4. IN DEFENCE OF WP

I. In an important article, 'Two models of grammatical description', C. F. Hockett sets out and contrasts grammatical statement made in terms of Item and Arrangement (IA) and that made in terms of Item and Process (IP), with examples drawn from grammars already published as well as examples specifically constructed for the purpose. These two models are selected as 'ideal' or 'cardinal' types of statement, to one or the other of which actual descriptions of languages recently published in the United States conform in varying degrees. In both types the minimal grammatical item is the morpheme, and this is the fundamental unit for all subsequent grammatical analysis.

At the beginning of his article Hockett remarks, however, that if one surveys the whole field of linguistic description the two models he discusses are not the only models possible or actually made use of. A third type or model of description is admitted to have had a long period of employment and a very respectable antiquity, that of Word and Paradigm (WP), and is familiar, in varying approximations, in the school book presentations of Latin, Greek, and Sanskrit, as well as of several modern European languages.

Hockett mentions WP only to refer it, almost apologetically, to a possible separate consideration at a later date, remarking that the bulk of contemporary American grammatical descriptions can be brought under the IA or IP models. He admits further that neither IA nor IP is as yet fully satisfactory, and that WP deserves the same consideration as the other two.

It is the feeling of the present writer that this reconsideration of WP as a valid model of grammatical description is long overdue, since neither Hockett nor anyone else seems as

yet to have taken up the suggestion made in the article just mentioned. It may well be that it offers the means of linguistic descriptions, at any rate of some languages, that are by various criteria preferable to those available by either IA or IP methods, and that linguistics would be impoverished if analysts were exclusively or even principally to subject their material to treatment by one of these two models.\(^1\)

II. WP as a model for contemporary grammatical description may be said to have acquired certain advantages and disadvantages from its very antiquity and traditional status in the presentation of familiar languages.

On the one hand, linguistics is more conscious, perhaps, than some other sciences, of the extensive and almost violent changes in methods and criteria that have characterized and made possible its progress in this century. This is nowhere more true than in America, where explicit rejection of non-formal criteria in grammatical analysis is preached (and on the whole practised) by almost every linguist and in almost every linguistics textbook. The main objection to the older grammars and grammatical treatments was their reliance on non-formal criteria, either derived from the misapplication of the categories of other languages (usually English or Latin) or from alleged meanings or conceptual categories, and often on a sloppy mixture of both of these. And since these older grammars were almost all, in so far as they consistently followed any model, of the WP type, WP itself became, though unjustifiably, associated with the inefficiencies and unscientific method that linguists see in the work of their predecessors and now, almost iconoclastically, strive to avoid. Indeed, as Hockett points out,\(^2\) IA is often more favourably rated than IP just because Process is an older term and still smells of tradition as well as of the earlier confusion of synchronic and diachronic analysis.

On the other hand, however, the long maintained traditional employment of WP type grammatical statement in Europe (and WP is the model implicit in the grammars of Dionysius Thrax and of Priscian) must argue for a certain conformity of WP with some of the commonsense intuitive ideas of grammar and grammatical structure of which speakers are themselves aware. In suggesting that this conformity, in so far as it is found, is an argument in favour of WP, we must be quite clear on what it is that is being put forward. No one would seriously support the view to-day that linguistic analysis and statement should be based on categories taken from sentiment linguistique or the native speaker’s intuitions. But it may fairly be claimed that a model in terms of which grammatical analysis can be firmly established on formal and overt criteria is in one respect preferable to other models equally admitting formal analysis, if the presentations taken from it conform to some extent to the unformulated intuitions of native speakers. In somewhat different words, intuitions of grammatical structure may be the result of a maximal convergence of various features in the language on particular units and associations of units, which are therefore likely to yield a strong foundation for the systematic organization of the grammatical statement.

III. In one sense WP is not a model in opposition to IA and IP in the same way that these two are in opposition to each other. It is claimed for IA and IP that one can account for all grammatical forms and constructions, from single morpheme to complete sentence, as either the result of the arrangement of items or of processes applied to items; nobody would suggest that word and paradigm can likewise cover the whole field of grammar. The salient difference between WP and both the other two models is the centrality it accords to the word as a fundamental unit in the grammar as a whole and as the basic unit of syntactic structure. IA and IP both start from the morpheme as the minimal grammatical element and also the basic syntactic unit, passing through the

---

\(^1\) It must be remembered throughout that the P of IP stands for Process, and the P of WP stands for Paradigm. If this is kept in mind the slight ambivalence of symbols will not cause trouble, and has the advantage of brevity and also of conformity with Hockett’s usage (now adopted by some other writers).

\(^2\) op. cit. 213–4.
word as relatively unimportant, and consequently regarding the traditional division between morphology and syntax as unnecessary or even misleading. This division, with morphology covering the formation of words and syntax the structure of sentences stated principally in terms of the relations between words and word groups, remains central to WP, in contradistinction to both the other models. On the other hand, IA stands at a greater distance from WP than does IP in one respect, in that WP is prepared to make use of 'process' as a term in grammatical description (cp. VI, p. 134).

It was certainly a weakness in the classical grammarians, and in many later writers who followed their example, that they barely recognized any grammatical unit below the level of the word, and certainly never set out with any rigour the establishment of the morphemes of a language. A formalized version of WP must recognize the morpheme as the minimal grammatical (not semantic!) unit of a language, but it makes the word the unit that carries in its paradigmatic and syntagmatic associations the main weight of grammatical description.

