Introduction

Pages ix–xxii of
Signal, Meaning, and Message: Perspectives on sign-based linguistics
Edited by Wallis Reid, Ricardo Otheguy and Nancy Stern
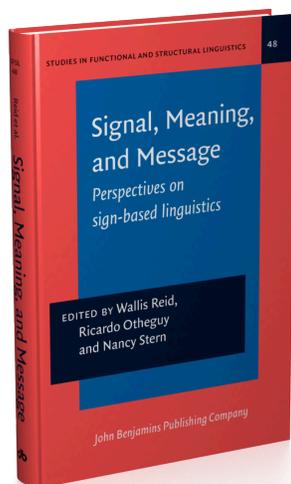[**Studies in Functional and Structural Linguistics**, 48]
2002. xxii, 413 pp.

Introduction: Sign-Based Linguistics

Wallis Reid
Rutgers University

This is the second volume of papers on sign-based linguistics to emerge from Columbia School linguistics conferences, the first being *Meaning as Explanation: Advances in Linguistic Sign Theory* (1995). In her excellent introduction to that volume, Contini-Morava outlines the defining characteristics of sign-based theory, weaving together themes that are prominent in not only Columbia School work but in the Jakobsonian and Guillaumist frameworks as well. As a comprehensive orienting discussion it cannot be improved on. Therefore this introduction will assume a basic familiarity with sign-based theory, especially as carried out by Columbia School scholars, and elaborate certain themes that relate to the articles in this collection.

1. Meaning as Explanation

Contini-Morava and Goldberg named their volume *Meaning as Explanation,* yet the casual reader might well come away not having grasped the true import of that title. The place of meaning in linguistic theory has been one of the most contentious in the field. Its treatment has ranged from indefinite postponement by Bloomfield, guarded entry by Chomsky, to the central place it occupies in cognitive grammar. On one point, however, there has been general agreement: a major goal of any linguistic theory is to offer an explicit account of meaning in language. Columbia School turns the tables here and makes linguistic meaning the *explanans* in linguistic theory rather than the *explanandum.* Systemic meaning functions to explain the distribution of linguistic forms rather than being itself the object of explanation. Meaning, then, is explanation

in both its technical and its everyday sense. In its technical sense – as a *signifié* – linguistic meaning is the immediate, theory-internal explanation for the appearance of a linguistic form; the form is there because the language user has chosen to signal its associated *meaning*. And in its everyday sense, it is the higher-level, explanation of that choice: the language user has chosen that *meaning* because it facilitates the communication of some intended *message*.

The analytical focus of Columbia School linguists over the past thirty years has been to develop this framework in the area of grammar, wherein grammatical morphology is treated as fully-fledged meaning-bearing units. Radmila Gorup's examination of Serbo-Croatian pronouns illustrates this two-tiered explanatory role of meaning. Gorup posits a three-member Deixis system for three forms that are glossed in English as 'this' and 'that'. Each member of this system signals a different level of attention, High, Mid or Low, with respect to the referent of the pronoun. The appearance of a particular pronoun in discourse is then attributed to the speaker's desire to signal its associated level of attention.

Clearly such an explanation would be vacuous without any account of why a speaker would want to signal a particular level of attention. Gorup provides this by first specifying a communicative function for the three meanings: namely, to help the hearer locate the referent. In this referent-finding strategy, High Deixis implies a hard-to-find referent, Mid Deixis implies medium difficulty and Low Deixis implies a referent that poses little or no difficulty. She then specifies the formal and semantic correlates of difficulty. Low Deixis is used when no particular referent is intended, as in a sentence that translates "Probably ONE who does not himself know who he is cannot love". Mid Deixis is for referents that have been previously and recently mentioned through a noun. High Deixis is employed for referents that are embedded in a difficult to process communication, as, for example, when the pronoun effects a shift in grammatical subject. Finally, she tests for the predicted statistical favorings of each deictic pronoun for the objectively defined context she has specified.

