MEANING CORRELATIONS AND SELECTIONS IN MORPHOLOGY-SYNTAX PARADIGMS

CHARLES F. VOEGELIN

Indiana University

A discussion on morphology-syntax is not easy to open, since languages which have an immediately apparent word morphology are not thereby less concerned with syntactic constructions than are languages without apparent or elaborate word morphology. Morphology and word morphology are the same thing; is there any other kind of morphology, strictly speaking, than that concerned with morphemic constituents of words? Some classes of words, in languages with apparent morphology, can be stated in terms of a variety of frames which have, however, a common feature: some constituents are substitutable and some are invariant. For example, each pair in English *man men, mouse mice, woman women* occurs in a frame in which the consonants are invariant while the vowels are substitutable; for some speakers the substitution appears only in the first syllable of the last pair, /wumən wimən/. Words exhibiting some such invariant constituent and either internal vowel substitution or else prefix or suffix substitution are obviously not precluded from syntactic constructions. Languages with little apparent word morphology are sometimes treated, conversely, as though the paradigm model did not apply to them.

But this is, of course, not true. Though less obviously so, the paradigm model is as applicable to stretches of successive words as it is to morphemic constituents in a single word.

Utterance frames are termed 'syntactic constructions' when the range or boundaries of the frame are selected by the linguist. When, however, the boundaries of the utterance frame are produced by the speaker and are heard by the linguist as phonological junctures, then the frame may be termed a **contour span**. In either case, one part of the frame is held invariant for testing the extent to which substitutions—whole words in syntax on the analogy to word-partials in morphology—are possible in another part of the

frame. Thus, in the following English syntactic construction, one may hold invariant the margin words of the frame, and make substitutions in the middle word of the frame: *birds eat oats, birds hate oats*. Or conversely: *birds eat oats, cows eat oats;* and for substituting the last word of the frame: *birds eat oats, birds eat ivy*.

Analysis within a word (morphology) appears to be easier, or rather—as the published record seems to attest—more successful than analysis within syntactic constructions. Why this should be so is not entirely clear, but probably a universal factor contributing to the relative difficulty or incomplete success of syntactic construction analysis is the disturbing factor of the native speaker's awareness that isolable words mark some meaning or other, however vaguely. On the other hand, the informant's unawareness of the meaning of non-free morphemic constituents within a word makes possible a relatively greater objectivity. The linguist's awareness of the value of non-free morphemes does not influence the informant's response to questioning. The informant's awareness, however vague, of the value of isolable words does influence the informant's response to questioning, and in a disturbing way.

Most disturbing of all is the informant's persistent intrusion of a problem which is technically known as **selection**. Thus, *birds eat ivy* fits in our English frame, even if birds do not include *ivy* in their diet. Now compare *humans drink cigarettes;* this fits, by both frame and by selection, in Turkish; it fits also in English—by virtue of its frame, but certainly not by selection. Anything which fits in a frame—syntactic construction or contour span—but which is possibly contrary to fact (*birds eat ivy*) may be dubiously accepted by the informant ('well, maybe'); and whatever is nonsensical for one language (*humans drink cigarettes*) is firmly rejected by an informant in that language ('we don't say that') even though, in another language, as Turkish, much the same selection would be possible and hence neither nonsensical nor unacceptable to an informant in the other language.

Linguistic analysis ultimately succeeds in discovering possible frames for syntactic constructions. Thereafter, it can determine whether an informant's reluctance to accept a new utterance reflects an improper selection in a possible frame, or an impossible frame for a possible selection. However, the native speakers themselves rarely distinguish between doubt or rejection attributable to impossifle selection, from doubt or rejection attributable to improper syntactic construction. Thus, the same English informant who rejects **humans drink**

cigarettes will also reject *birds drink waters* (reflecting the restriction on plural form of mass nouns, *waters*, where a counting noun might appear, as in *birds eat bananas*); he will also reject *birds eat a*, as well as *birds eat the* (because *a* and *the* never appear finally in English syntactic constructions).

In frames for morphemic constituents of words, the most recurrent word-partials—as prefixes or suffixes—are only indirectly associated with a referent, namely, through their association with the remaining word-partial of the given isolable word. Thus, the prefix ʔi- *my*, occurs in the isolable word frame ʔi-N (noun) in Hopi: ʔi-vóko *my dog*. Any Hopi-English bilingual speaker can cite the Hopi word for *dog*, and is accordingly said to be aware of the word-partial marking *dog* in Hopi. By the same logic, he is said to be unaware of the word-partial for *my* because he cannot translate the Hopi prefix for *my* when ʔi- is uttered by itself, nor can he cite this minimum morpheme for *my*. When asked to say *my* in Hopi, the informant says ʔi-hími *my belongings*, which fits in the frame ʔi-N. We know that the same prefix does not fit in the frame ʔi-V (verb), but we can experimentally prefix ʔi- to some V, as that for *running*, to find out how the informant will respond to possible morphemes in an impossible frame. The experimental result is that ʔi-V is neither translated *my running*, nor rejected outright or even said to be somewhat awkward; it is either not understood at all, or else revised with a replacement of the Hopi form for *I* in place of that for *my*. If he has a glimmering of what is wanted, the informant corrects from the impossible frame ʔi-V to the possible frame ni-V: *I'm running*.

An error in filling the frame of morphemic constituents of isolable words is apt to be corrected; it is probably less easy for the informant to correct comparable errors when the frame is a syntactic construction or contour span, and the constituents are isolable words. It is certainly less easy for the informant to make such correction when the error involved is an error of selection. If it is nonsensical, by selection, to say *humans drink cigarettes*, then is *humans eat cigarettes* more sensible? In non-directive eliciting, the informant will respond with rejection or dubious acceptance to more or less nonsensical selection; but in deliberately directive eliciting, we can tell the informant that he is to offer correction in place of his previous reactions. Then a correction of *humans drink cigarettes* might be *humans devour cigarettes* or *humans smoke cigarettes* or *humans consume cigarettes* or the like. The basis of correction would be different in the language of advertising

than in poetic language. Whatever the stylistic basis of correction, the fact that more than one correction is possible for a wrong selection in a right frame makes it easier for the informant to doubt or reject than to correct.

