SEMANTIC THEORY AND LEXICOGRAPHY

LEONARD E. NEWELL

Summer Institute of Linguistics

0. INTRODUCTION

The place of semantics in linguistic theory has been in question from the very beginning of the development of structural linguistics during the present century. Structural linguists have placed a high premium on rigor in their development of the science of linguistics, and partially as a result of this, have tended to regard the vast area of meaning as beyond the reach of the structural approach. Many have, in fact, rejected semantics outright, excluding it from consideration within the linguistic system. Hockett, for example, says, "the linguistic system of a language does not include the semantics. The system is abstract, it is a signalling system, and as soon as we study semantics we are no longer studying language but the semantic system associated with language."[1] Others view semantics as being within the province of linguistics but have done little to incorporate it into their models. Chomsky, for example, states, "Some of the features [of a lexical entry] are semantic features. These too [i.e. like phonological features] are drawn from a universal 'alphabet', but little is known about this today, and nothing has been said about it here. We call a feature 'semantic' if it is not mentioned in any syntactic rule, thus begging the question of whether semantics is involved in syntax."[2] Katz, Fodor and Postal have done important work in developing semantic theory but as recent as April, 1970 Wallace Chafe, in his *Semantically Based Sketch of Onondaga*, complained "there is no satisfactory available model of what semantic structure is like."[3] He continues "I do not believe that it can be described in terms of the 'phrase structure' model which has been used for 'deep' structure, and I have no choice but to present Onondaga in terms that will seem novel to the reader." Tagmemic theory has its lexical hierarchy which is intended to integrate semantics into linguistic theory but after twenty years tagmemicists are still working almost exclusively in the grammatical and phonological hierarchies.[4] Similarly, Lamb's development of Stratificational Grammar includes the field of semantics within the model with its sememic stratum (actually two strata as outlined by Lamb, 1966). However, Lamb has done little more than point in the direction of semantics and has provided some rudimentary tools with which to begin. He has not in any detailed way described what semantic structure is like. Probably the most detailed work in developing semantic

---

[1] C. F. Hockett, Language in Culture, ed. H. Hoijer, Chicago, 1954, p. 152.

[2] N. Chomsky, Aspects of a Theory of Syntax, Cambridge, 1965, p. 142.

[3] W. Chafe, A Semantically Based Sketch of Onondaga, Supplement to IJAL, p. 3 (April 1970).

[4] R. Longacre, D. L. Ballard and R. Conrad, members of the Summer Institute of Linguistics, worked on the development of Tagmemic theory to handle "deep and surface grammar" during a workshop in New Guinea, 1970. The results of this study are in two articles: The Deep and Surface Grammar of Interclausal Relations, to appear in Foundations of Language, and More on the Deep and Surface Grammar of Interclausal Relations (unpublished manuscript).

theory has been done, not by linguists at all but by anthropologists. This is the so-called 'componential analysis' developed by Conklin, Frake, Goodenough, Lounsbury, Wallace, Atkins and others. While it is true that their work has not provided a complete model for semantic description, it has made one major contribution—it has demonstrated the useful-ness of recognizing semantic units (lexemes, sememes or whatever) that can be further analysed into constituent parts in ways very similar to the relationship between the phoneme and its constituents.

Although the prospects for a full-blown linguistic theory in the near future to ade-quately handle semantics are perhaps not too promising, we must get on with the job of making dictionaries. I do not share James Sledd's pessimism, expressed at the Conference on Lexicography held at Indiana University in 1960, in which he stated that the "lexi-cographer must humbly muddle through to his own solutions of his own problems."[5] Lexicographers must acquaint themselves with what has been done and is being done in the field of semantic theory and apply it, as possible, to lexicography. The following are a few suggestions of how some aspects of semantic theory, developed to date, might be applied to dictionary making.