Treating pronouns in semantic terms is quite consistent with traditional paradigms of linguistic analysis. Sign-based linguistics has extended the explanatory scope of meaning to phenomena that are usually dealt with in non-semantic, syntactic terms. Alan Huffman tackles the problem of full-verb inversion in English. Rather than treating the normal *subject-verb* order (e.g.,

An explanation comes first) as canonical, and then attributing post-posed subjects (e.g., *First comes an explanation*) to syntactic triggers, Huffman treats both sequences as signals of grammatical meanings that relate to discourse focus. Each occurrence of *subject-verb* and *verb-subject* is attributed to the speaker's desire to signal one or the other of these two meanings; and that desire is, in turn, attributed to the discourse function of those meanings in structuring the ongoing narrative. Huffman tests for the claimed contrasts of discourse focus between the two word-order signals through qualitative and quantitative analysis of *O Pioneers!* and *Lord of the Flies.*

2.   Integration with Cognitive Grammar

Sign-based theory shares much with cognitive schools of linguistics, while still differing in important ways. The latter account for various aspects of linguistic form in terms of general cognitive and perceptual biases but do not posit sharply delimited linguistic signs of the Columbia School type. Yet these two modes of explanation need not be mutually exclusive, and Huffman's study shows how they can be synthesized. The preposing of an adverbial (e.g., *First comes an explanation)* is treated as an instance of "iconic"or "natural" word order, natural because it reflects the universal cognitive principle of putting orienting information first. The post-verbal position of *an explanation*, on the other hand, is a manifestation of a linguistic sign, a language-particular, morpho-semantic feature of arbitrary structure.

Ellen Contini-Morava's analysis of noun class markers in Swahili shows another way of integrating sign-oriented and cognitively-oriented approaches to linguistic structure. Like gender markers, Swahili noun class markers are associated with classes of nouns that share no single set of defining semantic characteristics. But unlike the European gender classes, the Swahili classes are far more numerous, ten by one count; and, more importantly, the classes have internal semantic structure. The meanings of the nouns form metaphoric, metonymic and associative connections that link them together into a semantic network. As a result, switching its marker can sometimes alter the interpretation of a stem. Nevertheless the markers prove resistant to analysis as straightforward meaning-bearing units because their effect is too variable or intermittent to yield a consistent semantic value. The analytical problem is what to say

about the markers. How can they have meaning when the class members themselves share nothing in common?

Contini-Morava begins by showing how three Swahili noun classes are structured semantically. The picture resembles the network of the semantically related senses of a lexical item as posited by linguists working in the framework of cognitive grammar. Next, she examines the communicative function of noun classes and proposes that they maximize the use of pronouns instead of nouns. She then considers four analytical options for the markers themselves and settles on one that recognizes the semantic diversity of each class while still positing a single value for each marker. The key to her solution is a new kind of linguistic meaning, one that reflects not cognition, but the mechanics of communication.

Michael Wherrity's study of Irish influence on American usage is a third instance of a cognitively inspired treatment. A curious phenomenon in the speech of the Irish and many Irish Americans is the frequent occurrence of *on* + personal pronoun or name to characterize an event as happening to the disadvantage of the party involved: *Now look what you went and did! You broke the table on me!* Although this parallels Irish Gaelic usage, Wherrity finds the notions *loan translation* and *calque* unsuitable here because there is nothing structurally innovative in the English of Irish Americans. Neither does the usage seem to reflect a change in the meaning of *on*. Consequently, this is not a case of structural change resulting from language contact.

Wherrity sees here conceptual, not linguistic, innovation. Contact speakers began expressing messages that were innovative for English speakers but which were commonplace for Irish Gaelic speakers. They could do this without systemically altering English because of what Wherrity calls "conceptual affinity". The new messages expressed with *on* could be seen as an extension of already established *on* messages, such as *to play a joke on someone*. English was always systemically capable of expressing the Irish-inspired messages; but no English speaker thought to do so until exposed to the Irish model.

3. A New Object of Explanation

The shift of meaning from object of explanation to *means* of explanation is part

and parcel of another shift. In generative linguistic theory, language use is, at best, a source of supplementary evidence for linguistic structure, occupying a distant second place to introspective judgments of grammaticality, structure and sentence meaning. By contrast, the analytical point of departure in Columbia School is the observable acoustic and graphic output of language users. This is Columbia School's true *explanandum*, the data to which linguistic meaning stands in an explanatory relation. Columbia School starts here not because of an interest in language use *per se*, but because it sees in sound (and graphic configuration) a range of phenomena that is truly *pre-theoretical;* that is to say, it can be observed and described without problematic assumptions about the nature of linguistic structure or, more importantly, without assuming in advance the identity of linguistic categories.

