague, 1939), 94-104; "Psaný jazyk a český pravopis" [Written Language and Czech Spelling], *Čtení o jazyce a poesii*, I (1942), 231-306; Kâre Kaiser, *Mundarten und Schriftsprache* (Leipzig, 1930).
[14] V. Fried, *Anglická interpunkce* [English Punctuation] (Prague, 1954, mimeographed).

a nineteenth-century novel; *five years* but *a five-year plan*; *upper middle class*, but *upper-middle-class life*, etc.   Other examples of variation are such as *sleeping car* or *sleeping-car*, *rolling mill* or *rolling-mill*, *dining room* or *dining-room*, *sitting room* or *sitting-room*, *an ill-built house* or *an ill built house*, *a well-known person* or *a well known person*.   If *well* or *ill* has a function as the adverbial determination of the verb, the hyphen is not written, e.g. *he was well bred, my friend got well known*.   In a number of cases writing the hyphen contributes to a semantic distinction: *forget-me-not* as against *forget me not*, *a bull's-eye* as against *a bull's eye*, *love-lies-bleeding* as against *love lies bleeding*, *four-crown stamps* as against *four crown stamps*, etc.   Here a clear distinction exists between a compound and a group of two or more words.   For more details see Chapter V, 1.

   In Czech the development of the spelling norm is moving in the direction of the correspondence of the written norm with the conceptual unity or diversity.   Thus, formerly, only *s počátku* 'at the beginning' was the norm, whereas nowadays both *zpočátku* and *z počátku* are admissible forms.   Likewise, *naboso* and *na boso* 'barefooted', *nadobyčej* and *nad obyčej* 'unusually', *nadrobno* and *na drobno* 'in small pieces', *naprázdno* and *na prázdno* 'idly', *naruby* and *na ruby* 'inside out', *přesčas* and *přes čas* 'overtime', but only *nadobro* 'entirely, completely', *nadosmrti* 'for the term of one's life-time'.   Since the forms written as one word are regarded as more progressive, they are quoted first in 'The Rules for Czech Spelling'.   It may be assumed that in the above-mentioned cases the development will follow the tendency to assert only one way of spelling, viz., as one word.   However, there are many cases in which the double spelling, as one word and as separate words, destroys the homonymy of forms such as *načisto* 'completely' as against *na čisto* 'clean (copy of)', *nadmíru* 'exceedingly' as against *nad míru* 'beyond measure', *nahoře* 'up, upstairs, above, etc.' as against *na hoře* 'at the top of the hill', *nato* 'thereupon, afterwards' as against *na to* 'upon it', *natolik* 'in such a degree' as against *na tolik* 'so much', *navrch* 'up, upwards' as against *na vrch* 'up the hill', *nejednou* 'many a time' as against *ne jednou* 'not once', *nejinak*

'just like that' as against *ne jinak* 'not otherwise', *opravdu* 'really, in fact' as against *o pravdu* 'for truth', etc. It is difficult to give English equivalents of these Czech forms. In these cases we are concerned mostly with the differentiation of an adverb (one word) from the connection of a noun and a preposition (two words). Likewise, no hyphen is written in compounds in which either both components form a close semantic unity, for instance, *hluchoněmý* 'deaf and dumb', *sladkokyselý* 'sour-sweet', or the first member of the compound determines the other, for instance, *církevněslovanský* 'Church-Slavonic', *literárněvědný* 'literary-scientific', *občansko-právní* 'concerning civil law', etc.

The written norm is often at variance with the delimitation of the word units. Thus, whereas in Czech the negation of the verb is written together with the verb as one word, in Russian the negative particle *ne* is written separately (cf., the Czech *neznám* with the Russian *ne znayu* 'I do not know'); likewise, in Serbo-Croatian (with the exception of the verbs *biti* 'to be' and *imati* 'to have', e.g., *nije* '(he, she, it) is not', *nemam* 'I have not', but *ne razumijem* 'I do not understand'. However, it is not certain whether the Russians understand this connection as one or two words. The situation is even more problematical in Serbo-Croatian. In Turkish the forms of the verb 'to be' in the present tense are written together with the word to which they belong, e.g., *Bu ev güzeldir* 'This house is beautiful', but *Bu ev güzel değildir* 'This house is not beautiful'. In the latter case *dir* is connected with *değil* but not with *güzel*. As the English 'I was not at home' is translated into Turkish as *Evde değil idim* or *Evde değildim*, it is evident that the negative particle *değil* behaves as an independent word, whereas the copula *dir* behaves as an affix. On the other hand, 'I had a house' can be translated into Turkish as *Evim var idi* or *Evim vardı*. In the past tense the verb 'to have' is, accordingly, expressed by means of *var* + *idi* or *vardı*, the latter being composed of *var* plus the copula *idi*. In evaluating word delimitation in Turkish we can say that it is the vowel harmony which plays the basic role, both in the spoken and in the written norm. Of course, we also meet with violations of vowel harmony, but they are not

very frequent, so that, essentially, it may be said that in Turkish
the vowel harmony characterizes the word unit.

   The number of examples which we have given from a few
languages could easily be increased.   However, they are sufficient
for us to come to the conclusion that the graphic aspect of language
utterance plays an important part in the delimitation of the word
unit.   And, since the graphic aspect of language is secondary
whereas the spoken utterance is primary (that is, the graphic norm
should be an image of the spoken norm), we can to a great extent
take the delimitation of the word units as the basis of the graphic
norm.   This conclusion is not valid without certain reservations
and limitations.   The graphic aspect of language cannot be
regarded as a perfect determinant of the spoken language.   Neither
is it invariable or unchangeable; it is always developing, even if,
relatively speaking, this development is slight and slow.   There can
be no doubt that the development of the graphic norm of language
tends to increase the informative value of the graphic expression.
It is the task of the graphic rendering of a spoken language to
segment the continuous flow of speech into units so advantageously
that the graphically fixed image of speech is given the maximum
informative value.