To be acceptable any model must be able to identify its units formally and state its operations in terms of which the grammar of a language may be analysed. WP can be formalized just as fully as IA or IP. Its basic units are Words, many of which can be grouped into Paradigms of morphologically associated words, and words enter into constructions, the maximal construction (at present and by tacit agreement) at the grammatical level being the sentence.²

Below the word in terms of size, the morpheme is recognized and established by the now well known techniques (including the grouping together of morphemic elements in complementary distribution, often called 'allomorphs'); but some notoriously 'difficult' cases of morpheme analysis may be simplified by the less fundamental status assigned to this unit in WP (cp. VI, pp. 132–3).

Formal criteria of the word as a unit are by now well known. Bloomfield's use of the 'minimum free form' as a central criterion of the word is familiar,¹ but perhaps exaggerated in view of the contextual unlikelihood of many words appearing as isolated utterances. More significant is the stable internal composition of the word, either absolute or relative, as compared with its relatively free mobility within larger sequences. This general criterion subsumes such other features as uninterruptability by pause or parenthesis in normal discourse, general immobility of the order of its internal morphological constituents or segmental morphemes (any exceptions being far fewer than the general mobility of words themselves in word groups), and limited and regulated extensibility at any point against the almost indefinite extensibility of word groups by the addition of other words.² To the word units of this sort there correspond to a large extent the 'minimum free forms' or smallest potential utterances in a language, and in many, though by no means in all, languages phonologically marked stretches delimited by accentual features of stress, pitch or the like, and sometimes by other prosodic marks.³ Since the word is by definition a grammatical abstraction, any such phonological criteria must remain logically secondary, even though in certain languages marginal cases may be decided one way or the other by reference to phonological criteria where the primary, grammatical, criteria are conflicting or ambiguous.

³ General immobility of internal morphemic order provides the main criterion of word unity in languages like Eskimo, where, as envisaged by J. H. Greenberg, Essays in Linguistics, Chicago, 1957, 32, there is in theory no limit to the extensibility of bases by derivational morphemes.

¹ cp. Z. S. Harris, Methods in Structural Linguistics, Chicago, 1951, 262, 281.
² 'Discourse analysis' very rightly seeks to pursue structural statement beyond the sentence level, but would scarcely be regarded as strictly and exclusively grammatical (cp. Harris, 'Discourse analysis,' Lang. 28 (1952), 1–30).
Words may be divided in most languages into variable words and invariable words, with the concomitant identification of root or base morphemes, forming generally large and open-ended classes, and by which variable words are brought together, and of affixes, forming smaller and closed classes, and by which the different forms of variable words are individually distinguished. The total set of word forms sharing a common root or base are grouped into one or more paradigms, and further paradigms may be similarly set up of corresponding sets of forms built on non-minimal or extended bases or stems.

A cross-classification of less importance to the grammar is that of monomorphemic words (i.e. free morphemes) and polymorphemic words. Monomorphemic words may be members of a paradigm of variables (unless extensive and undesirable use is made of zero morphemes to obviate this: e.g. dog, dog-s), and many polymorphemic words are invariable (not members of a paradigm: e.g. without, midway 1). Polymorphemic words and bases include some constituent morphemes that are irrelevant to the grammatical class of the whole and do not of themselves carry any grammatical category or limitation (e.g. English: receive, recapture, rewear, preelect, precursor). The constituent morphemes may be in part or wholly bound as in the examples just given, or free and linked together by phonological features (e.g. unitary stress) and/or a positional relation not found with the corresponding separate words in the language (e.g. blackberry, a comeback, etc.). To this general type belong the so-called compound words (those written with more than one character) in the Chinese-derived part of Japanese vocabulary. The treatment of such words and bases has little effect on any model, except that in a morpheme based syntax their union will be stated as the lowest layer of immediate constituents,

1 That is to say, paradigms in the usual sense of limited sets of affixes associated with sets (often unlimited) of bases or stems. In a wider, mainly lexical, sense, any polymorphemic word could be said to fall into a paradigm (without, within, without; midway, midstream, midday, etc.).

while a word based syntax will take their composition for granted. In either case their analysis and classification belongs primarily to the lexicon of the language.

IV. As has been said above, WP preserves the division of morphology and syntax more clearly than the two morpheme based models, and attaches more importance to it. Constructions, the province of syntax, of which the maximal construction is the sentence, are principally made up, in WP, of words in the first instance, and the places in a syntactic structure are filled primarily by words and word groups.1 Sentence types and construction types are analysed formally in the now familiar terms: exocentric, endocentric, subordinate, co-ordinate, superordinate, by the application of the usual tests of cohesion and substitutability (distributional parallelism) of part and whole, with the distinctive provision that the cohesion of a construction or group rests on its substitutability for or by a single word (of the same class as one of its components in an endocentric construction, of a different class in an exocentric construction).

It must be emphasized that by the inter-word relations within constructions one understands the interrelations of words not as individual lexical items (which is the province of collocation), but as members of classes that occupy places and fulfill functions in a construction by virtue of their class membership and any inflectional categories they may bear. In complementation to this the inflectional formations and the categories of which they are the exponents are relevant to the grammar of a language by virtue of the syntactic structures into which they enter and the constructions and structural relations within constructions that they mark.