The total inventory of words in any one language is impressively large to the compiler who, in effect, removes or segregates his compilation—bristling with problems of selection—from the rest of linguistic analysis. First, he lists words in a separate file or volume, the dictionary. Secondly, for purposes of further formal analysis, he reduces the unmanageably large number of words in the dictionary to a manageable number of classes and subclasses, as N for nouns. Thereafter, he can treat combinatorial possibilities without explicit reference to restrictions on selection.

Restriction by virtue of selection seems generally greater in syntax than in morphology. A mildly restrictive statement might apply to N after the Hopi prefix ʔi- *my* in the morphological frame ʔi-N; some few N may not appear in this frame, but surely more Hopi N are permitted in the ʔi-N frame than are precluded by selection. Such a restrictive statement would be much too mild for many analogous frames in syntax. Thus in the English frame exemplified by *humans smoke cigarettes*, only a few verbs are substitutable with *smoke*, and all other verbs in English are precluded by selection.

Turning now from selection to frame (irrespective of selection), we ask whether the formal structure of paradigms in morphology is or is not analogous to that of paradigms in syntax.

In the traditional morphological paradigm certain word-partials are brought together in such a way that at least one part is held invariant when combined with other replaceable word-partials. The latter comprise a substitutable set consisting of two or more members. The combination of invariant word-partial and substitutable word-partial constitutes an isolable word which exemplifies a particular paradigmatic frame.

It would, in principle, be possible to hold a single prefix invariant in respect to hundreds of substitutable morphemes found in the N class, for example in Hopi.

In practice, the converse procedure is followed in paradigmatic eliciting: a given morpheme from a large inventory class, as N, is held invariant; and this invariant is combined with morphemes from a low numbered inventory set, as possessor person markers. The usual practice is efficient, for it enables

analysis to proceed without introducing the problem of **selection** until all possible paradigmatic frames are established. Each paradigmatic frame is exemplified by a series of isolable words which share the same morphemic invariant and show successive replacements of the substitutable morphemes.

The paradigm commonly encountered in morphology, as identified above, is hereinafter used as a model for stating combinatorial possibilities among isolable words in syntactic constructions. The first parallelism to be noted is that between the **isolable word** which exemplifies the frame of a morphohological paradigm, and the **isolable contour** which exemplifies the frame of a syntactic paradigm. The boundaries of all kinds of contour spans are marked by junctures with distinctive intonation followed by pause. Contour spans, as identified below, are more objective than the traditional 'construction'.

The discussion of contours, which now follows, is relevant to stating (a) the maximum unit in syntax (which is certainly not the 'sentence'); (b) the formal analogy between constituents of paradigmatic frames in morphology and syntax.

One kind of juncture, here symbolized by #, marks the boundary of a stretch of speech uttered by itself without signaling that an immediately following utterance is to be expected. Such contours, with final boundaries marked by #, are termed **isolable contours**. In English, isolable contours may range in length from a single isolable word (as *fire#* or *please#* or *indeed#*) to short phrases (*thank you#* or *of course#* or *if you please#*) to long complex sentences, as in a technical exposition.[1]

Another kind of juncture, here symbolized by a double bar //, marks the boundary of a stretch of speech, and at the same time signals that an immediately following utterance is to be expected. Contours which end with such//junctures are termed **dangling contours**. In every language observed for contour spans, there is at least one juncture which signals an expectancy that a relevantly additional utterance will follow. In some Indonesian

---

1. Though languages so far encountered seem to present more varieties of junctures for marking **dangling contours** than for marking **isolable contours,** several kinds of isolable contours need to be recognized in terms of their combination with other contours in the same language. In Hopi, an isolable contour may begin in the middle of an isolable word (at a morpheme boundary), and hence is **dependent,** since it will appear only after a prior dangling contour. The sequence of one or more **dangling contours** and an **isolable contour** is termed a **complex isolable contour.** The sequence of two or more **isolable contours** (simple or complex) bounded by a distinctive paragraph final juncture is termed a **compound isolable contour.** It is found in Shawnee and other Algonquian languages. Enormously long **isolable contours** are also found in Dravidian languages.

languages, as Havunese, virtually every **isolable contour** is **complex,** since everyone includes a **dangling contour** prior to the simple isolable contour. In English (as in Hopi) each item in an enumeration, including numbers themselves, may be uttered as a **dangling contour** (for Hopi, sɨ·ka//lô·yöʔ //pá·yoʔ//taʔá ǂ; for English, *one//two//three//go* ǂ), In English (as in Turkish) prior clauses may end in//; such a **dangling contour** may be followed by a considerable pause before the conclusion of the **complex isolable contour** is uttered; the diagnostic part of the//juncture, however, is the intonation preceding the pause rather than the pause itself (for Turkish, *eğer siz gidersen//bende giderim* ǂ; and for English, *if you go//I guess I'll go too* ǂ).

The distinction between **isolable contours** and **dangling contours** is none too strong,[2] but it is powerful in integrating fractionated conversational structure; for example, it makes it possible to identify one **complex isolable contour** which is uttered by two speakers, A and B, in a triple conversational exchange:

(A) *I'd like a big piece of//*

---

2. **Dangling contours** from Indonesian languages are now being recorded by tape recorders and played back, contour by contour. Indonesian speakers are asked to repeat up to the point at which the tape recorder is turned off, namely at the//juncture for dangling contours. In the repetition from the playback, the//juncture or some other variety of juncture for **dangling contour** is replaced by the ǂ juncture for **isolable contour.**

Though no language has been found which lacks the distinction between one or more kinds of **dangling contours,** and one or more kinds of **isolable contours,** it must be said that the two have in common more important shared features which distinguish them: (1) Either may be included in a longer stretch of speech: a **compound isolable contour** is defined as a sequence of isolable contours bounded by paragraph final junctures, as in Shawnee; a **complex isolable contour** is defined as a sequence of at least one **dangling contour** before a **simple isolable contour,** as exemplified above (*if you go//I guess I'll go too*ǂ). (2) Either may include within its span pauseless junctures which show greater diversity in different languages than do pause junctures (open transition in Havunese is marked by pauseless junctures; but pauseless junctures mark the end of an alternation of stress sequence in Hopi); since hesitation forms are non-occurring at points marked by English pauseless juncture, +, *book+worm* for *scholar* would replace+by//if the hesitation form, -er-, were uttered for+(*he's just a book//...er...worm*ǂ). (3) Both **dangling** and **isolable contours** occur in utterance final, a position which—in terms of either the speaker's production or the hearer's expectation—is appropriately filled by an **isolable contour;** though utterance final position may be filled by a **dangling contour,** it is not so filled in the general case but only in a special case in conversational exchange (or as-though conversational exchange), namely, when (a) the speaker signals by **dangling contour** that either he or the hearer will continue the discourse, and (b) neither he nor the hearer do continue the discourse, thus allowing the **dangling contour** to remain in utterance final. Those who use the term 'terminal juncture' imply (falsely) that other junctures are necessarily non-terminable.