4. THE WORD AS A LINGUISTIC SIGN

One of the concepts about which linguists differ considerably is
that of the linguistic sign.   There are as many opinions as there are
different definitions of the sign.   Let us mention here just two
important discussions of the linguistic sign.   The first concerned
the arbitrariness of the linguistic sign and was published in the
first four volumes of *Acta Linguistica*; the other took place at the
International Symposium at Erfurt in 1959, where among a number
of other questions was discussed the question of what is to be
understood by a linguistic sign.   Because this is a question of basic
importance for our problem, we will report in more detail the
replies of some of the participants at the symposium in answer to

this question.   Thus, according to Wilhelm Appel:[15] 'Ein Zeichen ist etwas Wahrnehmbares, das willentlich gesetzt und mit einer bestimmten Bedeutung verbunden ist'.   V.A. Artemov:[16] "Unmittelbar für sich allein hat das Sprachzeichen keine sprachliche Bedeutung.   Davon kann man sich leicht überzeugen, wenn man mit einem Menschen in einer ihm unbekannten Sprache spricht." Jaroslav Bauer:[17] "Unter einem sprachlichen Zeichen verstehe ich jede Lautgebilde, das für den Träger der gegebenen Sprache mit einer bestimmten Bedeutung in fester Verbindung steht."   Jeffrey Ellis:[18] "A linguistic sign is a relation made in the mind (the product of a refined kind of conditioned reflex) between (expression) any token of a type (range) of sound-waves heard (or articulatory movement made) or written marks seen (or coding of these into other media) and some range of 'meaning'."   Karel Horálek:[19] "Das sprachliche Zeichen ist im Grunde ein konventionelles (unmotiviertes) Verständigungsmittel.   Die Konventionalität ist durch das Verhältnis der sprachlichen Lautform (Wörter, Morpheme) zu Bedeutung (und zugleich zur aussersprachlichen Wirklichkeit) gegeben."   Vladimír Skalička:[20] "Das sprachliche Zeichen ist eine Einheit, die phonologische Eigenschaften und Bedeutungseigenschaften besitzt."   B. Trnka:[21] "Das sprachliche Zeichen ist jeder akustisch-physiologischer Lautkomplex, welcher einen bestimmten Bedeutungsinhalt besitzt.   Ein solches Zeichen ist die Sprachäusserung."   Pavel Trost[22]: "Unter einem Sprachzeichen versteht man allgemein ein einzelnes Lautgebilde, das eine Bedeutung innerhalb einer für eine gewisse soziale Gruppe gültigen umfassenden Konvention hat."   We have quoted these views in order to show the diversity of the definitions of a linguistic sign.   Where most definitions agree — and this is important for our problem —

[15]   Cf. *Zeichen und System der Sprache*, I (Berlin, 1961), 16.
[16]   *Ibid.*, 29.
[17]   *Ibid.*, 34.
[18]   *Ibid.*, 47.
[19]   *Ibid.*, 76.
[20]   *Ibid.*, 139.
[21]   *Ibid.*, 143.
[22]   *Ibid.*, 148.

is in the relation of the linguistic sign to the meaning.   There are few definitions which reject meaning in relation to sign.

If we want to make a correct characterization of the linguistic sign, we must start with an examination of the concept of the sign in intersubjective communication and look for those attributes which are common to this means of communication and those which are not.   The sign in a general, broader sense, may be defined as something perceptible which is capable of indicating, representing, or reporting on something.   In every sign we can therefore distinguish between that which serves as a sign, i.e., the MATERIAL of the sign, and the SENSE or MEANING of the sign.   With regard to the material of the sign, we can divide signs, according to the particular sense by means of which they are perceived, into optical, acoustical, tactile, etc.   The meaning of the sign represents a special relation between the material of the sign and the thing meant.   If the basis of this relation is a real dependency, a natural connection, we can speak of a 'natural sign'.   A natural sign of this type is, for instance, the smoke which indicates the presence of fire.   If, however, the meaning (sense) rests upon an intentional, explicit designation, then we have an artificial, conventional sign. To this group of signs belong most linguistic signs, symbols and signals.

A SYMBOL, in a narrow sense, is a thing or action which is not only a phenomenon in connection with reality, but which, moreover, expresses or represents something else that is hidden.   Thus, for instance, the flag is a symbol of a national or state feeling. In a broader sense, most figurative expressions, allegorical sayings, conventional signs, etc. may be denoted as symbolic.   A SIGNAL, on the other hand, is an optical or acoustic sign, the sense or meaning of which is agreed upon or explicitly ensues from it.   Examples of signals are optical and acoustic signals, used to control traffic, etc.

An attempt to give more precision to the concept of the sign in its use in intersubjective communication has been made by Bohumil Trnka,[23] who excluded from this concept some of those

[23]   B. Trnka, "On the Linguistic Sign and the Multilevel Organization of Language", *TLP*, 1 (1964), 34.

entities which it seems to include.   Trnka uses as his starting point
the well-known medieval definition of the sign as being something
that stands for something else ('Aliquid stat pro aliquo').   As this
definition is too inclusive, Trnka excludes from it everything that
has nothing in common with communication.   First of all, he
eliminates the so-called operative signs applied to constant or
variable values in mathematics and symbolic logic and, secondly,
symptoms like clouds indicating rain.   All these stand for some-
thing different from what they are.   Indeed, they are far from being
linguistic signs.   They are, in fact, cognitive signs.   Still more
different from linguistic signs are entities that recall something
else by comparison or contrast, e.g., a person's portrait.   On the
other hand, symbols which Trnka calls 'low-grade signs' possess
some of the characteristics of more developed types of signs.   Like
words, symbols may, according to Trnka, be homonymous and
synonymous.   Signals are very similar to linguistic signs, but,
according to Trnka, they differ from them by lacking the subject-
predicate relation.

We will limit our discussion of the linguistic sign to words only,
disregarding the fact that word groups, sentences, clauses, and
utterances may also be signs.

To investigate the relation between word and sign we must first
determine whether it is possible to identify the word with the sign.
At first sight it would seem that we can answer this question in
the affirmative, and some definitions of the sign evidently admit it.
According to most definitions, a linguistic sign is every sound
formation (Lautgebilde) possessing meaning (Bedeutung).   How-
ever, there is a difference between the meaning of a word and the
sign in the broader sense.   The sign stands for a certain thing, a
certain concept,[24] whereas the word as a lexical unit frequently has
several different meanings.   When in context, however, the word
is explicitly defined.[25]   Unlike the word as a lexical unit, the sign

[24]   Symbols that may be, as stated above, homonymous and synonymous are
exceptions.
[25]   Cf. Jaroslav Popela, "K teorii jazykového znaku" [On the Theory of
Linguistic Sign], *Bulletin ústavu ruského jazyka a literatury*, VIII (1964), 6-7:
"... in the actual utterance ("parole") of a concrete speaker *individualization*

in the broader sense has the characteristic of unambiguity.   The
sign is absolutely objective, whereas the word is subjective.   We
can say that the word as a lexical unit, when isolated, has in most
cases a POTENTIAL sign character which becomes an ACTUAL sign
character only in context.[26]   However, the isolated word — the
lexical unit — lacks those additional expressive (stylistic and
emotional) features which make their appearance only when the
word is in context.   Unlike words, symbols do not change, even
when they are isolated.

Trnka's term 'low-grade signs' is valuable, as it implies the
GRADUALNESS of the sign.   It is precisely this character of gradual-
ness that, in our opinion, is typical of linguistic sings.   Some words
have a more sign-like character than other words.   Thus, for
instance, the word *carpet* has a more sign-like character than the
word *chair*, while the latter, again, has more of this character than
the word *seat*.   Similarly, the verb *to walk* is more sign-like than
the verb *to go*, etc.