In this respect syntax may be considered the core of grammar; a language may dispense with some morphological complexity, as in the case of modern English as compared with Anglo-Saxon, or some modern Romance languages as compared with Latin, and it may reduce its inflectional apparatus almost or quite to vanishing point, as have some

1 For reservations to this general statement see VII, pp. 140-2.
In every language words as members of word classes cohere in groups with specific distributional and functional relations with each other and with other groups to form sentences of different syntactic types. And the formal description of these constructions and sentence types (whether the starting point is made the word or the morpheme) does not differ markedly in its inventory of basic syntactic terms as between the so-called inflectional languages with complex morphology and languages of the type represented by Chinese.

Hockett suggests in passing that WP is "incapable of organizing efficiently the facts of a language like Chinese". To say that paradigms are not part of a model applied to a language whose words do not exhibit grammatical paradigms is labouring the obvious, but it may still be urged that the word as formally established is the most profitable unit to be taken as basic in the statement of the sentence structures of such languages; and this indeed is the case with two recent treatments of different varieties of Chinese by writers both of whom reject the facile and time-worn equation 'word = morpheme = segment written with a single character', and recognize the word as the unit of sentence composition and as comprising simple (monomorphemic) and compound words. It may be said that it is in its grammatical structure markers rather than in its basic grammatical structures that languages of the Chinese type differ most sharply from those of other types, and that therefore a word based syntactic analysis of such languages is still, though as an extreme case, on WP lines.

V. In terms of WP it has been argued that syntax is the core of grammar, that words are to be treated as the primary units of syntax, and that morphological formations (or more strictly inflectional formations) in languages exhibiting them serve as the markers of syntactic structures and groupings and the relations within and between them. This last proposition brings the paradigm into prominence in grammatical description. The paradigms of traditional presentations, whatever semantic glosses are attempted either for a whole paradigm or for individual members of it, bring together the inflectionally different forms into closed sets each of which marks one or more of the syntactic functions of the word class in the constructions into which it enters; paradigms thus represent interlocking systems of grammatical oppositions within particular syntactic fields. Paradigms are primarily and mainly of single words, but where short groups of words or phrases (e.g. Latin, and some Greek, perfect passives) are syntactically comparable to single words in the corresponding places of a different paradigm they are obviously to be included in paradigms themselves.

From what has been said about the centrality of syntax, it follows that word classes are primarily distinguished by their different syntactic functions; in inflecting languages these may in some (but not all) classes be associated with different sets of morphological paradigms. In these latter cases it matters little whether the class is predicated of the root or stem less inflection or of each of the words brought together in the paradigms.

Examples of the syntactic criteria by which paradigm categories are formally established are obvious and need no more than mention: Number in English, relation between subject (usually preverbal) noun or pronoun and present tense verb; case in Latin, relation of noun or pronoun to preposition (exocentric construction) or verb (endocentric construction); gender in French, relation between noun or pronoun...
and adjective in endocentric ('modifying') construction and exocentric ('predicative') construction, in German only in endocentric adjective + noun constructions.

The marking of a particular set of syntactic relations is the criterion by which a large number of inflectional forms sharing a single base or stem is divided multidimensionally among several crossing paradigms. A familiar example is the organization of the Latin adjectives of each grade (positive, comparative and superlative), which may exhibit up to 15 different forms, into a paradigm of number with two members, a paradigm of gender with three members, and a paradigm of case with six members, each paradigm covering a different set of syntactic relations of the adjective forms with other words in syntactic groups. As a result each form of the word class is uniquely specifiable as the meeting point of three interlocking paradigms.

Certain comments are needed here:

1. Diagrams can only clearly exhibit two dimensions (across and down) at a time. Thus an inflected word involving more than two paradigm categories (or three, if one resorts to perspective representation for the third dimension) must repeat particular paradigms under different cross headings. In the example just given this is done by duplicating the case and gender diagram under both numbers. This representational necessity, which is more noticeable with word classes involving more paradigm categories, for example verbs in many languages, does not affect the logic of paradigmatic organization.

2. Only inflectional paradigms have so far been considered. Derivational formations with a common base form can also be set out in paradigms, though these are mostly less regular and offer more variation as between individual bases in the classes involved. Some paradigms can be said to be on or near the borderline as between inflection and derivation. The category of grade in the Latin adjective is borderline in this respect; almost (though not quite) universal in their applicability to the class of adjectives, the comparative and superlative forms share the majority of the syntactic functions and relations of the positive forms, though a certain few constructions are unique to each. In the example of paradigm treatment given above grade was regarded as derivational and so not included in the (inflectional) paradigms of the adjective; the picture could easily be adjusted in terms of four inflectional paradigms instead of three.

3. To what extent should symmetry determine the assignment of single forms to more than one place in a multidimensional frame of paradigms? This must be decided by the degree to which the multiplication of grammatical homophones is held to be compensated by the greater overall simplicity of a generally valid frame. The general symmetry of singular and plural case paradigms in Latin nouns may be a justification for continuing to recognize separate dative and ablative cases in the plural even though no formal distinction is found in any declension. A like argument can be used to justify distinguishing nominative and accusative cases in neuter nouns and adjectives by reference to the formally separate corresponding masculine and feminine forms. On the other hand it would hardly be reasonable to posit three genders in the plural paradigms of German adjectives and articles, where no formal distinction is found in any of the cases, merely to provide symmetry with the singular paradigms.

4. Irregular, defective and suppletive paradigms, which by definition are always a minority, are established subsequently.

---

Footnotes:
1 The converse, of course, is not true. Some syntactic relations (more in some languages, e.g. English, than others, e.g. Latin or Greek) are marked by word order or by the class of words involved alone, as is wholly the case in languages of the Chinese type, in which particular classes of words (often designated 'particles') operate as construction markers.