(B) *cherry pie//*

(A) *Just what I wanted* ‡

For conversational texts such as this it seems futile to look for the maximum linguistic stretch known traditionally as 'the whole utterance'. When does 'the whole utterance' of a conversational text between partners living together ever end? When they lapse into the silence of a moment's hesitation, or when both fall into a brown study, or when one or the other takes a nap, or when both sleep for the night? Even if partner B were to die, A could take on a new partner (B') in order to continue the conversation; B' could do likewise, with A' in place of A. The same conversation can continue between a pair of persons so long as there is a continuity in which one of the pair remains (as A) in case the other should be replaced (as B' for B). It is possible to imagine an endless conversation between A (or A') and B (or B').

The maximum Phonological stretch in linguistic analysis can be objectively identified as one or another kind of **isolable contour.** As the longest phonologically bounded utterance stretch, the **isolable contour** provides a maximum for paradigmatic analysis. Shorter contours can of course also be used as frames, but it is easier to elicit syntactic paradigms in **isolable contour** frames than in such **dangling contour** frames as *I'd like a big piece of//.* Native speakers say this sort of thing readily enough in conversational exchange. But in the stimulus and response of eliciting, even well trained informants tend to respond to attempts to obtain a **dangling contour,** as in the preceding example, with some such reformulated **isolable contour** as *I'd like a big piece of pie* ‡.

Simple isolable contours are now analogized to **isolable words:** parallel instances will be cited to show that each may exemplify frames in paradigmatic analysis.

In the model for the morphological paradigm, the **isolable word** exemplifies the frame, and the morphemic constituents within this frame are classifiable in two ways, namely, (1) either (a) **invariant** or (b) **substitutable;** and (2) either (a) **free** or (b) **bound.** The first distinction is essential for setting up paradigms, as has already been noted. In many languages there is a positive correlation between morphemic constituents which are held **invariant** (la) and those which are classifiable as **free** (2a); and between morphemic constituents which belong to a **substitutable** set (1b) and those

which are classifiable as **bound** (2b). Any morpheme in a multimorphemic **isolable word** (e.g. English *books*) is classifiable as **bound** or non-free (e.g. English -s) by virtue of its non-appearance as the sole constituent in another **isolable word;** conversely, any morphemic constituent which does appear as the sole constituent in another **isolable word** (e.g. *book*) is classifiable as a **free** morpheme.

For the syntactic paradigm, the **isolable contour** is the frame whose constituents—**isolable words**—are certainly classifiable in the first way noted for the morphological paradigm, namely as either **invariant** (1a) or **substitutable** (1b). We seem to land in a contradiction—or in an abstract world of empty classes—if we press the analogy further, and say that the constituents of **isolable contours** are also classifiable as **free** (2a) or **bound** (2b), since the constituents of **isolable contours** are **isolable words** and the latter are defined operationally as the minimum form which trained informants are willing to cite as the sole constituents of **isolable contours.** Accordingly, all **isolable words** would seem to be **free** (2a); hence, the class of **isolable words** which are **bound** (2b) would never have more than zero membership.

It might be possible, nevertheless, to press the analogy further by setting up a distinction more or less equivalent to the **free** (2a) and **bound** (2b) distinction. The equivalence here proposed is based on inventory frequency (relatively high or low), and recurrent usage frequency (high or low) of the inventory word class or word subclass of the constituents of an **isolable contour.** Actual counting for 'high' and 'low' would be necessary only if the frequencies or word class inventories were in a middle range; when the inventory membership for a given class is at one or the other end of the frequency range for all inventories in a given language, it is possible to state an immediate equivalence: low inventory word classes with **bound,** and high inventory word classes with **free** morphemic constituents. In most languages actually encountered, this kind of equivalence can be generally stated immediately, since the inventory membership in word classes are generally either at the high extreme or at the low extreme. In Indonesian languages, citable words from such inventories as P (person markers) are equivalent to **bound** (2b) morphemes: both word class P and affixes (bound morphemes) belong to low numbered inventories (therefore also 2b). In Algonquian languages, citable words from such inventories as Na (animate nouns) and Ni (inanimate nouns) are equivalent to **free** (2a) rather than to **bound**

morphemes: Na and Ni, on the one hand, and **free** morphemes on the other, belong to high numbered inventories (therefore, also 2a).

A particular instance from any of the high numbered inventories is, however, apt to be less recurrent in usage frequency than a particular instance from the low numbered inventories. Conversely stated, but with the same implication, a particular instance from the low numbered inventory is apt to be more highly recurrent than a particular instance from the high numbered inventory, assuming the low numbered and high numbered inventories are both represented in the utterance frames used for counting or estimating usage frequency.

This frequency based distinction—inventory frequency with implication for usage frequency—serves a double purpose.

First, it permits us to complete the formal parallelism between morphemic constituents of **isolable words** and the analogous constituents of **isolable contours.** As already indicated, one constituent is held **invariant** while the other or others are taken to be **substitutable**—both for the **isolable word** type and the **isolable contour** type of paradigm frames. Further, **free** morphemes among **isolable word** constituents are taken to be equivalent to **isolable contour** constituents having membership in a high numbered inventory, while **bound** is taken as equivalent to low numbered inventory.