We cannot, however, maintain that EVERY word is a 'potential
sign'.   There are two classes of words that can be explicitly re-
garded as signs in the true sense of the word: proper names and
onomatopoeic words.   A proper name designates a certain indi-
vidual DIRECTLY, not through the class to which it belongs.[27]
Similarly, onomatopoeic words and sometimes some other simple
denominations DIRECTLY suggest their meaning.[28]   Homonymy is
not compatible with sign-character, as homonyms in isolation are
not capable of being semantically differentiated.   However, since

---

of the general sign both as to the acoustic and semantic aspect takes place...
The individualization concerns both the 'concretes' (concrete nouns) and the
'abstracts' (abstract nouns and other word classes which are always
abstract)...."

[26]   Cf. J. Popela, *op. cit.*, 16: "In fact only the actual sign (the sign in "parole")
indicates reality..."

[27]   Cf. B. Trnka, "Problém vlastních jmen" [On the Problem of Proper
Names], *Philologica Pragensia*, VI (1963), 88.   Much has been written about
the problem of proper names, but we should like to point to only one extremely
interesting paper on this problem: Pavel Trost, "Eigenname", *Acta Universi-
tatis Carolinae — Philologica*, 3, 59-61.

[28]   Cf. Jaroslav Popela, *op. cit.*, 6.

homonymy is usually eliminated in context, there is no need to
discuss it in this book. Synonymy, on the other hand, is a pheno-
menon of a highly sign-like character, as one 'signifié' has two
or more 'signifiants'. Of course, we know that complete identity
between synonyms occurs only rarely, if at all, so that we may say
that in most cases we cannot speak of linking several 'signifiants'
to one 'signifié'. It is preferable to speak of an increase of the
signness by means of a finer differentiation of various 'signifiants';
in other words, a greater individualization of the sign takes place
in this case.

We can therefore assume that the more denominations there exist
for certain things in a language, the more sign-like the latter is,
because this implies an increase in the differentiation and, conse-
quently, in the sign's approximation to unambiguity. From this
an important conclusion can be drawn with regard to the sign-like
character of particular languages, viz., languages may have a
greater or smaller signness. For example, let us compare Czech,
German, and English in the sphere of language denominations.
To the Czech word *obchod* there correspond three denominations
in German *(Handel, Geschäft, Handlung)* and four in English
(business, commerce, trade, shop). The situation is, however, more
complicated, since the English words *business, commerce,* and *trade*
have, in addition to the meaning corresponding to the meanings
of the above-mentioned Czech word, other meanings that are not
included in the Czech word or in the German words (the word
business means also *profession, duty, difficult matter, thing that
concerns one,* etc.; the word *commerce* means also *sexual inter-
course, card game*; the word *trade* means also *a kind of wind,* etc.).
For the Czech verb *počítat* there are two verbs in German *(rechnen,
zählen)* and four verbs in English *(calculate, reckon, compute,* and
count). Each of these English verbs is, so far as the sign-like
character is concerned, more specific than the Czech verb or the
two German verbs, but not quite unambiguous, because each has
still other meanings. Similarly, the Czech verb *nést* corresponds
to two verbs in German *(tragen* and *bringen)* and to four verbs in
English *(bear, wear, carry,* and *bring)*. To the Czech noun *duch*

corresponds only one word in German *(Geist)*, but five words in English *(spirit, mind, intellect, genius, ghost)*. Vilém Mathesius[29] sees in this phenomenon a greater atomization of English word stock, but it can also be taken as a tendency of the language towards a greater signness in word stock. In all the above-mentioned cases English appears to be of a more sign-like character in word stock than Czech, so far as the semantic area expressed in the Czech words is concerned. We may assume that this phenomenon of differences in the sign-like character of the word stock is a universal one in the languages of the world. Let us add further examples from two non-European languages. In the American Indian language Hopi there are, for example, denominations for various kinds of leaves; this testifies to the greater sign-like character of this language as composed, e.g., with Czech, which has only one denomination for 'leaf' *(list)*. A very interesting case is provided by the American Indian language Ponca, in which an unusually great sign-like character is connected with the use of articles. Ponca has animate and inanimate articles. Inanimate definite articles have different forms:

(1) *k'e* dor horizontal objects,
(2) *t'e* for standing objects and collective terms,
(3) *ǥ'aⁿ* for rounded objects,
(4) *ge* for dispersed objects.

Animate definite articles have the following forms:

A. Subject:
   1. *ak'á*, singular of animate objects at rest,
   2. *amá*, singular of animate objects in motion; plural.
B. Object:
   1. *t'an*, singular of animate objects at rest,
   2. *ǥiⁿ*. singular of animate objects in motion,
   3. *ma*, plural of animate objects,
   4. *ǥiñk'á*, singular of sitting animate objects,
   5. *ǥañk'á*, plural of sitting animate objects.

[29] Cf. Vilém Mathesius, *Čeština a obecný jazykozpyt* [Czech Language and General Linguistics], (Prague, 1947), 175-184.

We think we can generalize our observations into the form of a rule stating that the more denominations a language has for certain concepts, the more sign-like it is. However, it is not as simple as that. English is a conspicuous example of a language in which in the formation of the lexical stock two tendencies coincide which may be called 'the differentiating tendency' and 'the accumulating tendency'. The differentiating tendency is the tendency towards a finer differentiation, a narrowing-down of the meaning, that is, the tendency towards greater signness. On the other hand, the accumulating tendency is the tendency towards the accumulation of different meanings in one word form and therefore, towards sign vagueness. There is no doubt that a more extensive and detailed comparison of the lexical stock of various languages with regard to their sign-like character would considerably contribute to the comprehension of the semantic relations in the lexical stock of the languages concerned. In this way, we presume, a certain pattern in the lexical stock might be discerned.

Let us now return to the comparison of linguistic signs with signals. B. Trnka is right in asserting that signals are always 'one-level signs'; on the one hand there is the signal and on the other hand the action or state signalled by it, or, in Saussurean terms, there is a direct connection between the *signifiant* and the *signifié* without there being any subject-predicate relation. However, we are not quite sure whether or not this subject-predicate relation is an indispensable attribute of the relation between the *signifiant* and the *signifié*. Syntactical categories are no doubt characteristic of linguistic reality, but there are lots of cases where they are not indispensable for communication. Trnka himself admits that this relation does not occur with some 'verbal utterances', as for instance, with some military commands. And, let us add, there are other quite frequent cases (cf., e.g., inscriptions, titles, headings, headlines, catchwords, firm signs, etc.). We assume therefore that the subject-predicate relation is not a characteristic and indispensable attribute of the linguistic sign. In our opinion, the syntactical subject-predicate relation is not inherent in the linguistic sign but enters it from outside in the process of

the actualization of the potential sign-word in utterances or in context.