Why such interest in aspiring to an objectivity more appropriate perhaps to the physical sciences? Ricardo Otheguy lays the groundwork for this epistemologically austere position. In brief, Columbia School adopts Saussure's radical anti-nomenclaturism on the issue of linguistic categories, a position that represents a fundamental departure from the western tradition of language study. According to Saussure, language is not a set of names for objects and states in the physical world (meaning as reference*)*, nor a set of names for the categories and processes of thought (language as cognition*);* nor do the categories of language derive from philosophy or logic. If linguistic categories had such external motivation, then one could turn to the physical sciences, psychology, philosophy or logic for help in their discovery, as the western tradition has in fact done. But those venues are not open, says Saussure, because language is a system unto itself. The categories of linguistic meaning and structure are both language internal and language particular; they do not come from outside. It follows, they must be discovered by linguistic principles alone.

For Saussure, that linguistic principle was the linguistic sign, a fixed association between a *signifiant* and a *signifié*. In confronting a language, one does not know in advance what its structural, grammatical, morphological and semantic categories will be; all one has to go on is the expectation of a regular pairing of form and meaning. In practice, this means the analyst must wipe his mental slate clean, setting aside all the familiar categories of the grammatical tradition – subject, predicate, complement, direct and indirect object, noun, verb, adjective, agent, patient, recipient, past, present, future – and look for

categories that stand in the most regular relation to form. The traditional structural and semantic categories are not ruled out of consideration; they just have no privileged status. They can be entertained as hypotheses, but must earn their way as true linguistic categories by testing out successfully.

Charles Ruhl shows what emerges in the area of lexical meaning when *a priori* notions are set aside. Dictionaries typically offer multiple definitions for a word, definitions that are often justified in terms of *a priori* conceptual distinctions borrowed from classical rhetoric and philosophy: literal versus metaphorical; concrete versus abstract; basic versus extended. Interestingly, this tradition dates from the earliest dictionaries, which were bilingual, and which were explicitly designed as practical aids to translation, not as scientific contributions to linguistic theory. In place of the dictionary tradition, Ruhl proposes the *Comprehensiveness Principle*: "The measure of a word's semantic contribution is not accuracy (in a single context) but comprehensiveness (in all contexts)". The difference between this sign-based principle and the dictionary tradition is that between monosemic- and polysemic-oriented research.

Polysemy initially appears more credible than monosemy, concedes Ruhl; monosemy seems implausible, belied by the facts. Yet it is polysemy, he argues, that ignores the facts. Polysemy seems undeniable when the database is restricted. For example, **breaking** *sticks* and **breaking** *in a new man* would seem to mandate two distinct semantic values for *break*. Ruhl then augments the database with forty-nine examples gathered from natural language use. These create a seamless interpretive continuum between *breaking sticks* and *breaking in a new man*, a continuum that now belies the idea of two discrete "senses" of *break*. Ruhl then delivers the goods and presents a monosemic analysis of *break*, supported by an additional eighty-three examples. The conclusion is that the polysemy approach misattributes to the word in question information that in actuality is coming from the surrounding context.

Joseph Davis shows how a Saussurean skepticism about traditional categories plays out in his study of pronominal reference in Italian. The third-person pronouns *ess+, loro* and *sé* are traditionally characterized as a disjunctive demonstrative, a reciprocal and a reflexive respectively. But Davis finds these descriptions do not match the forms' actual use. Most striking is that all three are employed for *both* reflexive and non-reflexive reference. In fact, a broad survey shows that their use in texts in no way consistently corresponds

with the categories subject, predicate, reflexive, demonstrative, anaphor or pronominal. Davis thus abandons these as useful grammatical categories and searches for other ways to describe how the three pronouns differ. He finds the answer in the notion of information load. Each of the pronouns signals a different amount of grammatical information (gender and number) about its referent. Davis then shows that writers choose among the three pronouns by assessing how much information a reader needs to find the referent of the pronoun in the context.