Secondly, it gives us an objective basis for determining which constituents to hold **invariant** in an **isolable contour.** and which to take as **substitutable** for purposes of eliciting paradigms. If those constituents which belong to a high numbered inventory are held **invariant,** while those which belong to a low numbered inventory are taken as **substitutable,** then eliciting for paradigms in syntax is freed from the vacillation which often characterizes work in syntax—vacillation between random sampling and unmangeable exhaustiveness (eliciting which seeks the permutation of all combinatorial possibilities and thereby encounters an initial difficulty in distinguishing between restrictions on frame and restrictions on selection from high numbered inventories). In its place, the paradigm model requires only as many distinctions to be elicited as there are substitutable constituents in a low numbered inventory; and the paradigm frames—for some **isolable contours** as well as for the traditional inflection among **isolable words**—can be established before the problem of selection is approached. The paradigm model is applicable to syntax as well as to morphology; it is, however, applicable only when con-

stituents—of **isolable words** or of **isolable contours**—include the combination of an **invariant** and a **substitutable** in each exemplification of the paradigm frame.

COMBINATORIAL POSSIBILITIES OF PARADIGM INVENTORIES

In a frame exemplified by English *look at mine!* [V-P 41 ♯], *look at* is held invariant (V-P for 'verb-preposition' in the square brackets indicates that, by selection, other V-P than *look-at* might be held invariant in this frame), and *mine* is a substitutable member of the paradigm 40 inventory: *mine* (41), *yours* (42), *his* (03), *hers* (44), *its* (05*), *ours* (46), *theirs* (47).

So also, in *my house?* [31 N//], *house* is held invariant (N for 'noun' in the square brackets to show that, by selection, other N than *house* might be held invariant in the same frame), and *my* is a substitutable member of the paradigm 30 inventory: *my* (31), *your* (32), *his* (03), *her* (34*), *its* (05*), *our* (36), *their* (37).

The rationale of numbering paradigmatic items in parallel paradigms, and the reason for postponing immediate constituent analysis of these items is given here, before citing two other paradigms parallel to paradigms 30 and 40 in English.

Two digits are used for each paradigm number. When the second digit is 0, the resulting number (as pdg. 30 or pdg. 40, above) refers to the set as a whole rather than to any particular paradigmatic item in the set; such numbers as pdg. 30 or pdg. 40 permit an indication of item membership in a given inventory, without indication of successive distinctions possible in that inventory (as between 31, 32, etc. or as between 41, 42, etc.). When the first digit is 0, the resulting number (as 03 or 05* above) indicates intersection: paradigmatic items with membership in two or more paradigms. Thus, English *his* (03) is one item and *its* (05*) is another paradigmatic item; each one is listed in the inventory of pdg. 30 and again listed in the inventory of pdg. 40 because each is substitutable with the non-intersecting items in these inventories. (31, 32, 34*, 36, 37 of pdg. 30 are non-intersecting with 41, 42, 44, 46, 47 of pdg. 40.) Both digits are higher than 0 when numbers refer to non-intersecting paradigmatic items.

By immediate constituent analysis the morphemic components of 42, 44, 46, 47 might be stated as 32, 34*, 36, 37 plus an -s suffix (*your* (32)+-s= *yours* (42), and so on); the same analysis does not quite hold between *my*

(31) and *mine* (41), and the intersecting members which are included both in the pdg. 30 and pdg. 40 inventories—*his* (03) and *its* (05*)—would call for a still different analysis. It seems more efficient to postpone such analyses until meaning correlations are introduced (see below); analysis here is confined to the combinatorial possibilities between invariants selected from lexicon inventories (as from N, V, and V-P) and paradigmatic items (isolable words taken as the substitutable constituents in such paradigm frames as [30 N//] and [V-P 40 #]).

After comparing the paradigm frames [10 V//] and [V-P 20 #], it will be possible to revise the tentative numbers assigned to items 34* and 05* in the preceding paradigms (30 and 40), and 24* and 05* in paradigms 10 and 20.

In the frame exemplified by *I slipped?* [11 V//], *slipped* is held invariant (other V than *slipped* might, by selection, be held invariant in the same frame), while *I* is a substitutable member of the paradigm 10 inventory: *I* (11), *you* (02), *he* (13), *she* (14), *it* (05*), *we* (16), *they* (17).

So also, in *look at me!* [V-P 21 #], *look at* is held invariant (other V-P than *look at* might, by selection, be held invariant in the same frame), while *me* is a substitutable member of the paradigm 20 inventory: *me* (21), *you* (02), *him* (23), *her* (24*), *it* (05*), *us* (26), *them* (27).

Intersections between paradigmatic items in 10 and 20 are given a number with 0 as the first digit, and then entered in the inventories of both paradigms: *you* (02) and *it* (05*). When these paradigms are compared in pairs (30, 40 and 10, 20) *her* (34*) turns out to be non-intersecting with *hers* (44), just as *she* (14) turns out to be non-intersecting with *her* (24*). But by looking at all four paradigms, it is apparent that *her* (24*) and *her* (34*) intersect and must therefore be restated as a single paradigmatic item; the item index number is revised from 24* and 34* to 04. But the single paradigmatic item *her* (04) remains in the inventory of paradigm 20 as well as in the inventory of paradigm 30, since it is substitutable with non-intersecting items in each paradigm.

Another index number revision is needed for *its* (05*), intersecting in paradigms 30 and 40, and *it* (05*) intersecting in paradigm 10 and 20; since the items *it*, *its* are non-identical, they must be numbered separately. It is desirable to have the index numbers for *it* and *its* revised so as to suggest that the paradigmatic items numbered are neither identical nor wholly

different; the index number for *it* is revised to 05.1, that for *its* to 05.2, because it turns out that *it* (05.1) and *its* (05.2) mark an identical successive distinction (see below).

Compare now affix paradigms in Hopi-Tewa.