V.A. Zvegincev[30] seems to approach nearest to our conception of the sign and of its relation to the word, and he is right to point to the fact that de Saussure's conception of language as a system of signs takes no account of the specificity of the linguistic sign as compared with other signs. Zvegincev assigns the following characteristics to the concept of the 'sign': unproductivity, un-systematicality, autonomy, and unambiguity. In these character-istics, Zvegincev infers, the sign essentially differs from the word. According to him, the word is something more than the sign. We can accept Zvegincev's standpoint for the most part, especially as far as the last-mentioned characteristic of the sign is concerned, viz, its unambiguity. This is, in our opinion, the basic character-istic of the sign. The other characteristics of the sign, as given by Zvegincev, can, however, be adopted with reservations. Thus, though it is certain that signs are predominantly unproductive, it must be admitted that productive signs do occur. Whether we can speak of the systematicality of signs is a problem which is difficult to solve and thus far we have not been able to take an unambiguous stand in relation to it. On the other hand, we think that we can postulate the autonomy of the sign. Zvegincev, however, in his analysis of the concept of the sign and its relation to the word did not go so far as to investigate the sign-like character of the word in different situations and on different levels. The word as a lexical unit is usually only a potential sign, whereas the word in context is an actual sign. Words as lexical units possess a different degree of 'signness'. Words may, of course, be fully sign-like, but such words are relatively infrequent (an example is proper names which directly designate a particular individual). However, for a great majority of words it is their sign-like gradualness that is typical. In context, this sign-like gradualness disappears as words become real signs and are individualized. We admit though, that there are extreme cases where even the lexical word possesses

[30]   V.A. Zvegincev, *Problema znakovosti jazyka* (Moscow, 1956), 6 ff.

a sign-like character, that is to say, it complies with the basic criteria of the sign, unambiguity and autonomy, and in some cases even with the secondary criteria.

V

THE DEFINITION OF THE WORD

1. THE DEFINITION

Having discussed some of the problems connected with the conception of the word as a linguistic unit and having especially devoted our close attention to some criteria of the word, we will try to give a definition of the word which would be broad enough to be applied to the greatest possible number of the world's languages. This definition reads as follows:

> The word is the smallest independent unit of language referring to a certain extralinguistic reality or to a relation of such realities and characterized by certain formal features (acoustic, morphemic) either actually (as an independent component of the context) or potentially (as a unit of the lexical plan).

Let us now examine to which criteria this definition conforms and to what extent it includes the moot points that have so far been the causes for the failure of the definitions of the word. If we write that the word is "the smallest INDEPENDENT unit of the language", then it means that by using the word 'independent' we are differentiating the word from the bound morpheme which is not independent. The free morpheme is, of course, identified with the word. By including in the definition not only the reference to extralinguistic reality (that is, the meaning) but also the relational character of words, we also bring in synsemantics, which are excluded from all definitions based on meaning. To the other part of our definition, in which certain formal morphemic features are indicated, conform even cases where in a context a part of a word unit is separated from the other word units; this is, for instance,

the case for separable prefixes in German verbs, e.g., *aufstehen*, *einladen*, *zurückkommen*, etc.   The case of English verb compounds with adverbs, e.g., *put up*, *take in*, *set on*, etc. is a similar example.

Of the general criteria of the word the only one which, in our opinion, can be accepted is that of the replaceability of the word by another word or words.   The criterion of displaceability is, as we have already mentioned, not valid in all languages to the same extent when one considers closer or looser position of words in syntactical wholes.   Similarly, the criterion of separability by other words is not always reliable; we refer to our earlier discussion of the genitive *s* in English, which, according to Vl. Horejší, is an independent word, because it is possible to separate it from its base by another word, e.g. *the king's* as against *the king Arthur's*. Of course, if *s* has to be regarded as a word unit, it must be an INDEPENDENT unit.   However, this *s* behaves like any other suffix and the difference between *the king's* and *the king Arthur's* consists in the fact that the *s* in the latter example belongs to the whole complex, *the king Arthur*.   This is evidence of the fact that a grammatical morpheme, in this case the genitive *s*, can belong to a whole group of words.   It would be a word if it could be isolated. It is similar to the case in the word *forget-me-nots*, where *s* also belongs to the whole complex.   Consequently, the complex *the king Arthur's* may be regarded as a compound noun purely on the basis of the connection of all its components by means of the genitive *s*.

The German *am*, *zur*, *fürs*, etc., each of which Horejší regards as two words, are, in accordance with our definition, single words just like the Italian and Spanish *del*, *al*, Portuguese *do*, *ao*, and Czech *bych*, *abych*, *abys*, *kdybych*, *kdybys*, etc.

We also refer to Czech word connections of the type *na jaře*, *na podzim*, each of which, in Horejší's view, should be regarded as one word, because the preposition *na* and the noun cannot allegedly be separated by another word.   A.I. Smirnickij (in an article cited in Chapter Three, Footnote 31) shows a similar case from Russian: *na-sud*.   He argues as follows: if *na* is a part of a word, *sud* is also a part of a word, and, *vice versa*, if *sud* is a

word, *na* is also a word (cf. p. 193). We should prefer to say that it is the cohesion of both words which is decisive here, as in cases like *nakonec* 'in the end', *naoko* 'for make-believe', etc. Moreover, we can apply here the criterion of replaceability (often with a morphological change in the basic word), e.g., *na jaře* 'in spring' — *do jara* 'till spring', *nakonec* 'in the end' — *kukonci* 'towards the end', *napříště* 'for the next time' — *do příště* 'till next time', *nadnes* 'for today' — *dodnes* 'up to now, till the present day', etc. We see that the criterion of replaceability may be asserted even when the criterion of separability cannot be used. Of course, the question remains whether adverbs like *nakonec*, *nadnes*, etc., are, in the last analysis, no more than compounds the meaning of which does not equal the sum of the meanings of both parts. A similar case is that of the adverbs *přesto* 'notwithstanding' and *zatím* 'meanwhile'; they must be regarded as single words and not as groups of words, even if we can apply here the criterion of separability (*přes všechno to* 'in spite of all that', *za vším tím* 'behind all that'). *Přes to* 'over that' and *za tím* 'behind that' are something else when written separately. In one way they have a different meaning from that of the adverbs *přesto* and *zatím* and, in another way, each component of this connection is the bearer of a certain morphological function (*přes* has the function of a preposition and *to* has the function of a demonstrative pronoun; the same can be said of the expression *za tím*), whereas in the case of the adverbs *přesto* and *zatím* it is always the whole which is the bearer of a morphological function.[1] In the chapter dealing

---

[1] A similar problem is discussed by Karel Svoboda in his paper "Pojem slova" [The Concept of the Word], *Listy filologické*, 67 (1940), 173-178, where he writes: "Other groups of words, though close, such as prepositional cases, reflexive pronouns and periphrastic verbal forms cannot be regarded as single words as their components still have their meaning (e.g. *do vsi* 'to the village'), another word may be inserted between them (*do naší vsi* 'to our village') or may be displaced (e.g. *ptáte se, co se ptáte?* — the *se* is the expression of the reflexive character of the verb *ptát se* 'to ask'). Sometimes, however, doubts may arise as to whether such a group should not be regarded as a single word, because *bát* isolated from the expression *bát se* 'to be afraid' has no meaning. Similar doubts may also arise with some word groups, e.g. whether *při tom*

with the word in the written norm of the language we have pointed to this double spelling in Czech.   This differentiation of the double function by means of different spellings may be regarded as a welcome step towards a greater correspondence between the conceptual content and the graphic norm of the language.