2 cp. B. Bloch and G. L. Trager, Outline of Linguistic Analysis, Baltimore, 1942, 56.
(logically and temporarily) to the regular paradigms, principally on the basis of syntactic parallelism with the corresponding members of regular paradigms (e.g. French *aime*, *aimes*, *aime*, *aimons*, *aimiez*, *aiment*, and therefore *vais*, *vas*, *va*, *allons*, *allez*, * vont*).

VI. The main distinctive characteristics of a formalized WP model of grammatical description would then be the following: the word is taken as the basic unit of both syntax and morphology, and variable words are grouped into paradigms for the statement of their morphological forms and the listing of their various syntactic functions. It will now be reasonable to consider some of the advantages that word based WP offers in certain cases over morpheme based IA or IP.

In the first place in a grammar considered as part of a complete structural description, the central position accorded by WP to the word may lead to a greater ‘congruence of levels’ in languages wherein the word is itself the domain of phonological features, prosodic or other, and therefore a focal point for the statement of phonological structuring. This is not stressed here, however, for two reasons: 1. There are several languages in which the word as a grammatical unit does not appear to correspond clearly with any phonologically marked or significant unit, and 2. It is as a model of strictly grammatical description that WP is being examined.

In these terms WP seems to offer advantages in formal description of two principal types: 1. The word as a unity is more easily susceptible to grammatical statement than is the individual bound morpheme, and 2. Certain problems of morphemic composition (morphophonology or (American) morphophonemics) become less tiresome in WP.

These may be considered in turn:

1. As has been said, the morpheme must be recognized as the minimal element of grammatical structure; but this does not imply that it is the most suitable element to bear the assignment of all the grammatical functions fulfilled by the words into whose composition it enters. Some entities that are clearly to be assigned morphemic status may be seen in several languages to bear conflicting and even contradictory grammatical functions when considered in isolation. This is often accompanied by conflicting ‘meanings’ also, but as one must insist that grammatical methods are not determined by semantic considerations, and that the morpheme is not to be defined as a semantic element in the first instance, this consideration is to be ruled out as irrelevant at this point. On the other hand words anchored, as it were, in the paradigms of which they form a part usually bear a consistent, relatively simple and storable grammatical function. The word is a more stable and solid focus of grammatical relations than the component morpheme by itself. Put another way, grammatical statements are abstractions, but they are more profitably abstracted from words as wholes than from individual morphemes.

Examples of this may be considered from a number of languages:

In Sundanese one finds an extreme example of a language whose individual affix morphemes are very few in number, but in which most of them combine with stems and with each other in forming a considerable array of grammatically different word forms, giving a very rich set of possibilities of word structure. Most affixed words are built up from combinations of root (most of which occur also as free forms without affixation) and one or more of the following: -an, -kyn, -yn, ka-, pa-, pay- (may-), pi- (mi-), sa-. Reduplication of the whole root or of its initial syllable is also freely employed; other affixes are rarer both in variety of uses and in frequency of occurrence with roots.1

These affixes are found in both derivational and inflectional formations, the former being the more numerous and diverse,
and paradigms of derivational possibilities for roots of the noun and the verb classes can be set up, though with considerable lexically determined limitation on the use of a particular formation on a particular root. As an example, -an may form denominal nouns (laut, sea; lautan, ocean), denominal verbs (tanduk, horn, tandukan, to be horned), deverbal verbs (asup, to enter; asupan, to put in), and deverbal nouns (asupan, to be complete; asupanen, end); and similarly diversified derivational functions of other affixes are found. Combinations of more than one affix exhibit a separate derivational power from that shown by either used singly. Thus parJ- + verb root yields a deverbal noun (derJe, to hear; parJderJe, what is heard); -kYn suffixed to an intransitive verb root forms the corresponding transitive verb (sare, to sleep; sarekYn, to send to sleep); but parJ- and -kYn together affixed to a verb root produce a 'benefactive' form (bawa, to carry; parJbawakYn, to carry (something) for someone.

One sees that by themselves several of the most commonly used affixes have no set grammatical function or non-contradictory group of functions, but serve to form morphemically complex words possessed of very definite grammatical functions and belonging to definite word classes. It is to be noted that one is not dealing with whole series of grammatical neutrals; though grammatical neutrals (word forms serving both as nouns and verbs) are found fairly freely among roots and occasionally among derivative formations, the great majority of the formations referred to above have a single and perfectly precise grammatical function and this is a property of the word as a whole unity, and not predictable from its component morphemes in isolation (unless an intolerable degree of homophony among the affixes were permitted).

In written Japanese a distinction of categories central to the verbal system is that of Conclusive and Attributive, conclusive forms corresponding somewhat to the indicative tenses of verbs in other languages and mainly occurring in sentence final position, attributive occurring before nouns to form endocentric nominal groups, and also before certain particles and in some other positions in the sentence; thus: conclusive, kono koto o sirubesi, (one) should know this (thing); attributive, sirubesi koto, a thing that (one) should know. This distinction is marked by the final morpheme in each verb form, but among the morphemes most frequently so used -ki and -si are each found as both attributive and conclusive. Thus we find: attributive yukisi, ykitariisi, (which) went, and ykitariisi, (which) did not go; conclusive yukiki, ykitariiki, went, and ykitariiki, did not go; but attributive yubeki, (which) should go, and yubeki, (which) wants to go; conclusive yubesi, should go, and yubesi, wants to go. In the inflected adjective, which in Japanese belongs to the class of verbs both in morphology and syntactic function we find: attributive hayaki, early, hayaki, (which) was early, and conclusive hayasi, is early, hayasi, was early.