Walter Hirtle's analysis of English *do* provides a third example of setting aside traditional distinctions, this time syntactic. In this exploratory study, Hirtle searches for the semantic unity underlying the use of *do* in its main verb capacity, *Sally **did** the dishes*, its suppletive capacity, *We asked Harry to make dinner and he **did** so*, and its auxiliary capacity, *Harry **did** make the dinner*. Hirtle sees in these three uses a progression of semantic specification. As an auxiliary, *do* simply indicates duration and nothing more; as a suppletive, *do* indicates both duration and change; as a main verb, where it contrasts with *make, do* indicates, duration, change and transitivity, though unmarked for result. The problem here is how to capture all this while still maintaining the essential unity of the sign.

Hirtle's Guillaumist framework allows him to do so by means of a process conception of linguistic meaning. A sign has a potential meaning, together with operative conditions that realize different portions of it in discourse. This potential meaning is more than just a semantic common denominator of all its uses. Rather, it is semantic trajectory, structured from minimal to maximal specification. In selecting a sign, a speaker mentally moves through this trajectory and stops at the point corresponding to the degree of actualization that represents the experience he has in mind. This process, called ideogenesis, yields the various actual meanings of a word observable in discourse.

4.   Non-Random Distribution of Forms

When the distribution of linguistic forms in authentic discourse becomes the explicit object of explanation, the procedures of data collection and analysis assume a heightened importance. Early in Columbia School tradition, the chi-

square test of statistical significance became established as a regular feature in
the presentation of quantitative data. In a second paper, Joseph Davis offers an
insightful rethinking of the role of statistics in Columbia School analyses and
argues that chi-square has often been inappropriately used. The barrier to the
use of chi-square is the lack of statistical independence among observations in
connected text. Moreover, statistical inference in general is both inappropriate
and unnecessary because Columbia School envisions no ideal *langue* or
linguistic community, no population from which a random sample of linguistic
tokens might be drawn. Accounting for observations in selected texts is an end
in itself, and there is no need to claim greater generality, since the demographic
prevalence of a linguistic feature is typically of little or no interest in these
studies. Davis concludes by offering a statistical test that might be used when
researchers do have some reason to infer generality from a sample to a
population of texts (e.g., a genre).

5.   Phonology

A second basic analytical problem confronts the sign-based linguist, namely
the phonological composition of the linguistic sign itself. This area of research,
traditionally known as phonotactics, parallels the grammatical analyses in this
volume in that both address, broadly speaking, the problem of asymmetrical
distribution. Grammatical analysis addresses the problem of the distribution of
linguistic signs in texts, and phonological analysis addresses the problem of the
frequency and arrangement of phonological units within the linguistic sign. But
the explanatory factors are more complex. In grammar the primary explanatory
mechanism is communicative: a linguistic meaning and its semantic
contribution to messages. Phonology appeals to three additional factors: the
physiology of the vocal tract, acoustics, and people's tendency to do things in
the easiest way.

 Yishai Tobin's study is an illuminating example of this functional
phonology. It addresses the fact that in the world's languages some word-
initial consonant clusters are very frequent while others are rare or non-
existent. For example, many one-syllable words in English begin with *tr* and
sl, while only three begin with *sf* and none begin with *tl* or *sr.* Tobin argues
that such synchronic facts about signals are the result of diachronic pressures

operating on speakers' selection of lexical items. The pressures involve the difficulty of articulating the individual phonemes in a language and the difficulty of articulating sequences of the phonemes. These principles act over time to shape the inventory of linguistic morphology in the language just as natural selection shapes the morphology of a species, gradually weeding out specimens that are functionally less advantageous. In language, these pressures gradually weed out hard-to-pronounce words by disfavoring their selection in speech whenever lexical alternatives exist.