Finally, a few words on compounds.   It is not necessary to deal here with the problem of compounds in detail as, on the whole, we can, in accordance with our definition of the word, regard as a safe criterion of the compound as a single word the fact that the semantic cores of the compound cannot be separated by suffixes or prefixes.   Thus, for instance, the plural ending is added to the whole word, e.g., *forget-me-nots*.   However, it may be objected, how do we evaluate the compound *brothers-in-law*?   It seems that in this case we must resort to the semantic criterion, even if this criterion is not quite satisfactory, as shown by B. Trnka,[2] when he points to groups of words like *I have come, I had come, it will rain, the best men*, etc.   Here we must, of course, distinguish between the formal and the semantic aspects.   With regard to the formal aspect, we can, in agreement with B. Trnka, distinguish between

---

'by, at, with, it' is one or two words.   The existence of such doubts is not surprising; in the development of language some kinds of words merge into looser or closer wholes and it is not always possible to determine when the whole is already close (a new word).''   This quotation is in agreement with our assertion that meaning can hardly be regarded as the only criterion of the word, if we can admit it as a criterion at all.   If we use the semantic view as our starting point, then *bát* by itself has, of course, no meaning.   It is not until the reflexive expression *se* is added that it obtains its full meaning.   However, it is interesting that even such a word may become meaningful; in colloquial speech we very often hear *neboj* instead of *neboj se* 'do not be afraid'.   This tendency is possible in the case of verbs if the opposition between a reflexive and a non-reflexive verb does not exist; cf., for instance, *bát se* 'to be afraid', *stydět se* 'to be ashamed', *smát se* 'to laught', *divit se* 'to wonder', etc.   The form without *se* will not acquire reflexive meaning in the case of pairs of verbs, one having the reflexive *se* and one not; cf., for instance, *zavřít — zavřít se* 'to close', *hledat — hledat se (s čím)* 'to look for', *šít — šít se (s čím)* 'to sew (something)', *utnout — utnout se* 'to cut off', *chválit — chválit se* 'to praise', etc. Of course, we can also apply the criterion of displaceability according to which we can safely ascertain that cases like *bát se* must be regarded as two words.
[2]   Cf. *op. cit.* in fn 8 on p. 49, p. 176.

compounds, collocations and groupings.[3] All these formations
consist of independent words but, semantically, they are units the
meaning of which is not a sum of the meanings of their components;
this is why we regard compounds as single independent words.

2. THE PSYCHOLOGICAL CRITERION

Apart from the formal criteria of the word, there are some cases
in which we can apply only the psychological criterion.   This
means that a certain sound complex is felt either as one or as two
or more words.   This is the case for English words such as *fathers-
in-law*, *commanders-in-chief*, *meeting-place*, *apple-tart*, etc. Ac-
cording to the formal criteria, these compounds should be regarded
as being composed of two or more words as it is possible to insert
between their components other words or grammatical morphemes
and the components of the words can be replaced by other words.
In this case, however, the psychological criterion will be decisive,
the formal criteria being subordinated.   The conclusive factor will
be the way in which the speaker and hearer apprehend the word,
viz., as a compound word or as a complex consisting of two or
more words.   Of course, the influence of the written language is
also a factor which cannot be underestimated here.
     The psychological criterion is dealt with by Sapir,[4] according to
whom linguistic experience shows that there is not the least difficulty
in apprehending a word as a psychological reality.   Thus, an
American Indian quite unaccustomed to written words can without
any difficulty dictate a text word by word.   When asked, he can

---

[3] Cf. B. Trnka, *ibid.*, 175-179. Cf. also Jiří Nosek, *Základy mluvnice
moderní angličtiny* [Fundamentals of Modern English Grammar], (Prague, 1960,
mimeographed), 67-68. Compounds are lexical units composed of two or
more semantic cores which are expressed, in the language system, by inde-
pendent words the components of which need not be separated. Collocations
are groups of independent words which may be separated from each other by
other elements (for instance, *I shall come*, *I shall not come*). Word groups are
complexes the meaning of which is different from that expected from the
meaning of their members.
[4] *Language*, 34.

isolate words or repeat them as units. On the other hand, he rejects the isolation of roots or grammatical elements on the grounds that "it has no sense". Our experiences with naive speakers convince us, remarks Sapir, far better than any theoretical argument, that the word has a certain plastic unity.

Our analysis of the problem of the word as a linguistic unit has convinced us that although it is necessary to take into account both the psychological and the formal criteria, they must nevertheless not be mixed. The word as seen conceptually is something other than the word as seen formally. However, these two viewpoints can be united in our minds and each can assert itself equally or one can prevail over the other.

The psychological view begins predominantly as an individual one, as the evaluation of a certain speech complex differs from individual to individual. But as soon as a certain evaluation gradually begins to assert itself in the minds of all or most members of the language community, it becomes a norm.

3. THE GRAMMATICAL FORM OF THE WORD

There have been doubts as to whether words as components of the lexical fund of a language and as bearers of lexical meaning have their own grammatical form. This question has been answered by A.V. Isačenko[5] in the negative. His view point has been criticized and it has been argued that it would be dangerous to create such amorphous elements in language.[6] We are of the opinion that Isačenko's view cannot be completely rejected. Its deficiency lies in the fact that it generalizes about something that is not valid for all languages. But there are indeed languages in which the word as the bearer of lexical meaning does not have its own grammatical form. This is to a considerable extent the case with English, which very language Isačenko possibly had in mind, as may be concluded from examples quoted by him. The

---

[5] Cf. *op. cit.* in fn. 9 on p. 15.
[6] Cf. *O vědeckém poznání soudobých jazyků* (Prague, 1958), 117.