None of the words quoted is grammatically ambiguous (though there are some homophonous attributive and conclusive forms, e.g. yuku, goes, which goes); but the grammatical significance of the morphemes -ki and -si is only determined within the complete word in its paradigms of category oppositions.

Similar situations may be seen in English and other European languages where a single morpheme shape, or class of such shapes in complementary distribution, may as an isolate serve quite different grammatical functions, but without grammatical ambiguity in the whole words of which they form part. Thus we have in English -er serving as comparative adjectival suffix and as 'agentive' deverbal noun formative, and -(e)s ([-s], [-z], [-iz]) as third person singular present tense suffix in verbs and as plural suffix in nouns; and in German

\[1\] In the Nipponsiki romanization here used si ([wi]) = shi, ti ([wi]) = sit, o = so of other transcriptions in the examples cited here and elsewhere in this paper. For further details of the grammar see B. H. Chamberlain, A Simplified Grammar of the Japanese Language, revised McIlroy, Chicago, 1924.
the prefix ge- acts both as the formative of collective and general nouns (e.g. Gestein, Gerede) and (often with other inflections) as past particle marker in verbs (gekommen, gemacht).

Latin and ancient Greek are well known for the difficulties they reveal, unlike languages of the type represented by Japanese and Sundanese mentioned above, when one attempts to find a unique segmentability into serial morpheme segments; most morphemically complex words admit a number of morphemic analyses almost equally justifiable by comparison with other forms elsewhere in the language. But in these languages we find morphemic elements which in isolation are grammatically ambiguous or even contradictory but help to form words revealing a clear-cut grammatical function in the set of oppositions maintained by the members of their different paradigms.

In Greek verbs with a sigmatic aorist and future tense -σ- may be abstracted as the tense morpheme. If the bound morphemes are to bear grammatical descriptions individually, -σ- must either be assigned a hazy 'non-present' grammatical significance only clarified in the word, or at the cost of an uneconomic homophony of different morphemes, a distinct future morpheme and aorist morpheme, both with the shape -σ-, must be posited. A similar problem arises with the singular members of the perfect and aorist indicative paradigms of some -σ- verbs (δεσιωκα, etc., εδωκα, etc.; τεθεικα, etc., εθηκα, etc.).

If, however, the word is taken as the basic grammatical unit of description (as in the traditional grammars), then λόγω and διωκα, with the rest of their paradigms, are respectively the bearers of the categories of future and aorist tense, together with the other relevant categories, and likewise


2 M. S. Rupírez, 'Neutralization of morphological oppositions', *Word* 9 (1953), 249, suggests that the 'momentary' feature of the aorist stem as contrasted with the 'durative' feature of the present stem is neutralized in combination with the future category.

8€ωκα and εδωκα, as whole words, bear the categories of perfect and aorist respectively. Their morphemic composition as including -σ- and -κ- morphemes is recognized, but categories need not be assigned at that level to these morphemes, as would be necessary if the morpheme is to be regarded as the fundamental unit of grammatical description.

A number of related problems have been left untreated here, including the possible morphemic identification of the -σ- element with parts of some non-sigmatic (strong) aorists and futures, the morphemic segmentation of the rest of the words cited, and the morphemic status of the augment and reduplication; but these omissions do not affect the significance of the examples in comparing a word based and a morpheme based grammatical analysis.

A similar case of morphemic ambiguity (among others that could be mentioned) arises in the future and imperfect tenses of Latin verbs of the first and second conjugations (*amabo, amabam*, etc.), unless use is made of the morphemic distinction, adopted by A. A. Hill, produced by taking -b- as the future tense morpheme, and -ba- as the imperfect tense morpheme, which runs counter to the entirely parallel syllabic structure of all the word forms concerned.

2 We now turn to the second reason for preferring, in some cases, a WP treatment to an IA or IP one: WP avoids some of the difficulties in morphophonology (morphophonemics), in the relating of grammatical structuring to phonological structuring, which beset IA (and to a lesser extent IP).

In languages of an almost pure agglutinative type, of which Turkish is often taken as an example, it is usually possible both to segment polymorphemic words into serial segmental morphemes without difficulty and to assign each to a par-

1 The names given to these grammatical categories are here used as labels only; the grammatical significance of the categories lies in their 'sequential' and other syntactic relations with the rest of their sentences, rather than in the temporal and aspectival meanings that may be assigned to them, important as these are in the total analysis of the language.

ticular grammatical category as its exponent. Other languages, such as Latin, Greek, and Sanskrit, and to some extent English, freely admit cumulation of categories in a morphemically indivisible segment.

It has already been made clear that a formalized version of WP must admit the morpheme as the minimal unit of grammatical formation, below the word level; but if the word is to be the focus of grammatical statement, the segmentation into morphemes need not take into account any need for a parallel representation of the grammatical categories applicable in every word in any class (though, of course, it may do so if a clear statement on these lines is possible). Whatever the morphemicization of *amabamus* and *bonarum*, the categories of tense, mood, voice, number, and person, and of case, number and gender are exhibited by the verb and noun forms as whole words taken together. Phonological structure does not cut across morphemic analysis in words like English *took* and *goose*, or French *au*. The grammatical categories are borne by the words in their paradigms, and no objection need be taken to paradigms varying in morpheme composition from word to word within a word class. Thus *rang* (the bell) and *ranged* (the bird's leg) are both past tense forms in the paradigms *ring, rings, ringing, rang, rung*, and *rang*, *rings*, *ringing*, *ranged*; and *rang* and *rung*, like *took* and *goose*, are 'perfectly respectable' monomorphemic words, as their whole phonological composition argues.\(^1\) The 'tactical parallelism' between *took* and *baked* that prevents Hockett\(^2\) admitting their morphemic non-parallelism is not binding on WP as in this model it is the word that is the (sy)n-tactical unit; and we are spared the subterfuges of zero suffixes with conditioned allomorphic variation, infixes (a formation otherwise unknown in English), morphemes 'consisting of vowel change', and suffixes that are not suffixed.\(^3\)