The minimum frame for the four Hopi-Tewa paradigms treated below is the isolable word; the treatment which follows is based on John Yegerlehner's analysis which includes not only the single word frames, but also frames for which two or more isolable words are relevant.[3]

An inventory of paradigmatic items is given for paradigms 10 through 40. One stative verb stem (SV) is held invariant, and all items from pdg. 10 and 20 are substitutable in respect to this SV invariant. Another stem, active verb (AV), is held invariant when items from pdgs. 30 and 40 are shown to be substitutable in respect to the AV invariant. Hence it is not necessary to treat the divisive function of pdgs. 10, 20 versus pdgs. 30, 40 as a meaning correlation; it is possible to state this function in combinatorial terms: stems combinable with pdgs. 10, 20 are classified as SV and stems combinable with pdgs. 30, 40 are classified as AV, not to mention an overlap stem class (SAV) combinable with all paradigms treated here.

The inventory of items for pdg. 10 is listed in three groups of three items each; the items in each group distinguish *singular, dual,* and *plural;* the use of digits for indexing differs from that employed for English, above, by the addition of a digit in parentheses for *singular persons* (1), *dual persons* (2), and *plural persons* (3). Thus, for the paradigm 10 inventory: 11(1) dín-, 11(2) gáh-, 11(3) gíh- for *first person* quantified as *singular* (1) *dual* (2), and *plural* (3); ʔǘh- 12(1), dę́h- 12(2), ʔóh- 12(3) for *second person* quantified as indicated by the parentheses; 13(1) ʔę̀ù̧-, 13(2) dén-, 13(3) dóó- for *third person,* likewise quantified for *singular* (1), *dual* (2), and *plural* (3). One possible frame which includes pdg. 10 is [10–SV], exemplified by the selection of one SV (-hwì *be trapped*) which is held invariant while any item in the inventory of pdg. 10 may be prefixed to it, as dín-hwì *I have something trapped* [11(1)–SV], gíh-hwì *we have something trapped* [11(3)–SV], and so on.

The inventory for pdg. 20 is 21(1) ʔò-, 21(2) gà-, 21(3) gì- for *first person* quantified as *singular* (1), *dual* (2), *plural* (3); 22(1) ʔú-, 22(3) ʔì- for *second person,* either *singular* (1) or *plural* (3), but not *dual;* 23(1) nà-, 23(3) dì-

3. John Yegerlehner, Phonology and Morphology of Hopi-Tewa, Ph. D. Dissertation (Indiana, 1957).

—102—

for *third person*, either *singular* (1) or *plural* (3), but not *dual;* 24(2) dà-quantified as *dual* but neither ar *singular* nor as *plural* for *non-first person* —hence there is a possible ambiguity as to whether 23(2) marks *second* or *third perscn*. The selection of -hwì from the SV subclass for pdg. 10 in [10–SV] may be repeated here in the frame [20–SV], exemplified as ná-hwì *it's trapped* [23(1)–SV], and so on.

Since -hwì *be trapped* is selected from the SV subclass of verb stems, it is incompatible in the frame [30–AV]; our selection from the active verb (AV) subclass is -mùn *see*, which is held invariant in frames which include either pdg. 30 or pdg. 40. A possible frame which includes pdg. 30 is, as noted, [30–AV]; this frame is exemplified by díí-mùn *I see myself* [31(1)–AV] or, of course, by any item in the pdg. 30 inventory which may be prefixed to the AV selected: 31(1) díí-, 31(2) ʔàn-, 31(3) ʔíbí- for *first person* quantified as *singular* (1), *dual* (2), *plural* (3); 32(1) bì-, 32(3) ʔóbí- for *second person*, *singular* (1) and *plural* (3) respectively; so also, 33(1) ʔì-, 33(3) díbí- for *third person*, *singular* (1) and *plural* (3) respectively; but 34(2) dèn- *non-first person*—i.e. either *second* or *third person*—quantified as *dual* (2).

The selection of -mùn from the AV subclass for pdg. 30 in [30–AV] may be repeated here in the frame [40–AV], exemplified by dó-mùn *I see it* ([41(1)–3–AV] in which –3 after 41(1) marks *third person*, here translated as *it*, and so likewise in the inventory—–3 serves to index *third person goals* or *recipients*, without committment as to sex or number. The frame [40–AV] may be exemplified by any item in the pdg. 40 inventory prefixed to the particular AN which is held invariant: 41(1)–3 dó-〜dò-, 41(2)–3 ʔán-, 41(3)–3 ʔíí- for *first person* quantified as *singular* (1), *dual* (2), *plural* (3) respectively, and functioning as *agentive* or *actor* in respect to a non-quantified *recipient* or *goal third person;* 42(1)–3 náá-, 42(3)–3 ʔobíin- for *second person singular* (1) and *pluarl* (3) respectively, as *agentive* in respect to *third person goal;* so also, 43(1)–3 mán-, 43(3)–3 díí- for *actor third person singular* (1) or *plural* (3) in respect to *goal third person;* but 44(2)–3 dén- *actor non-first person*—either *second* or *third person*, but in either case specifically *dual* (2) —in respect to a non-quantified *third person goal*.

SELECTION AND MEANING CORRELATIONS

Up to this point, discussion has been restricted to combinatorial possibilities (invariants selected from lexicon inventories and substitutable items from

paradigm inventories—in frames); and the membership of frame constituents has been given above—in listed paradigm inventories for English and Hopi-Tewa, and in non-listed lexicon inventories (V-P, N, V for English and SV, AV for Hopi-Tewa). This kind of combinatorial possibility structure might be characterized as *fait accompli* structure—operationally demonstrated and entirely successful, as far as it goes. It does not go far enough to give any relevant meaning correlations, since all that is relevant is the fact that every paradigmatic item in a particular paradigm inventory is substitutable with every other item in a frame which includes at least one selection from a particular lexicon inventory, the so-called 'invariant'.

Meaning correlations are desirable only if they are relevant to one or another kind of structure. The information for such correlations are (in a rough and ready way) obtainable without difficulty—from native speakers (informant's testimony), from observations of meaning 'overlap' and 'remainder' in partially similar paired utterances (linguist's inference), or more generally, from construing texts in the tradition of philology. Meaning correlations, in general, seem to be irrelevant of unnecessary, or at least less relevant or less necessary, when analysis is based on a closed corpus than on an open corpus.