In biological evolution, no morphological feature is inherently advantageous; thick fur is a blessing in the arctic and a burden at the equator. In the case of phonological evolution on the other hand, articulatory advantage is intrinsic because the governing principles reflect the universal physiology of the human vocal tract and people's fine motor skills. Tobin invokes three such principles in accounting for consonant clusters: (a) a sequence of similar articulatory gestures is easier to pronounce than a sequence of dissimilar gestures; (b) each additional articulator required to produce a sound increases its difficulty; (c) reuse of the same articulator in proximate phonetic environments is difficult. The principles produce predictions that can legitimately be tested on any language whatsoever. Tobin does so in impressive detail. Data are presented from thirty Indo-European languages, three Semitic languages, three Finno-Ugric languages and one Caucasian language. All three principles are consistently supported, though with differing degrees of strength. The occasional failures are due, Tobin argues, to sound change that has occurred too recently for the evolutionary pressures to do their work.

Abdul Azim extends the explanatory scope of Tobin's functional phonology in an examination of aspiration in Urdu. Urdu is renowned for its rich inventory of stop consonants. In addition to the familiar p, t and k phonemes, Urdu has a retroflex t, a palatal c and a post-dorsum q. All but the last have aspirated versions (p^h, t^h, t^h, c^h, k^h) and voiced versions (b, d, d, j, g). Finally the voiced series itself has an aspirated version (b^h, d^h, d^h, j^h, g^h), a feature unique to the Indic family of languages. The four sets of stops differ strikingly in both their position within the morpheme and their frequency throughout the lexicon. These differences are shown to scale in ways that match their incremental articulatory complexity and their different functional loads at the beginning and end of words.

Azim then carries the study to a higher level by moving from the asymmetric distribution of phonemes to the systemic and typological asymmetry of the phonemic inventory itself. This asymmetry is manifest in two anomalous facts. First, Urdu has no garden variety *h* phoneme – simple aspiration – despite using aspiration as a "complicating factor" in two sets of stops; this is like a national cuisine that made liberal use of wine in its sauces, but never served a glass of wine with a meal. Second, Urdu does have a *voiced h*, an articulation that is both typologically rare and articulatorily unnatural. These two facts gainsay any simple relation between ease and frequency in phonological structure. Here Azim appeals to "ease" in a different guise, to wit, ease of acquisition, but now played out systemically. In brief, a phoneme whose acquisition provides training for the acquisition of *other* phonemes enjoys a favored ecological niche. Azim then offers an explanation for the absence of a simple *h* that is a *tour de force* of non-causal, systemic, functional reasoning.

Mark Elson addresses the issue of morphological change in verbal paradigms, a problem that is also, broadly speaking, distributional. When speakers reshape a paradigm, replacing an existing form with a newly created form, what are they doing? Are they restructuring it in terms of its constituent lexical and grammatical morphemes? Or are they simply choosing one already-inflected member as base and then re-inflecting it? In other words, are they operating at the morpheme or the word level? Elson first takes issue with Bybee and Brewer's hypothesis (1980) that certain forms of verbal paradigms are autonomous, i.e., may, in their entirety, serve as the base for generation of other forms. Bybee claims that the third person singular is the most autonomous, followed by the first person singular, and second person singular, explaining this ranking by establishing a hierarchy of autonomy based on three factors: semantic complexity, opacity and frequency of occurrence.

Elson proposes an alternative view of paradigmatic innovation, one more in keeping with a sign-based framework. He argues that, under certain circumstances a member of a verbal paradigm may be *under*-specified grammatically; that is to say, it expresses less than the totality of the grammatical meaning associated with it. In such a case, the expressive potential of the under-specified form encompasses that of the more specified forms, thus making it semantically suitable to serve as a word-level base for formal innovation. For example, a form dubbed 'first person, singular' might be specified, in actuality, for only 'first person'. Such a form is compatible with its use as a first person

plural, and therefore it may serve as a word-level base for a new first person plural form. Elson's proposal, though in the separate area of morphological change, starts from the same premise as the synchronic grammatical analyses of Contini-Morava, Davis, Gorup, Hirtle and Huffman: namely, that the distribution of a form is a function of the meaning it bears.

6. The Future of Minimalist Linguistics in a Maximalist World

Since Columbia School exemplifies a minority position within linguistics, the question of survival naturally arises, and in the final paper of this volume Robert Kirsner contemplates that question. What future can this minimalist approach to language expect in a world of maximalist theories? He begins by articulating features that distinguish it from the competition. A major strength is its straightforward approach to meaning and to the essential problem of delineating the actual units of language. Its hard-nosed focus on these issues contrasts with the imprecision of many other functionalist approaches.