If morphemes are the Items that have to be Arranged, then clearly something must be found comparable to the *-ed* of *baked* in *took*; but if *took* is just part of the paradigm of the verb *take*, no such need arises. Among paradigms of wholly polymorphemic words one may plausibly regard Latin first and second conjugation futures (*amabo, monebo*, etc.) as containing one more morpheme than the futures, grammatically equivalent in every way, of the third and fourth conjugations (*regam, audiam*, etc.).

The number of morphemes in words belonging to comparable paradigms may vary, and categories may be serially or cumulatively expressed. In consequence the 'empty morph' as an oddity disappears.\(^1\) If the segmentation requires or is facilitated by a recurrent partial without any grammatical category applicable to it, this is of no consequence; it is part of the word to which it belongs and which, as a whole, exhibits the categories relevant to it as a whole.

Process and Arrangement both have their place in a grammatical description. Process is most naturally applicable to the morphological formation of words, and arrangement to the syntactical relations (not linear order, though this is often, in varying degrees, a mark or exponent of syntactical relations) of words as members of classes bearing specific grammatical categories in constructions.

Process adequately covers the variety of formal associative relations found between root and stem forms (III, p. 121) and members of the paradigms, and provides a single means of describing the formation of these words by reference to a constant or common element (except in suppletive paradigms) underlying them, whether or not the root or stem form itself serves as a paradigm member or not. The processes employed in languages include affixation (*fix-ed, ἐπίτροπον*), alternation of vowel and/or consonant structure (*sing sang, bring*, *brought*), reduplication in whole or part of base form (*λέ-λυ-κα*, Malay *akar-akar*, roots), compounding *black-berry, Japanese san-eu*, landscape (*mountain-water*), stress difference

\(^1\) cp. Bazell, 'The correspondence fallacy in structural linguistics', *Studies by Members of the English Department, University of Istanbul* 3 (1962), 4–5.  
\(^2\) 'Two models', 224.  
\(^3\) Hill, *Introduction*, 140–1.
(export, export) and tone difference (Ibo isi-, head, ok-, rat, isi ok-, rat's head 1).

Subtraction is also a possibility, 2 though the advantages of keeping to the constant (in regular paradigms), and therefore shorter, form as the descriptive base probably outweigh the occasionally briefer statement of formation rules made available by recognizing subtraction.

It may be said that in affixation at least, Arrangement is as appropriate a descriptive term as Process. To this one replies that the intra-word bond is always much closer than the bond between words, in the restriction or complete exclusion of further affixation of bound morphemes, and in the almost complete fixity of internal order, whereas syntactic word groups are often variable in relative position, usually admit the further insertion of other words, and are almost always further extensible in one or both directions by other words. This difference between affixation and word grouping has its counterpart on the phonological level in that junction or juncture features are more frequently admissible (if not always found) between rather than within words. Affixation, even where the fusion of the elements is at a minimum (as in the purest agglutinative type languages) is a much tighter union than syntactic juxtaposition, and this justifies its inclusion in the class of processes describing the morphology of members of paradigms. Conversely, in ‘allegro forms’ where syntactically separate words are phonologically fused (e.g. English [wount] (won't), [ai am ‘get it] (I'm going to get it)), there is almost always the grammatically equivalent (and semantically near equivalent) non-fused sequence available ([wil not], [ai am ‘goin’ ‘to get it]).

In keeping intra-word Process and inter-word Arrangement as separate terms, it would also seem better to restrict ‘exocentric’ and ‘endocentric’ to syntactic (inter-word) structures, ‘derivation’ and ‘inflection’ being more suitable terms with reference to word structure and covering approximately the comparable situations; and the collocation of ‘semantic’ with ‘endocentric’ and ‘exocentric’ in the classification of compound words 1 only brings confusion by reintroducing semantic criteria into what should be formal analysis and classification.

VII. We must now consider the possible objections that may be made, and the disadvantages that may accrue to WP, or to a model of grammatical analysis wherein the word is the unit of morphological description and the basic unit of syntactic analysis.

Objections may be either theoretical, and in fact almost aesthetic, or practical.

Firstly it may be argued that with its use of Process terminology, WP, like IP, implies a historic perspective and confuses synchronic linguistics with diachronic linguistics. This should by now have lost its force. ‘Process’ as used today in any model of descriptive grammar is simply a means of relating formally one word or form to another word or form by the most economical statement, and implies no sort of time dimension. 2

Secondly, and more seriously, WP would appear at first sight less tidy and economical in requiring both Process and Arrangement as separate terms (in morphological and syntactic description respectively) than either IA or IP with their exclusive use of one or the other. The tidiness and economy of these models, however, is more apparent than real when it is seen that Arrangement in IA includes (by

---


3 In this way words like English won't are descriptively contractions in a sense in which French au is not a contraction of a le.
whatever devices, with or without zeroes) internal vowel and consonant alternation and all accentual differences, and that Process in IP is made to cover such patently diverse things as 'past tense formation', 'predication', 'addition' and 'modification', as well as the markers thereof such as 'vowel change' and the like.¹

Thirdly it may be felt that a descriptive model is intrinsically less desirable if it makes a non-minimal element, the word, basic in the hierarchy of structures at the same general level of analysis, as against both the other models that make the minimal grammatical element, the morpheme, also the basic element of structure. The failure of the ancient grammarians of Latin and Greek to recognize the morpheme, and of their traditional followers of more modern times to take proper account of it, shielded them from the need to consider this possible objection to their method and so weakened its appeal.