1. IDENTIFICATION OF SEMANTIC UNITS

One of the primary considerations in making dictionaries is, of course, the unit to be chosen as the entry form. Some dictionaries are prepared without a clear notion of what kind of a form is being entered. Some lexicographers believe the morpheme is the eligi-ble unit for entry. Malone of John Hopkins University, for example, states "The main thing is to make the morpheme, not the word, our unit of determining what items to include in the inventory."[6] If it were true, as it was once claimed, that the morpheme is the union of a form and a meaning, then it would certainly be the unit to choose. But this concept of the morpheme has been generally abandoned by linguistis as too simple to adequately handle the relationship between form and function. It is obvious, for example, that constructions often have meanings only vaguely reflecting their constituent parts: English 'kick the bucket', 'a flash in the pan'; Samal[7] *taytayan tikus* 'upper plate of house wall' (*taytayan* 'bridge', Malay *tikus* 'rat'). Many constructions of this type involve, not phrases, but words derived by compounding, affixation, etc. Philippine languages are very rich in such constructions: Tagalog *bahay-bata* 'placenta' (*bahay* 'house', *bata* 'child')· Casiguran Dumagat *ebuked* 'more primitive negrito living deep in forest' (*e-* 'place of', *buked* 'forest'), *kabetwan* 'river bed' (*ka...an* 'location', *bito* 'stone'); Batolan Sambal *baboybali* 'domestic pig' (*baboy* 'pig', *bali* 'house'); Agta *ulolag* 'worm' (*ulag* 'worm', reduplication affix 'diminutive'); Batad Ifugao *himbaluy* 'husband and wife' (*hin-* 'unit', *baluy* 'house'), *hinqama* 'father and child' (*hin-* 'unit', *qama* 'father'), *hinqina* 'mother and child' (*hin-* 'unit', *qina* 'mother').

Long lists could be prepared for any one Philippine language of similar constructions. To file and give the meaning of the constituent morphemes of these forms would be quite inadequate to handle the meaning, since the meaning of the forms do not consist of

---

[5]J. Sledd, Comments on Problems in Lexicography, IJAL 28:147 (1962).

[6]K. Malone, Structural Linguistics and Bilingual Dictionaries, IJAL 28:113 (1962).

[7]Philippine language examples are from the field notes of the following members of the Sum-mer Institute of Linguistics: K. Pallesen (Samal), T. Headland (Casiguran Dumagat), C. Houck and H. Minot (Batolan Sambal), R. Mayfield (Agta), L. Newell (Batad Ifugao).

the sum total of the constituent parts. Thus it is necessary to recognize a semantic unit (lexeme, sememe or whatever) and this must be the entry form of the dictionary. But there must be strict control on the entry of constructions. It would be quite possible to clutter up a dictionary with constructions better handled in a grammatical description of the language. Grammarians have too long relegated to the dictionary embarrassing constructions that cast doubt on an otherwise neat and regular description. For this and other reasons, it seems quite imperative for the linguist to work both in the grammar and lexicon of a language simultaneously. Exactly where the lines between grammar and lexicon are to be drawn would be decided upon by analysis in both areas of language rather than on the basis of work in one area (either grammar or lexicography) as is so often the case.

2. COMPONENTS OF SEMANTIC UNITS

Perhaps the most important concept of semantic theory as it relates to dictionary making is that of subcomponents of meaning units. As mentioned earlier, anthropologists have developed the concept of semantic components with much success in dealing with limited lexical sets such as kinship terms, color, etc. In describing kinship relationships, for example, dimensions such as sex, seniority of generation, etc. are used to define lexical terms. I believe lexicographers can and must use similar methods to analyse and describe all lexical items of a given language even though the lexical sets with which they work may not be so easily identified as those chosen for study by anthropologists. It makes no more sense to attempt to analyse the meaning of a given lexemic unit in isolation than it does to attempt to isolate a phonological unit of a given language and analyse it without reference to the phonological system of that language. Contrast in minimal environment is a powerful analytical tool and can be used just as effectively in analysing the semology of a language as is done in the phonology. The following are two simple lexical sets of Batad Ifugao.[8] Each lexeme is listed, followed by letters indicating the shared and contrastive components (A, B, etc.) of each member of the set, and then by a tentative gloss reflecting the components indicated by the letters.

Components of the Lexical Set:

 A open.

 B degree of opening: B_1 little, B_2 wide.

 C permanence of opening: C_1 momentary, C_2 permanent.

 D plane of motion: D_1 horizontal motion, D_2 vertical motion.

Lexical Set:

bughul	A ·	B_2	· D_1	'opens wide in a horizontal motion'
loqwab	A ·	B_2	· D_2	'opens wide in a vertical motion'
t̄iab	A ·	B_1	. C_1	'opens a little, momentarily'
toŋabel	A ·	B_1	· C_2	'opens a little, permanently'

Components of the Lexical Set:

 A spit.

 B distance: B_1 far, B_2 near.

 C consistency: C_1 liquid, C_2 solid.

---

[8]For an initial attempt to define lexemes in terms of components, see my Batad Ifugao Vocabulary, New Haven, 1968.

Lexical Set:

buldid	A ·	B_1		'spits far'
tukpa	A ·	B_2	· C_1	'spits near, of a liquid'
bula	A ·	B_2	· C_2	'spits near, of a solid'

Following some such analysis as that above, relatively rigorous definitions can be written for each lexeme that will indicate both the membership by each lexeme within a set (shared components) and also their contrastive lexical constituents.