The *fait accompli* structure, illustrated above, is obtained from a half-open corpus. The items listed in the English as well as the Hopi-Tewa inventories 10, 20, 30, 40 are obtained by eliciting which is continued as long as new items are found for any one of the paradigm inventories, and which is discontinued when it turns up the same items over and over again, without discovering any additional items for the paradigm inventories. This leads one to predict that future utterances (hence 'open corpus') will likewise fail to show up any new items for the substitutable half of the paradigmatic frame; this constitutes the half-open part of the *fait accompli* structure. In contrast, the half-closed part is concerned with selection freedoms and restrictions from the lexicon inventories V and N and V-P (in English); selection is left unstated beyond the fact that one selection from each (hence 'closed corpus') is tested for the exemplification of each paradigm frame. Thus *look at me!* [V-P 21 #], is tested for *look at you!* [V-P 02 #] and so on for other items in the paradigm 20 inventory. But this does not guarantee that another selection in [V-P 20 #], as *wink at me!* [V-P 21 #], can be exemplified for all the items in the inventory of paradigm 20; in fact, though it makes no sense in English to say *wink at you!* [V-P 02 #], it is possible to say

something like this in a closely similar frame: *wink at you?* [V-P 02//].

Whether based on eliciting (half-open corpus) or wholly on texts (wholly closed corpus), the *fait accompli* structures are predictive in a probablistic sense—give some guidance for generating new utterances—only to the extent that lexicon selection is tested. Wholly predictive structures, resulting in a grammar for generating untested utterances, are obtainable in more than one way: by extending the *fait accompli* structure from paradigm inventories to a manageable part of a lexicon inventory (a particular domain within a lexicon inventory), and then predicting that permissive selection as well as restrictions on selection in the tested part of the domain will hold for the untested part of the domain;[4] by extending the discoveries made by Harris and Chomsky[5] on convertibility between two *fait accompli* structures which contain the same lexical items ('transformations') to converting restricted selection to permissive selection by restating the same lexical items from an impossible contour frame (*wink at you!* [V-P 02 #]) to a possible contour frame (*wink at you?* [V-P 02 //]). Predictions for transformations between two *fait accompli* structures can be made without recourse to meaning; predictions for extending a *fait accompli* structure from the sampled part of a domain to the remainder or untested part of the same domain can be made only (a) after extensive eliciting which is sensitive to meaning correlations, (b) after the loci of meanings are identified (see below).

For example, with English V as transitive verb (*drop it!* but not **drop you!*) or with V-P, as above, eliciting reveals that 02 is generally less combinable with selections from V or V-P than are other members of the pdg. 20 inventory. A greater combinatorial probability would be reached if the frame were revised from [V-P 20 #], as above, to [V-P 20 (minus 02) #]. A few random selections of new V-P (as *sit on, talk to, run from, play with* as well as *wink at*, given above) all show combinatorial possibilities with almost any item from the inventory of pdg. 20 (except 02): *sit on him!* (but not **sit on you!*), *talk to me!* (but not **talk to you!*), *run from them!* (but not **run from you!*), *play with us!* (but not *play with you!*). The restriction on [V-P 20 (minus 02) #] is not a structural restriction but rather a selection

---

4. C. F. Voegelin and Florence M. Voegelin, Hopi Domains, A Lexical Approach to the Problem of Selection, IUPAL 14 of IJAL (1957).

5. Relevant reference to Harris is included in Noam Chomsky, Syntactic Structures (The Hague, 1957).

restriction; the only true structural statement is [V-P 20 #], without restriction (without "minus 02").

This particular restriction is, nevertheless, relevant to combinatorial possibility statements since the locus of the probabilistic restriction can be nicely placed in terms of the *fait accompli* structure, i.e., the paradigmatic frame which has been tested to work with at least one selection from a lexicon inventory (*look at you!* [V-P 02 (in pdg. 20) #]).

By analogy to restrictions on selection, meaning correlations are not part of *fait accompli* structure and hence not included in frame constituent formulae; they are nevertheless relevant to combinatorial possibility statements since (like the selection restriction involving 02, above) these meaning correlations can be ascribed to a particular locus in the *fait accompli* structure.

LOCI OF SUCCESSIVE DISTINCTIONS

The four parallel paradigms cited from English can be said to mark exactly the same successive distinctions; those cited from Hopi-Tewa show some same and some different successive distinctions.

Among parallel paradigms, the total of all paradigmatic distinctions are found (a) when one pair of items after the other is compared within a single paradigm (as in English pdg. 30, the distinctions between *my* and *your*, *your* and *his*, *his* and *her*, etc.) until all items included in the inventory are compared—these are labeled **successive distinctions;** (b) when the common meaning of all items in a given paradigm—the meaning which is shared or non-variable or constant for that paradigm, hence labelled *k*—is compared with the common meanings, the *k* of other parallel paradigms (as in Eng. *k* of pdg. 10 is *actor*, *k* of pdg. 20 is *goal*, etc.). Assuming now that the loci for the common meaning—the *k* of each paradigm—are identified for English and Hopi-Tewa (see below), it is possible to state what remains as the loci of **successive distinctions.**

The successive distinctions of English parallel paradigms are all concerned with *person*, and in addition, with *number* of person (*singular* or *plural*) when the person is *first* or *third*, and in addition, with *gender* of person (*masculine*, *feminine*, *neuter*) when the person is *third singular*. Among the four parallel paradigms of English, the successive distinctions are identical; loci for *singular first person* in 11, 21, 31, 41; loci for *plural first person* in 16, 26, 36, 46; loci for *singular third persons* (with three items for the three

genders in each paradigm) in 13, 23, 03 (in pdgs. 30, 40), and in 14, 04 (in pdgs. 20, 30), 44, and in 05.1 (in pdgs. 10, 20), 05.2 (in pdgs. 30, 40); loci for *plural third person* (non-committal as to *gender*) in 17, 27, 37, 47; loci for *second person* (non-committal as to *number*) in 02 (in pdgs. 10, 20), 32, 42.[6]

Compare now the four affix paradigms in Hopi-Tewa.