same may also be said of a number of other languages, e.g., of Chinese and the Polynesian and Melanesian languages.   On the other hand, Isačenko's view is not valid for words in flexive languages such as Czech.   We can see that from language to language there are great variations in the determinateness of the word and that these variations are of a typological character.   Words as lexical units are linked to the grammatical form in different ways. There is a whole range of possibilities, from complete amorphousness (e.g., the English *back*) through determinateness of the word class alone (e.g., the German word *Steuer*: we know that it is a noun, but we do not know whether it is a noun of feminine gender ('tax') or neuter gender ('helm') or whether it is singular or plural), determinateness of word class and gender (e.g., the Italian *città*), determinateness of word class, gender and number (e.g., the Italian *case* is a feminine plural noun), determinateness of word class, number and case (e.g., the Czech *hladovec* 'starveling' is a singular nominative noun) to completest form of determinateness, i.e., the determinateness of word class, gender, number and case (e.g., the Czech *učitelka*).   Isačenko is, however, right when he assumes that the word participates in two areas or plans of language: in designation, where its possible grammatical force is irrelevant, and in what is called syntax, where, on the contrary, all its grammatical possibilities are mobilized.[7]   In the same paper (p. 90) Isačenko mentions the slight grammatical force of the word as a word class in English (quoting as an example the word *back*) and asks whether a word which is not classed in a word class possesses any meaning at all.   If it does, it is only a very diffuse meaning.   However, the word has an unambiguous function in the context.   In our opinion, these doubts are quite legitimate.   They only confirm that meaning cannot be the basic criterion of the word.   If we want to preserve the unity of the word we must interpret it primarily from the formal point of view, since from the semantic point of view the function of the word is different in context and out of context.   Out of context the word is, in an extreme case, an amorphous unit and

[7]   Cf. *op. cit.* in fn. 15 on p. 58, pp. 89-90.

sometimes considerably polysemantic; in context it becomes con-
cretized in two directions: it acquires a precise meaning and ranks
with word classes.

4. THE PROBLEM OF THE SO-CALLED WORDLESS LANGUAGES

Our definition of the word presupposes the existence of the word
as a universal fact.   However, there are some scholars who do not
regard the existence of the word as universal in languages and
quote languages (especially some American Indian ones) which do
not seem to distinguish words.
    We shall limit our discussion to one of the most relevant works
which denies the existence of the word in some American lan-
guages.   It is the valuable study by Tadeusz Milewski,[8] who, using
as his basis the definition of the word as "a syntactical group
formed according to the pattern of all the syntactical groups of the
given language", asserts that some American languages (e.g., the
Na-dene, Tlingit, Haida group, the group of Athapaskan languages,
the languages of Maidu and Takelma of the Penutian family, Hopi
of the Uyo-Aztec group and a large group of Sioux dialects) do not
distinguish words.   Milewski argues that in all the languages of the
world there are three kinds of morphological units: morphemes,
syntactical groups and sentences.   Further, there are, according to
Milewski, signals which delimit syntactical groups, that is, they
allow the hearer to segment the sentence into parts.   However,
the difference between languages lies in the fact that in some
languages the means of differentiation of syntactical groups is
common to all types of such terms, whereas in other languages
each syntactical group has its own definite and special delimitative
signals, those relating to the predicate being different from those
relating to the remaining parts of the sentence.   These are, accord-
ing to Milewski, languages without words, as the word is defined
as a syntactical group which is differentiated inside the sentence

[8]   Tadeusz Milewski, "The Conception of the Word in the Languages of
North American Natives", *Lingua Posnaniensis*, III (1951), 248-268.

by means of the same delimitative signals as those used for differ-
entiating other syntactical groups of the sentence.   Consequently,
the language of words is, according to Milewski, a language in
which syntactical groups possess certain constant characteristics
independent of the function of a given group of morphemes in the
sentence.   In all the above-mentioned American languages, quoted
as examples, the predicative group is formed quite differently from
the remaining syntactical groups, the latter usually consisting of
one or two morphemes in a strictly determined order, as for
instance, in Hupa: (1) adverbial prefix, (2) second adverbial prefix,
(3) deictic prefix, that is the pronominal subject in the third person,
(4) first modal prefix (including the pronominal object), (5) second
modal prefix, (6) pronominal prefix, that is the pronominal subject
if the first or second person is involved, (7) third modal prefix,
(8) root, (9) suffixes.   In each of these nine positions only certain
groups of morphemes may occur and this allows us to find the
beginning and the end of the predicative group.   If Milewski's
starting point is the syntactical function of that complex which
we call the word, then he is right to speak of wordless languages,
on this syntactical level at least.   However, the possibility of
delimiting the beginning and the end of the so-called predicative
group enables us to separate this complex of morphemes from the
sentence, which is already tantamount to a criterion of the word
as it points to the independence of this complex.   In our definition
of the word there is nothing that would inhibit the classification of
such complexes with words.   Complexity cannot be an obstacle to
the identification of the word, anymore than semantic complexity
can be when a certain whole has the meaning of a part or of a
whole sentence.   We must, after all, point out that language units
in various languages have different boundaries that often overlap
the boundaries of other units.   Thus the boundaries between
different units often disappear.   We know that a phoneme may
also be a morpheme, a word or a sentence, a morpheme may be
a word or a sentence and a word may be a sentence.   These are,
however, extreme cases.   In the so-called wordless languages it is,
of course, possible, from the syntactical standpoint, to distinguish

on the one hand a predicative group and on the other hand the other syntactical groups which are formed in a way other than that applicable to the predicative group. Is there, however, apart from the difference in complexity, any fundamental difference that would preclude us from regarding both kinds of syntactical groups as certain units that have something in common, not only at the narrow one-language level (which, in this case, would be a syntactical level), but also at the higher level of the language system in general? We are of the opinion that we must answer this question in the negative. After all, nothing prevents us, in the case of languages in which it is possible to make a distinction between a predicative group and the other syntactical groups, from completing our universal definition of the word by adding a specific syntactical criterion in the sense that, in the above-mentioned groups of American languages, there are, at the syntactical level, word wholes of a double kind: predicative and non-predicative.

We do not regard it as possible that a language could exist which does not distinguish words, for the very simple reason that we have not so far met with a language in which it would be impossible to give in isolation, for instance, the names of objects. There is no language which lacks lexical words. The word is, however, a lexical-grammatical unit and it is in this very grammatical function that it may be constituted differently in different languages. Therefore, in addition to the universal criteria that are implicitly or explicitly expressed in our definition, use may be made in different languages of various accessory criteria that will complete our universal definition of the word and make it more precise. One criterion of this type is, for instance, the vowel harmony in Turkic languages (as long as words of other than a Turkic origin are not involved). C.E. Bazell[9] shows the specificity of these criteria by comparing the interrogative particle *li* in Russian with the interrogative particle *mi* in Turkish. The general syntactical rules for the interrogative particle are very similar in both languages: no

[9]   C.E. Bazell, "On the Historical Sources of Some Structural Units", *Miscelánea homenaje a André Martinet, Estructuralismo e historia*, I (La Laguna, 1957), 27-28.

particle in either language occurs after a pause, both languages allow the relatively loose insertion of morphemes between these particles and the preceding morpheme (in most contexts), and no morpheme may be substituted for them.   The only argument for regarding the Turkish particle *mi* as part of a word is in terms of the vowel harmony (cf. *gelir-mi, olur-mu, öldür-ür-mü, alır-mı*). And this argument is a sufficiently valid one.