By defining a lexical form in terms of components, it is possible to assure that the definition includes the sufficient and necessary lexical components in order for that form to denote and no more. Definitions of dictionary entries tend to fall into one of two categories—either they are verbose and vague, highly flavored by a number of synonymous terms, or they are little more than word lists with one word glosses. The constant cry of linguists and lexicographers is for precision in describing semantic units. Chomsky speaks of semantic features being drawn from a universal 'alphabet'. Weinreich says, "Ideally we might wish for an 'absolute' metalanguage which is entirely independent of the object language or any natural language.[9] Goodenough (1956:209) says, "I am convinced that further development of a notation for sememes [i.e. symbols for lexeme components] will open the way for . . . rigorous analysis of the content of conceptual systems." It is not, of course, practical for the lexicographer to express meaning in terms of a metalanguage, but what can be done is to carefully control the vocabulary used. If English is being used, for example, as the language to define lexemes in language X, then the English vocabulary used should, as much as possible, be single component lexemes. E.g. Batad Ifugao *toqol* 'look intently' (not 'stare'), *yuqgud* 'stooping posture' (not 'slouch'). English lexemes such as uncle, grandfather, gallop, wail, stallion, etc. would be avoided in defining lexemes in language X.

3. COLLOCATION

The theoretical approach described above is an intensional approach to meaning. That is, the lexeme is conceived of as a bundle of distinctive features consisting of the necessary and sufficient components to define the lexeme. But the definition of a lexeme involves more than a description of its distinctive features. It is also necessary to describe its collocational range and restrictions. Unfortunately, however, semantic theory has not been developed to date to adequately handle collocation. Anthropologists have given us practically no help and linguists have done little more than to suggest, Lamb expresses it, that language has a semotactic pattern. Martin Joos (1958:138) has demonstrated that semantic collocation is indeed patterned. Yet his methods can hardly be applied in any quantitative way and this, of course, is absolutely necessary for the lexicographer because of the sheer bulk of material with which he works. Nevertheless, a good dictionary will, for each entry, include both a definition and a statement about collocational restrictions. For the present, we must do the best we can in describing the latter, making use of whatever pieces and bits of information that may be available to us from linguists and others.

---

[9]U. Weinreich, Lexicographic Definitions in Descriptive Semantics, IJAL 28:37 (1962).

[10]J. Sledd, op. cit., p. 147.

4. CONCLUSIONS

During the Conference on Lexicography held at Indiana University referred to above, Professor Sledd made the statement that "structural linguistics, after thirty years, has produced neither a single big structural grammar of English, nor a single small structural history, nor a good English dictionary."[10] This inditement is painfully accurate. And if structural linguistics has failed for English, how much more is this true for other languages of the world. I believe this situation can and must be changed. If we who are involved in making dictionaries for Philippine languages will apply semantic theory to our work, there is little doubt that the results will be better dictionaries than those, in general, produced to date. Such dictionaries will, in turn, provide invaluable sources of information for linguists working to develop semantic theory.

REFERENCES

HAROLD C. CONKLIN. 1955. Hanunóo Color Categories. Southwestern Journal of Anthropology 11:339–344. Albuquerque.

_____. 1957. Ethnobotanical Problems in the Comparative Study of Folk Taxonomy. Paper read at the Ninth Pacific Science Congress, 1957. Bangkok.

_____. 1962. Lexicographical Treatment of Folk Taxonomies. In F. W. Householder and S. Saporta (eds.), Problems in Lexicography. Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication 21.

_____. 1964. Ethnogenealogical Method. In Ward H. Goodenough (ed.). Explorations in Cultural Anthropology: Essays in honor of George Peter Murdock. New York, McGraw-Hill.

J. A. FODOR and J. J. KATZ (eds.). 1964. The Structure of Language: Readings in the Philosophy of Language. Englewood Cliffs, N.J., Prentice-Hall.

CHARLES O. FRAKE. 1961. The Diagnosis of Disease Among the Subanun of Mindanao. American Anthropologist 63:113–132.

WARD H. GOODENOUGH. 1956. Componential Analysis and the Study of Meaning. Language 32:195–216.

MARTIN JOOS. 1958. Semology: A Linguistic Theory of Meaning. Studies in Linguistics 13.3–4:53–70. Buffalo.

J. J. KATZ and J. A. FODOR. 1964. The Structure of a Semantic Theory. Language 39: 170–210. Reprinted in Fodor and Katz (1964).