The answer to this has in effect already been given. In many ways, and quite apart from any phonological markers, the word is a unique entity in grammar, and not just a stage in the progression 'from morpheme to utterance'. As a grammatical element the word is unique in its relative fixity of internal morphemic structure, its focal status in relation to syntactically relevant categories, and, in inflected words, the stability of its paradigms. All of these factors make it a strong basis for grammatical description, both morphological and syntactic. The assumption of a simple ascent in order of size from single morpheme to complete sentence, ignoring or blurring the distinction of morphological structuring and syntactic structuring, achieves its apparent simplicity at the cost of neglecting or distorting patent structural features of languages.

The most serious objection to WP is the practical one, namely that at least in some languages the word is not in fact the best unit for statements of syntactic structure, and that word boundaries repeatedly cut across any reasonable analysis into immediate constituents.

It is agreed that immediate constituent analysis underlies most schemes of syntactic statement, and that the basis of grouping units into constituents is principally replaceability or syntactic equivalence between a group and a single unit.¹ If the morpheme is taken as the single unit, then many examples appear of immediate constituents cutting across word boundaries; for example, in large books, as a typical English adjective noun construction, large book is replaceable by book in almost every case, whereas the word books (or any other pluralized noun) is only replaceable in the important preverbal (subject) position by a single morpheme from the very small class of monomorphemic pluralia tantum like cattle. But WP takes the word as the basic syntactic unit, and the question to be asked is to what extent are word boundaries cut across by immediate constituents when syntactic equivalence to a single word is taken as the main criterion of constituent status. Wells indeed makes it a point of principle that the word should be treated as a unitary constituent in all cases except when this clearly engenders a conflict or an avoidable complication.²

One may compare the possible immediate constituent analyses of the same sentence on the basis of the morpheme as the minimal constituent and on the basis of the word as the minimal constituent. Harris's example 'My most recent plays closed down' offers the following possibilities³:


The line diagramming used here seems the most suitable for present purposes; it can, of course, be easily and uniquely converted into the Quine system of dots as given in Harris's text.
The union of play and -s and of close and -d in the second analysis is accounted for elsewhere in the grammar, in the morphology of the noun and verb paradigms; for the rest the two analyses are equally efficient and the word based one is justified on the same general ground, substitutability of the group with the single unit (in this case the word). My may be treated as a more 'outer' constituent than most recent by virtue of its virtually closing the endocentric nominal construction in one direction; only all or both and some adverbs like only and exclusively can precede the possessive pronouns in a larger endocentric construction with the final noun as head. Harris appears to treat my in this way mainly to preserve a parallelism between order of constituency and order of morphemes in the text; this breaks down in the verb group where he has to admit -d as an infix (a category avoided in English grammar by a WP type analysis). However, in some languages, e.g. Latin, on either method immediate constituent structure and positional order of units would bear scant correspondence.

In many cases the immediate constituent analysis on both a morpheme and on a word basis will be very similar apart from the treatment of intra-word composition elsewhere, with a consequently less complex constituent structure. In some cases however a construction may be exocentric by one analysis and endocentric by the other. A sentence in written Japanese yields the following possible analyses:

\[
\text{teki gun-tai o kakoma-ba ti-ten o sir-as-e}
\]

(enemy troop object if (you) position obj. let (us)

unit particle surround part. know !

(if you surround the enemy troop unit report its position)

The curved lines link lexically united morphemes in compound words (III, p. 121) and need not be further considered here; siras- is a causative stem formed on the verb root sir-, to know. In the upper (morpheme based) analysis teki --- kakoma-ba is syntactically equivalent to a single verb base (mi-, see, yuka-, go, etc.), and the whole clause teki --- kakomaba is replaceable by a monomorphemic adverb (e.g. ima, now) and is consequently exocentric, since neither constituent, verb base and expansion, or hypothetical suffix is substitutable for the whole clause. In the lower (word based) analysis the whole clause is replaceable by any hypothetical (subordinating) verb form (miba, if (you) see, yukaba, if (you) go, etc.), and the clause is therefore endocentric. Within the nominal group of the English example above, most recent play-s is exocentric on a morphemic analysis and endocentric on a word analysis. Probably exocentric constructions will always be more numerous in a morpheme based constituent system than in a word based one.

In general one may claim that WP permits an equally efficient and sometimes simpler immediate constituent analysis in grammatical statement. Constructions that are ambiguous in isolation in another model are also ambiguous in WP; old men and women is resoluble one way or the other either by context, linguistic or non-linguistic, or is made explicit by pause features not ordinarily present: old, men and women; old men, and women.\(^1\)

There remain, however, in several languages some hard case in which immediate constituent analysis must cut across word boundaries. In English there are sentences on the pattern of the by now notorious example the king of England's hat, the man I saw's daughter, commoner in colloquial spoken English than in more formal discourse or in writing. The 's is most suitably treated as a bound form and as part of the word to

\(^1\) cp. Hockett, 'Two models', 219–20. It may be pertinent here to protest at the tendency among certain writers to-day to assert that junctural and other phonological features are always present at certain grammatical boundaries and in particular constructions, and can therefore serve as their criteria. Such features are better treated as merely available and intermittently employed as additional marks of grammatical differences.
which it adheres, as it is neither itself a potential free form nor the syntactic equivalent of one, nor is there available a corresponding form consisting of more than one consonant as there is with the smallest ‘weak forms’ of other English words. But this morpheme, excluded from the category of words, may belong grammatically, as in the examples above, in immediate constituency to a noun disjoined from it in the sentence. It is interesting to see that a corresponding situation is found among the Scandinavian languages in Danish and Norwegian but not in Swedish.