For Hopi-Tewa, the same successive distinctions are marked for *first*, *second* and *third persons* quantified for *singular* (1) in pdgs. 10, 20, 30; and in pdg. 40 a non-quantified *third person goal* or *recipient person* is added, as indicated by −3 in the formulae. The loci for these successive distinctions are simply the person marking components of the items remaining after the *k* of 10, the *k* of 20, the *k* of 30, and the *k* of 40 are extracted. Accordingly, the loci for *first person* are d- in 11(1), 41(1)−3 and ʔo- in 21(1); those for *second person* are ʔi̧- in 12(1), 22(1) and b- in 32(1) and náá- in 42(1)−3; those for *third person* are ʔi̧i̧- in 13(1) and na- in 23(1) and ʔ- in 33(1) and mán- in 43(1)−3.

For *persons* quantified as *dual* or *plural* in Hopi-Tewa, a pair of loci are identified for each successive distinction, one locus for *person*, and one locus for *dual* or *plural*. Thus, the loci for *dual* are -a- in 11(2), 21(2), 24(2), 31(2), 41(2)−3 and -e- in 13(2), 34(2), 44(2); the loci for *plural* are -i- in 11(3), 21(3), 22(3), 23(3), 31(3), 33(3), 41(3)−3, 43(3)−3, and ʔó- in 12(3), 32(3), and -óó- in 13(3).

While the loci for *persons* quantified for *dual* (2) or *plural* (3) may be identical, they differ in all instances from the loci for *persons* quantified as *singular* (1), as given aboven above. Thus, beside d- and ʔo- above, additional loci for *first person* are g- in 11(2)(3), 21(2)(3), and ʔ- in 31(2)(3), 41(2)

---

6. It is possible to reorganize items from the English paradigms under three main groups in terms of the successive distinctions (meaning correlations). When restricted to paradigms 20 and 30, these three successive distinction groups are then convertible into combinatory statements with reference to compounds whose first member is an item from one of the three groups and whose second member is either *self* or *selves*.

In the first group—successive distinctions for *third person*—paradigmatic items (from pdg. 20) marking *singular* distinctions (23, 04, 05.1) serve as first member, before *self* as second member in the compounds *himself, herself, itself;* and the item marking *plural* successive distinction (27) serves as first member, before *selves* as second member: *themselves.*

In the second group—successive distinctions for *first person*—compounds show the same second members (*self, selves*) after items (from pdg. 30) marking *singular* (31) and *plural* (36) respectively: *myself, ourselves.*

In the third group—successive distinction for *second person non-committal as to number*—a single paradigmatic, item (32) appears in compounds both before *self* and *selves: yourself, yourselves.*

(3)-3; so also additional loci are identifiable for *second* and *third persons* quantified either as *dual* (2) or *plural* (3).

As noted, the successive distinctions among the four Hopi-Tewa paradigms are identical when the persons are quantified as *singular;* they are not identical when the persons are quantified as *dual* or *plural*. For the latter, the successive distinctions in pdg. 10 are *first, second* and *third persons* (*dual* and *plural* respectively): 11(2) and 11(3), 12(2) and 12(3), 13(2) and 13(3). This differs from the successive distinctions in pdg. 20 and 30 (which do not differ from each other); *first person* is either *dual* or *plural*, while *second* and *third persons* are only *plural* (not *dual*) and, conversely, *non-first person* is only *dual* (not *plural*): 21(2) and 21(3) as well as 31(2) and 31(3); 22(3) and 23(3) as well as 32(3) and 33(3); 24(2) as well as 34(2). The successive distinctions of pdg. 40 differ from those of pdgs. 20 and 30 only by the addition of a *third person non-quantified recipient* or *goal* marked by each paradigmatic item (-3 in the formulae).

LOCI OF *K*

The meanings assigned to *k*s for English paradigms 10, 20, 30, 40 are narrowed so as to be applicable primarily to the exemplification frames cited: *actor* for *k* of pdg. 10; *goal* or *end-point* for *k* of pdg. 20; *possessor relative to adjacent specification of what is possessed* for pdg. 30; *possessor relative to omitted or non-adjacent specification of what is possessed* for pdg. 40. A single overarching *K* for all four parallel paradigms in this contra-constant system may be said to have a common meaning, *person;* though no particular locus can be found for this *K*, more than one loci can be identified for the *k* of each paradigm in the contra-constant system.

The loci of the *k*s of the four paradigms are found (a) uniquely in the frame, (b) redundantly in the frame and in additive contrasts, (c) in non-additive contrasts between pairs of paradigmatic items which mark the same successive distinction (q.v. above).

(a) Intersection is found where a given successive distinction (*second person non-committal as to number* and one *third person singular* for each of the genders), is marked by a single paradigmatic item which is included in two paradigms; *you* (02) in pdgs. 10, 20; *his* (03) in pdgs. 30, 40; *her* (04) in pdgs. 20, 30; *it* (05.1) in pdgs. 10, 20; *its* (05.2) in pdgs. 30, 40. For any pair of paradigms sharing a single item, the *k* distinguishing one paradigm

can be marked neither by an additive nor a non-additive contrast in item, since the same item serves for both paradigms; hence, the only locus remaining for distinguishing one *k* from the other lies in the contrast between the paradigmatic frames. For example, the *k* of pdg. 10 and the *k* of pdg. 20 show their frame loci by the contrast between [10 V//] and [V-P 20 ♯] for *you* (02) both in *you slipped?* [02 (in 10) V//], and in *look at you!* [V-P 02 (in 20) ♯]; so also, *it* (05.1) fits equally well in both frames. Similarly, the *k* of pdg. 20 and the *k* of pdg. 30 show their frame loci by the contrast between [V-P 20 ♯] and [30 N//] for *her* (04) both in *look at her!* [V-P 04 (in 20) ♯] and in *her house?* [04 (in 30) N//]. Finally, the *k* of pdg. 30 and the *k* of pdg. 40 show contrast in frame loci between [30 N//] and [40 C-A ♯] for *his* (03) both in *his house?* [03 (in 30) N//], and in *his is green* [03 (in 40) C-A]; so also, *its* (05.2) fits equally well in both frames (*is green* is held invariant in the exemplification of C-A, copula-adjective).