VI

CONCLUSION

The aim of our work has been a dual one: first we have wanted to verify the validity of various criteria of the word unit in different languages and, second, we have wanted, on the basis of that verification, to give a universal definition of the word. In verifying the validity of various criteria of the word, we have ascertained limitations to their application in different languages. This has induced us to give a very general definition of the word, as it appears that, though we cannot accept the theory of wordless languages, nevertheless there are great differences with regard to the place of the word in the systems of many of the languages of the world, ranging from languages displaying the undisputed presence of word units which are easily separable from the speech flow to languages in which the word units are vague and can be separated only with difficulty, their independence disappearing at the syntactical level. We are confronted here with linguistic concepts the importance of which has been stressed in an increasing number of papers recently: the concept of vagueness and the concept of the central and peripheral systems of language. We will not dwell upon these concepts but simply refer the reader to the second volume of *Travaux Linguistiques de Prague* (Prague 1966) which is devoted entirely to the problem of the center and the periphery of the language system. Let us limit ourselves to a quotation from František Daneš's paper:[1]

The situations in which our analytic practice meets the difficulty or even impossibility to determine unambiguously the place of the given item in the system, are very often connected with classing linguistic units with different categories. Mostly, the difficulties lie in the assumption that each element contains all properties of the given category and that

[1] František Daneš, "The Relation of Centre and Periphery as a Language Universal", *TLP*, 2 (1966), 10-12.

it contains them to the full extent (in other words, that the categories are clear-cut and unambiguously delimited). Thus, in German two opposite word-formative categories may be distinguished, viz. the derived words and the compounds. But in classing, e.g., the noun *das Schulwesen* or the adjective *kugelförmig* we are at a loss if we are to decide with which of the two categories these words should be classed... Similar situations occur in determining the word (as a systemic unit). Thus, in Czech there exists the form *se*, roughly corresponding to the Russian verbal morpheme *-sja*, but distinguished from it by two formal properties: it is spelled as a separate word and is "a free form" whose position in the sentence is governed by rules of word order (*se*, being an enclitic, is placed after the first word or word group of the given clause, irrespective of the position of the verb). Should it be treated as a "free" morpheme or as a "purely grammatical (non-lexical)" word?... All such insoluble dilemmas and useless disputes can be clearly removed if one gives up the common notion of strict compartmentalization and simply admits an obvious idea, viz. that the organization of linguistic elements in the "system of systems" has a different nature, a different structure, than is usually assumed: the classes (and sub-classes) of elements should not be regarded as "boxes" with clear-cut boundaries but as formations with a compact core (centre) and with a gradual transition into a diffuse periphery which, again, gradually passes (infiltrates) into the peripheral domain of the next category... The system of language might thus be presented as space with an uneven density of elements structured according to the principle of "Centre — Periphery — Transition" or as masses of centres with their "fields of gravitation" (of various extent and power) which are in contact, mutually infiltrate one another and are hierarchically ordered...

This quotation evidently confirms our view of the problem of the word unit as we have treated it in this book. At the same time, our problem is that of language universals. There can be no doubt that the linguistic concept of the word unit is such a language universal but the position of the word within the systems of various languages differs from language to language. In some languages the word unit is in the center of the language system, in other languages the word is more or less distant from it and in some languages it is on its periphery. Similarly, the vagueness of the word unit increases from the center to the periphery of the language system.

BIBLIOGRAPHY

Akhmanova, O.S., *Očerki po obščej i russkoj leksikologii* (Moscow, 1957), 67 ff.

Artymovič, A., "Fremdwort und Schrift", Charisteria Gu. Mathesio ... oblata (Prague, 1932), 114 ff.

Bally, Ch., *Linguistique générale et linguistique française* (Paris, 1932), 288 ff.

Bazell, C.E., "Historical Sources of Structural Units", *Miscelánea homenaje a André Martinet: Estructuralismo e historia*, I (Editada por Diego Catalán, Tenerife, Canarias, Universidad de la Laguna, 1957), 19-29.

Benveniste, E., "Les niveaux de l'analyse linguistique", *Proceedings of the 9th International Congress of Linguists* (The Hague, 1964), 266-293.

Bloch, B. and G.L. Trager, *Outline of Linguistic Analysis* (Baltimore, 1942), 54.

Bloomfield, L., *Language* (New York, 1933), 179.

Collinder, B., "Das Wort als phonetische Einheit", *Språkvetenskapliga sällskapets i Uppsala Förhandlingar, Jan. 1937-Dec. 1939* (Uppsala, 1939), 63-75.

Daneš, F., "The Relation of Centre and Periphery as a Language Universal", *TLP*, 2 (1966), 10-12.

Dixon, M.W., *What Is Language? A New Approach to Linguistic Description* (London, 1965), 163 ff.

Ducháček, O., *O vzájemném vlivu tvaru a významu slov* [On Mutual Influence of the Form and Meaning of Words] (Prague, 1953).

Ebeling, C., *Linguistic Units* (The Hague, 1960).

Falkovitsh, E., *Gramatik far dervaksene, I* (Moscow, 1930).

Filipec, J., *Česká synonyma z hlediska stylistiky a lexikologie* [Czech Synonyms from the Standpoint of Stylistics and Lexicology], with a German summary (Prague, 1961), 183.

Firth, J.R., "The Technique of Semantics", *Transactions of the Philological Society* (1935), 53.

Fried, V., *Anglická interpunkce* [English Punctuation] (Prague, 1954), mimeographed.

Gardiner, A.H., "The Definition of the Word and the Sentence", *The British Journal of Psychology*, general section, XII (1921-22), 352-361.

Gardiner, A.H., *The Theory of Speech and Language* (Oxford, 1932).

Garvin, P.L., *On Linguistic Method* (The Hague, 1964), 32.

Graff, W.L., "The Word and the Sentence", *Language*, 5 (1929), 163-188.

Groot, A.W. de, *Structurele Syntaxis* (Den Haag, 1949), 21-31.

Haas, W., "On Defining Linguistic Units", *Transactions of the Philological Society*, 1954, 60.

Hammarstrøm, G., *Linguistische Einheiten im Rahmen der modernen Sprach-wissenschaft* (Berlin, 1966).

Hansen, A., "On the Preservation of the Word-Identity, *TCLC*, I (1944), 48-65.

Harris, Z.S., *Methods in Structural Linguistics* (Chicago, 1951), Chapter 17, § 3.

Hausenblas, K., "Über die Bedeutung sprachlicher Einheiten und Texte", *TLP*, 2 (1966), 61.