However, what appear as strengths to its practitioners strike many in the larger community as defects, and Kirsner examines the major criticisms of Columbia School, in particular those coming from linguists working in the framework of cognitive grammar, in considerable detail. The fundamental criticism has been its penchant for postulating relatively sparse meanings for linguistic forms and then appealing to pragmatic factors to explain how such sparse meanings nonetheless communicate concrete messages. The approach is said to be too reductionist and the analyses ultimately unconvincing. A related criticism is that of analytical control. There is no independent information from psychology to provide any control as to what is a reasonable inferential gap between meaning and message. Without solid psychological guidance, an analysis might attribute too much to inference and build too little into the meanings proposed. "[It is] difficult," writes Kirsner, "to prove that any single analysis is correct unless one could show that the meanings have enough substance, enough 'oomph' to them, to help the hearer actually arrive at or 'get' the message."

Kirsner responds to a number of these criticisms. He points out that the lack of independent psychological confirmation is not a shortcoming unique to Columbia School. Chomskyan principles of universal grammar and the

semantic networks of cognitive grammar are in as much need of independent psychological support as are Columbia School meanings. All linguistic theories to date are theories of knowledge, not neurological structure. As for the suspicion that Columbia School meanings are too sparse for successful communication, one can equally well ask whether the massive polysemy that results from positing more precise meanings is too complex for successful communication. Finally, the objection to sparse meanings begs the question of the essential nature of language. Columbia-style meanings look unsatisfactory when language is assumed to be a self-contained representational system of literal sentence meaning. If, on the other hand, language is a communicative tool that functions in concert with encyclopedic knowledge, past experience, and the larger linguistic and extralinguistic context, then Columbia-style meanings look just about right.

To Kirsner's discussion we would like to add an observation of our own. The fundamental reason for Columbia School's lack of appeal in the linguistic marketplace is the modesty of its analytical goals, its unique and restricted object of explanation. A Columbia School analysis offers an explanation of linguistic *behavior*, namely, the speaker's choice of linguistic forms. The meanings it posits are those that best explain the choice of particular forms. They are not concomitantly part of an explanation of the hearer's inference of messages – or the communicative process itself – for the principled reason that the locus of those phenomena is the human mind in general and its contingent relation to its past and present experience.

But while Columbia School is offering a theory of linguistic behavior, it is not *behaviorist*. Far from shying away from unobservable mental entities, abstract linguistic meanings are its core explanatory constructs. And rather than seeking the cause of linguistic behavior in the external settings of language use, it finds the cause in a speaker's stimulus-free desire to communicate some particular message. Admittedly, a theory of linguistic behavior is unfashionable in the Chomskyan era. Yet it is also a strength because it provides a basis for analytical control in linguistic analysis which psychology has so far failed to provide. All analytical decisions trace back to facts of language use, not guesses about how the mind works.

What, then, can Columbia School contribute to linguistics when it is such a minority position? The answer, says Kirsner, lies in its therapeutic function and in its exquisite empirical methodology. By proposing analytical problems

and solutions which could simply not be conceived in other frameworks, it serves as a corrective force in the field as a whole. Columbia School's skepticism towards all received categories of analysis forces scholars who take such categories for granted to reconsider, however unwillingly, exactly what they are doing. Finally, its concern with both qualitative and quantitative validation liberates the linguist from the conceptual and empirical shackles of the sentence-based perspective still held by most other schools of linguistics.

References

Contini-Morava, Ellen and Barbara Sussman Goldberg (eds). 1995. *Meaning as Explanation: Advances in linguistic sign theory.* Berlin: Mouton de Gruyter.
Contini-Morava, Ellen. 1995. "Introduction: On linguistic sign theory." In E. Contini-Morava and B. S. Goldberg (eds), 1-39.
Bybee, Joan and M. Brewer. 1980. "Explanation in Morphophonemics: Changes in Provençal and Spanish preterite forms." *Lingua* 52(3/4): 201-242.