In still more colloquial English one finds almost nonce forms like *extract-of-beefed his bread*, where *beef* could not appear alone as a transitive verb.

Expressions such as *bandy-legged, heavy-fisted*, and the like are rather borderline cases; they are normally stressed as single words and are hardly extensible like word groups. *Dark and shaggy-bearded man* would ordinarily be construed *dark and shaggy-bearded man*, and sequences of this type with the first adjective not collocable with the final noun, though just possible, are unlikely (*sandy- and shaggy-bearded man*).

Enclitics in other languages (e.g. Latin and Greek) present similar difficulties for word based immediate constituent analysis, unless their relative freedom of position (other than in sentence or clause initial position), which is greater than that of English’s, and their word-like segmental structure may be held to justify word status despite their accentual dependence on the immediately preceding word. The inflected Japanese form *-rasi*, looks like (spoken form *-rasii*), is similar in this respect. It is a bound form, cannot occur initially, and is accentually treated as part of the word it immediately follows. In a sentence such as (spoken Japanese) *kodomorasisi*, it looks like a child, *-rasii* coheres grammatically with *kodomo*; in *kodomo norasiisi*, it looks like a child’s, it coheres grammatically with the group *kodomo no* (and *norasiisi* would not be a possible group by itself), but is phonologically bound to *no*. Bloch and Wells, despite their general principle of word based constituents, do not regard *-rasii* as a word, but as a suffix of unique structural patterning.

In Yurok, a native language of North California, a cross cutting of word boundaries by immediate constituents does occur, but in one type of construction and then as a free, but less frequent, variant of sentences without the cross cutting. *ki toktomoyel ki ?umuc megetolk*epel, they are now old enough to look after themselves, is in free variation with *ki toktomoyel ki muc ?umegetolk*epel, where *u*- in both sentences coheres grammatically with the rest of the verb form *megetolk*epel, to give a pronominal prefix verb form, as is used in subordinate clauses. Thus we have:

\[
\text{ki ?umuc megetolk} \text{epel and ki muc ?umegetolk} \text{epel.}
\]

If these types of cross cutting were frequent in languages it would be a serious and indeed fatal objection to a word based immediate constituent analysis, and a word based syntax, as a whole. But in the languages cited, those elements giving rise to this cross cutting are in a small minority, and in default of further evidence it would seem that they are best treated separately, and that marginal elements should not be allowed to disturb the choice of a model of description that displays considerable advantages over most of the system.

A possible example of a language ill adapted to a WP type of grammatical analysis is Mixteco as described by Pike.

---

1. The form [-iz] (horse’s mane) is, of course, only a contextually bound variant (allomorph).
that language sequences of morphemes belonging to different immediate constituents may all cohere phonologically into unitary stretches, and grammatical and phonological boundaries constantly fail to coincide. However, a different picture of this language might emerge if phonological features were not treated as definitive, and the word were defined in grammatical terms as a morphemic sequence with positional mobility within the sentence and absence or restriction of morphemes insertable within its boundaries, and with potential occurrence as a free form (or complete sentence) or having a regular lexical correspondence with such a free form (many lexical items are said to appear in two forms, one as used alone and in certain constructions, and a shorter form as used in certain other constructions).

VIII. Of the three models, IA, IP, and WP, no one has as yet been worked out to be equally suitable for every part of a grammatical system in every language, a fact hardly surprising in view of the immense complexity of language. But, if no model is wholly "right", no model that can be stated and applied in formal terms is "wrong", in the sense that a model of grammatical description relying on intuitions and on classifications based on unverifiable categories of "meanings" is, in this day and age, "wrong". WP is a workable model of formal grammatical description, offering many advantages in analysis, not to mention greater clarity in the allied field of pedagogical presentation.

This is not to say that traditional paradigm presentations exhibit the morphology of their languages in the best possible way; there is clearly scope for the examination of several different ways of organizing the material of a language into paradigms. It may also be that while each of the models discussed in this paper is feasible with every language, one of them is more appropriate with certain languages; possibly Mixteco, at least on Pike's analysis, is not a "WP language", and certainly on the evidence we have considered some languages are less suitably "IA" or "IP languages".

WP when reworked in terms of current formal criteria deserves proper consideration as a means of stating and analysing grammatical systems. It is due for a more sympathetic hearing than it has received in recent years, and increased representation among the descriptive grammars of languages newly made available to scholarship.

1 Hockett, "Two models", 233.
2 Nothing is said in this paper of the effect of applying transformational analysis on the choice of grammatical models (see N. Chomsky, Syntactic Structures, 's-Gravenhage, 1957). Chomsky, whose work, of course, post-dates Hockett's "Two models", seems himself to operate more or less in morpheme based IP terms with transformations as the major class of Processes; but it would appear equally possible and fruitful to apply transformational procedures to a word based syntactic frame.