(b) Similar contrasts could be shown between non-intersecting items; thus, the *k* of pdg. 40 is marked not only in frame loci, as above, but also (redundantly) by the suffixation of /-n∼-z/ which is, accordingly, said to be one of the loci for *k* of pdg. 40. Compare *my* (31), *mine* (41) where /may/ +/-n/=/mayn/. And for suffixation of /-z/ (orthographic -s), compare *your* (32), *yours* (42); her (04 in pdgs. 20, 30), *hers* (44); *our* (36), *ours* (46); *their* (37), *theirs* (47). Less simple instances of suffixation are found, as between *you* (02 in pdgs. 10, 20) and *your* (32); here /-r/ might be said to be one of loci of *k* of pdg. 30, though /-r/ is suffixed to /yu-/ rather than to /yuw/. So also, as between *he* (13) and *him* (23) or *his* (03), it would not be /hiy/ but /hi-/ to which /-m/ or /-z/ are suffixed. Likewise, between *they* (17) and *them* (27) or *their* (37), it would not be /ðey/ but /ðe-/ to which /-m/ or /-r/ are suffixed. Where an intersecting paradigmatic item includes a possible suffix, as /-s/ of *its* and /-z/ of *his* (05.2 and 03, respectively, in pdg. 30 and pdg. 40), the suffix can obviously not be set up as one of the loci of *k* for either paradigm, since each item in question, with its included suffix, is identical rather than divisive for the two paradigms.

(c) All the paradigmatic items remaining to be treated show non-additive contrasts as loci for different *k*s of pdg. 10, 20, 30 with the same successive distinctions: *I* (11), *me* (21), *my* (31); also *she* (14), *her* (04 in pdgs. 20, 30); and *we* (16), *us* (26), *our* (36). Though the successive distinction of item 14 is the same as that for item 04—*singular third person, feminine*—the locus

of *k* of pdg. 10 (or rather, one of the loci of this *k*) is identifiable by the complete difference in shape between *she* (14) and *her* (04). Though partially similar, *me, my* are nonetheless regarded as showing the same kind of contrast (non-additive); here the loci of *k* are differentiated by the non-additive contrasts, while the successive distinction is the same (*first person singular*). The morphemic distinctiveness of *I, me, my* is thereby attested (and the argument for their allomorphic status is weakened); so also with another same successive distinction (*first person plural*): *we, us, our.*

A given paradigm, or an item in a paradigm, may have multiple loci of *k*—frame loci, additive loci, non-additive loci. Conversely, *k*s are uniquely crucial whenever it is possible to represent items from two different paradigms in the same combinatorial relationship for a pair of frames, as [V-P 20 ♯] and [V-P 40 ♯], and when the same successive distinction is represented in both frames—as *singular first person* both in *look at me!* [V-P 21 ♯], and in *look at mine!* [V-P 41 ♯].

In summary, the loci of *k* of pdg. 10 are non-additive—contrast of items 11 with 21, 31, and contrast of 13 with 03, and contrast of 14 with 04, and contrast of 16 with 26, 36 (and by absence of additive loci for item 17); and (uniquely) in frame contrasts—pdg. 10 vs. pdg. 20 (for items 02, 05.1).

The loci of *k* of pdg. 20 are additive—/-m/ included in items 23, 27; and non-additive—contrast of items 21 with 11, 31, and contrast of 26 with 16, 36; and (uniquely) in frame contrasts—pdg. 20 vs. pdg. 10 (for items 02, 05.1), and pdg. 20 vs. pdg. 30 (for item 04).

The loci of *k* of pdg. 30 are additive—/-r/ included in items 32 and 37; and non-additive—contrast of items 31 with 21, 11, and contrast of 36 with 26, 16; and (uniquely) in frame contrasts—pdg. 30 vs. pdg. 40 (for 03, 05.2), and pdg. 30 vs. pdg. 20 (for 04).

The loci of *k* of pdg. 40 are additive—/-n/ included in item 41 and /-z/ included in items 42, 44, 46, 47; and (uniquely) in frame contrasts—pdg. 40 vs. pdg. 30 (for 03, 05.2).

In contrast to the English paradigms summarized above, the loci for *k* in the four Hopi-Tewa paradigms are all additive, as would be expectable for an affix paradigm system in contrast to an isolable word paradigm system.

The Hopi-Tewa meaning of *k* of pdg. 20 is stated in very general terms by Yegerlehner, since the statement was made wide enough to cover more examples than those cited here: *'recipient in a specified state'*. The meaning

of *k* in pdg. 10 is *'recipient, in a specified state, is possessed'*, thus differing from *k* of 20 only by the addition of *'is possessed'* (cp. above, the similarly close meanings of Eng. pdgs. 30 and 40). For Hopi-Tewa, the meaning of *k* of pdg. 30 is *'agent and recipient are identical'*; that of *k* of pdg. 40 is *'one recipient is undergoing specified action performed by another'*.

Of the four *k*s in Hopi-Tewa, only *k* of pdg. 20 has a single locus: low tone. When this additive component, /ˇ/, this *k* of pdg. 20 is extracted from the vowels of paradigmatic items 21, 22, 23 (for (1), (2), (3), respectively), the remaining vowels may be termed linear vowels; but adjacent consonants in the paradigmatic items also remain. The remaining linear vowels and adjacent consonants serve as loci for the successive distinctions (treated above). This statement in analysis can be reversed to obtain a synthesis in statement: the additive component of low tone /ˇ/ (locus of *k* of pdg. 20), when combined with linear vowels and adjacent consonants which mark successive distinctions (in pdg. 20), constitute the paradigmatic items listed in the inventory of pdg. 20 (above).

The loci of *k* of pdg. 10 are -ín- in 11(1), and low tone with adjacent consonant, /ˇh/, in 12(2) and (3), 12(1), (2), (3), and /ʔų̀ų̀-/ in 13(1)—'portmanteau morph' in the sense of marking both *k* of pdg. 10 and successive distinction 13(1)—and high tone with adjacent consonant /ˊn-/ in 13(2), and -óó in 13(3)—another 'portmanteau morph',

A similar variety is shown for the loci of *k* of pdg. 30, and a greater variety for the loci of *k* of pdg. 40. Details may be found in the forthcoming publication of John Yegerlehner's analysis (IJAL, 1958 and 1959).