_____ and P. POSTAL. 1964. An Integrated Theory of Linguistic Descriptions. Cambridge, Mass.: M.I.T. Press.

SYDNEY M. LAMB. 1966. Outline of Stratificational Grammar. Georgetown University School of Languages and Linguistics. Washington, D.C., Georgetown University Press.

---

[10]J. Sledd, op. cit., p. 147.

FLOYD G. LOUNSBURY. 1956. A Semantic Analysis of the Pawnee Kinship Usage. Language 32:158–194.

_____. 1964a. The Structural Analysis of Kinship Semantics. Proceedings of the Ninth International Congress of Linguists. The Hague, Mouton.

_____. 1964b. A Formal Account of the Crow- and Omaha-type Kinship Terminologies. In Ward H. Goodenough (ed.), Explorations in Cultural Anthropology; Essays in honor of George Peter Murdock. New York, McGraw-Hill.

_____. 1965. Another View of the Trobriand Kinship Categories. In E. A. Hammel (ed.), Formal semantic analysis. American Anthropologist 67, No. 5, Pt. 2:142–185.

ANTHONY F. C. WALLACE and JOHN ATKINS. 1960. The Meaning of Kinship Terms. American Anthropologist 62:58–80.

CORRECTIONS

Below are corrections to be made in Volume 1, Number 1 (June 1970) of *The PJL*:

Article by Yao Shen

PAGE 127: The third line after the third quotation should read "difference between Keats, an intellectual poet-lover who was impatient with the lack, and Faustus, an intuitive poetic-lover, who was unaware of the lack."

PAGE 128: Paragraph 2, line 8 should read "the religious, while their emotive implication, identical with the poet and the religious is neglected.

PAGE 128: Add to REFERENCES: *Complete Prose Works of John Milton, Volume II, 1643–1648*, (edited by Ernest Sirluck), New Haven: Yale University Press.

Article by Isidore Dyen

REPLACE ALL *o* in Cebu words with *u* throughout the article.

PAGE 4, 1.2 up: Read "everywhere" for "exery-where".

PAGE 5, 1.8 down: Read "Tagalog *qa-ŋ ba · hay*" for "Tagalog *q-aŋ ba · hay*".

PAGE 5, 1.10 down: Read "*so ator*" for "*s-o ator*".

PAGE 5, 1.11 down: Read "*aa-ŋ ba · hav*" for "*qa · aŋ ba · hay*".

PAGE 5, 1.13 down: Read "*so ator*" for "*s-o ator*".

PAGE 6, 1.6 down: Read "*so*" for "*s-o*".

PAGE 8, 1.3 down: Read "*n ~ ni*" for "*n – ni*".

PAGE 9, 1.9 down: Read "*in Maranao*" for "*on Maranao*".

PAGE 9, 1.15 down: (under Tag.) Read "*sa kanya*" for "*sa kaniya*".

PAGE 9, 1.20 down: Read "Tag. *qibigay ko sa kanya*" for "*qibigay so sa kaniya*".

PAGE 9, 1.15 up: Read "*ni – ni*" for "*ni – n*".

PAGE 9, 1.14 up: Eliminate entire line and instead read: Cebu *ku, mu* and Hil., Cebu *ta,* and except for the presence of Tag. *qa · kin* and

PAGE 9, 1.6 up: Read "*qa-ŋ ba · hay*" for "*qa ŋ ba · hay*".

PAGE 9, 1.5 up: Read "*qa-ŋ balay*" for "*qa ŋ balay*".

PAGE 9, 1.4 up: Read "*qa-ŋ balay*" for "*qa ŋ balay*".

PAGE 10, 1.4 down: Read "Tag. *ba · hay ko*" for "Tag. *ha · hay ko*".

PAGE 10, 1.7 down: Read "*qitu qa-ŋ qa · kin.*" for "*qitu qa ŋ qa · kin.*"

PAGE 10, 1.8 down: Read "*qitu qa-ŋ qa · kon.*" for "*qitu qa ŋ qa · kon.*"

PAGE 10, 1.9 down: Read "*kini qa-ŋ qa · kuq.*" for "*kini qa ŋ qa · kuq.*"

ADD THE FOLLOWING REFERENCES:

Bloomfield, L. 1917. Tagalog Texts with Grammatical Analysis. Part 2: Grammatical Analysis. Urbana.

Dyen, I. 1965. A lexicostatistical classification of the Austronesian languages. IJAL Memoir 19.

Kauffman, J. 1939. Principles of Visayan grammar. Iloilo.

McKaughan, H. 1958. The inflection and syntax of Maranao verbs. Manila.

Wolff, J. U. 1966. Beginning Cebuano. New Haven.