Hiorth, F., "On Defining 'Word' ", *Studia linguistica* (1958), 1, 1-26.

Hockett, Ch. F., "Linguistic elements and their relations", *Language*, 37, No. 1, 29-82.

Hořejší, Vl., "Postavení morfologie v mluvnici a její obsah" [The Position of Morphology in Grammar and Its Content], *Časopis pro moderní filologii*, 39 (1957), 75 ff.

Isačenko, A., "Slovo a věta" [The Word and the Sentence], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 89.

Jakobson, R., *Základy českého verše* [Fundamentals of Czech Verse] (Prague, 1926).

Kaiser, K., *Mundarten und Schriftsprache* (Leipzig, 1930).

Kelly, J., *A Practical Grammar of the Ancient Gaelic, or Language of the Isle of Man, Usually Called Manx* (Douglas, Isle of Man, 1870).

Kruisinga, E., *Handbook of Present-Day English*, II, 3 (Groningen, 1932), 4.

Laziczius, J., "La définition du mot", *Cahiers Ferd. de Saussure*, 5 (1945), 32-37.

Marty, A., *Satz und Wort* (Reichenberg, 1925).

Mathesius, V., *Čeština a obecný jazykozpyt* [Czech Language and General Linguistics] (Prague, 1947), 175-184.

Mathesius, V., "On the Potentiality of the Phenomena of Language", *A Prague School Reader in Linguistics* (Indiana University, Bloomington, 1964), 1-32.

Mikuś, F., "En marge du Sixième Congrès International des Linguistes", *Miscelánea homenaje a André Martinet*, I (Tenerife, 1957).

Milewski, T., "The Conception of the Word in the Languages of North American Natives", *Lingua Posnaniensis*, III (1951), 248-268.

Naes, O., "Versuch einer allgemeinen Syntax der Aussagen", *Norsk Tidsskrift for Sprogvidenskap*, XI (Oslo, 1939), 157.

Neustupný, J.V., "Accent in Japanese and Russian", *Ar. Or.*, 27 (1959), 131.

Nosek, J., *Základy mluvnice moderní angličtiny* [Fundamentals of Modern English Grammar] (Prague, 1960), mimeographed, 67-68.

Otto, E., *Sprache und Sprachbedeutung*, Abhandlungen der deutschen Akademie der Wissenschaften in Prag, Philosophisch-historische Klasse, 7 (Prague, 1943), 8.

Poldauf, I., "Tvoření slov" [On the Formation of Words], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 146.

Polivanov, E., "Zur Frage der Betonungsfunktionen", *TCLP*, 6 (1936), 79.

Popela, J., "K teorii jazykového znaku" [On the Theory of Linguistic Sign], *Bulletin ústavu ruského jazyka a literatury*, VIII (1964), 6-7.

Reichling, A.J.B.N., *Het Woord* (Nijmegen, 1935).

Rosetti, A., "Sur la définition du 'mot' ", *Acta Linguistica*, IV (1944), 51.

Rosetti, A., *Le mot* (Copenhague-Bucarest, 1947), 38.

Sapir, E., *Language* (New York, 1921), 26.

Sechehaye, A., *Essai sur la structure logique de la phrase* (Paris, 1926), 14.

Seidel, M.E., "Wort und Satz", *Rev. des ét. indoeuropéennes*, IV (1947), 50.

Seiler, Hj., "On Defining the Word", *Proceedings of the Ninth International Congress of Linguists*, Cambridge, Mass., 1962 (The Hague, 1964), 767.

Setälä, É., "Le 'mot' ", *Journal de la société finno-ougrienne*, 43 (1932), 17-34.

Skalička, Vl., "O příčinách jazykových změn" [On Causes of Linguistic Changes], *O vědeckém poznání soudobých jazyků* (Prague, 1958), 72.

Skalička, Vl., "Komplexnost jazykových jednotek" [Complexity of Language Units], *Universitas Carolina, Philologica*, Vol. 3, No. 1, 15-25.

Smirnickij, A.I., "K voprosu o slove" (Problema 'otdel' nosti slova'), *Voprosy teorii i istoriji jazyka v svete trudov J.V. Stalina po jazykoznaniju* (Moscow, 1952), 196.

Svoboda, K., "Pojem slova" [The Concept of the Word] *Listy filologické*, 67 (1940), 173-178.

Togeby, K., "Qu'est-ce qu'un mot?", *TCLC*, V (1949), 105.

Trnka, B., "Peut-on poser une définition universellement valable des domaines respectifs de la syntaxe?", *Actes de sixième congrès international des linguistes*, Paris, 19 au 24 Juillet 1948 (Paris, 1948), 19-30.

Trnka, B., *Rozbor nynější spisovné angličtiny, II: Morfologie slovních druhů (částí řeči) a tvoření slov* [Analysis of Present-Day Literary English, II] (Prague, 1954), mimeographed, 5 ff.

Trnka, B., "On the Linguistic Sign and the Multilevel Organization of Language", *TLP*, 1 (1964), 34.

Trnka, B., "Problém vlastních jmen" [On the Problem of Proper Names], *Philologica Pragensia*, VI (1963), 88.

Trost, P., "Eigenname", *Acta Universitatis Carolinae-Philologica*, 3, 59-61.

Vachek, J., *Jazykovědný pohled na sovětské zkoušky srozumitelnosti telefonního přenosu* [Linguistic View of the Soviet Tests of Intelligibility of the Telephone Transmission] (Prague, 1957), mimeographed, 22.

Vachek, J., "České předložky a struktura češtiny" [Czech Prepositions and the Structure of Czech], *Naše řeč*, XIX, Nos 6-10, reprint, p. 5.

Vachek, J., "Český pravopis a struktura češtiny" [Spelling and Structure of Czech], *Listy filologické* (1933), 287 ff.

Vachek, J., "Zum Problem der geschriebenen Sprache", *TCLP*, 8 (Prague, 1939), 94-104.

Vachek, J., "Psaný jazyk a český pravopis" [Written Language and Czech Spelling], *Čtení o jazyce a poesii*, I (1942), 231-306.

Vinogradov, V.V., "O formach slova", *Izvestija akademii nauk SSSR, Otdelenie literatury i jazyka*, 3 (1944), 31-44.

Weinreich, U., "Stress and Word Structure in Yiddish", *The Field of Yiddish*, I (New York, 1954).

Zaretski, A., *Yidishe Gramatik* (Vilna, 1929), 13.

Zvegincev, V.A., *Problema znakovosti jazyka* (Moscow, 1956), 6 ff.

INDEX OF NAMES

JANUA LINGUARUM

STUDIA MEMORIAE NICOLAI VAN WIJK DEDICATA

Edited by C. H. van Schooneveld

SERIES MINOR

MOUTON · PUBLISHERS · THE HAGUE