

On the Semantics of Polysynthesis

Author(s): Emmon Bach

*Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals* (1993), pp. 361-368

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

---

*The Annual Proceedings of the Berkeley Linguistics Society* is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

## ON THE SEMANTICS OF POLYSYNTHESIS

Emmon Bach  
UMass (Amherst)  
BLS 2/13/93

I. Preliminaries.

The basic question I want to pursue here is this:

- a) What differences, if any, are there between the kinds of meanings that are expressed in the grammar of words and the grammar of phrases?

Or more succinctly: Where do languages express what? We will pursue this question in the context of a traditional typology that characterizes languages according to the number of morphemes that can go into individual words. We'll first explore a bit a number of ways in which languages (and subsystems of languages) can be "polysynthetic," and then narrow our focus to a manageable subpart of this large topic, looking at some of the kinds of meanings that are expressed in lexical and grammatical affixes in a few languages.

Let me first give the word to Edward Sapir:

An analytic language is one that either does not combine concepts into single words at all (Chinese) or does so economically (English, French). In an analytic language the sentence is always of prime importance, the word is of minor interest. In a synthetic language (Latin, Arabic, Finnish) the concepts cluster more thickly, the words are more richly chambered but there is a tendency, on the whole, to keep the range of concrete significance in a single word down to a moderate compass. A polysynthetic language, as its name implies, is more than ordinarily synthetic. The elaboration of the word is extreme. Concepts which we should never dream of treating in a subordinate fashion are symbolized by derivational affixes or "symbolic" changes in the radical element, while the more abstract notions, including the syntactic relations, may also be conveyed by the word. A polysynthetic language illustrates no principles that are not already exemplified in the more familiar synthetic languages. It is related to them very much as a synthetic language is related to our own analytic English. The three terms are purely quantitative -- and relative, that is, a language may be "analytic" from one standpoint, "synthetic" from another. I believe the terms are more useful in defining certain drifts than as absolute counters.

Edward Sapir (1921: p. 128 [paper bound edition])

We'd better replace "concept" or "idea" in such characterizations by "morpheme" or "morphological operation." Consider for example this Westcoast (Nootka, Nuuchahnulth) root and its gloss as given in Sapir and Swadesh (1939: 256):

1. √yah- 'to shout in a prescribed manner in the woods as a daily morning practice for a mother of twins for a year after their birth'

Although simple words formed from this root surely contain a lot of concepts, I doubt if Sapir would take this root to exemplify polysynthesis. In his

Language (141 f.) he discusses words like the following in the context of the typology I'm invoking, again from Westcoat (my retranscription):

2. ?inikwiłm'inih?isita `several small fires were burning in the house'

analysis: (forms from Sapir and Swadesh, 1939)

?ink(--) `fire; burning'

`-ił(.-),m. ... `in the house, on the floor'

-minh plural [incremental suffix]

-?is, -?ic- diminutive [incremental]

-(m)it [?] past [incremental]

-ma; -a (with past -(m)it)(. for variable vowel marking) 3p indic

Probably everyone would accept this example and the language from which it is taken as bona fide examples of polysynthesis.

So we have:

analytic - one morpheme or morphological operation per word

synthetic - more than one morpheme or morphological operation per word

polysynthetic - many morphemes or morphological operations per word

Given this general comparative attribute of the the analytic-synthetic dimension, there are still lots of different ways in which languages can exhibit polysynthesis. One way relates to the relative independence of the elements that go into a complex word. If the elements are all more or less free forms (or closely related to free forms, perhaps roots or stems) we have what is usually called compounding. People don't usually include compounding under the heading of polysynthesis. For example, Chinese is usually adduced as near the analytic extreme of the analytic - synthetic scale. Yet most words in Mandarin Chinese are synthetic, as a perusal of a modern dictionary will confirm. It's just that the pieces of the complex words are usually free forms or potential words, and the language has a minimal number of derivational affixes (just one, if I'm not wrong). So we have to amend the traditional definition of synthesis (and polysynthesis) and restrict its application to affixal or bound elements. For the rest of this paper I will exclude straight compounding from consideration and concentrate on derivational and inflectional synthesis, except for purposes of comparison.

Another important dimension rests on the distinction between inflectional or grammatical markings on the one hand and derivational or word formation processes on the other. Even if you don't accept a sharp difference in kind between these two sorts of affixations and other operations (as e.g. di Sciullo and Williams, 1987, do not) there is still a pretty clear difference between elements on one end and the other of a scale from grammatical to word-formational.

Finally, there is a different scale of comparison (and one that Sapir seems to appeal to in the very notion of polysynthesis): from the concrete and specific meanings to very abstract relations such as those encoded in typical grammatical paradigms. The few linguists that I know of who have tackled the basic question I started from -- where do languages express what? -- seem to have looked at the question mainly with an eye on things on the grammatical end of this scale (Bybee, 1985; Carlson, 1983).

For a space of possible meanings, I will take the structured universe of possible denotations standardly assumed in many model theoretic approaches, to wit:

3.       A: set of possible individuals;  
           BOOL: truth values { 0, 1 } (possibly supplemented by  
   some impossible objects);  
           W: set of possible worlds;  
           F: hierarchy of functions built out of the preceding.

As a matter of convenience, I will also assume a loosely categorial framework for our comparisons of the meanings of words and affixes. By way of illustration consider the following analysis of a short English sentence, which shows what I suppose to be the canonical external semantics of Nouns, Verbs, and Adjectives:

4.       Every                    little                dog                barks.  
           DET                    ADJECTIVE        NOUN             VERB  
           <<e,t>,<<e,t>,t>> <<e,t>,<e,t>> <e,t>       <<<e,t>,<<e,t>,t>>,t>

This example also serves to illustrate the Montagovian category system I'll follow for general meaning-types:

5.       e : entities  
           t : truth values  
           <b, a> : functions from b-type things to a-type things

(Mostly, I'll ignore the intensional aspects of our theme, to be coded into categories using the further symbol g either syncategorematically as in Montague, or as representing the further primitive category of possible worlds.) When I say "loosely" I mean that I will entertain the possibility that some functional categories are expressed as operations with possibly no direct segmental identity as pieces of an expression.

With these preliminaries given we can now consider some further questions:

- b) How are the categories of subword elements or operations related to the categories of phrases?

That is, are the sets of meanings of phrases and words a subset or superset of those of parts of words, or are they incomparable, and if so do they overlap? An example of an affixal meaning that is never phrasal might be furnished by the subject-object portmanteaus that occur in quite a few languages (e.g. the transitive pronominal affixes of Iroquoian languages). An example of a phrasal meaning type that is never encoded into a single affix might be the meaning of declarative sentences (say of category t, that of truth-values). As a part of this question, or closely related to it, is the question whether the relations among elements within a word show the same possibilities as those in higher, phrasal constructions.

Next, we might ask about semantic connections that might cross the word-phrase boundaries:

c) How are the interactions among the meanings of subparts of words and phrasal elements constrained?

For example, if there are scope-relevant elements within words to what extent can there be scopal interactions between these elements and elements at higher levels? Here also belong questions about words as anaphoric islands and so on. Note for example the limitations in English on the interpretations of sentences like (7):

6. Some animals are omnivorous. ( $\neq$  everything has some animal that eats it!)

Finally, let me mention a pervasive idea about (some) polysynthetic languages:

d) Are complex words in (some) polysynthetic languages just little (or not so little!) sentences?

This question is obviously related to previous ones. If the full range of structured sentence meanings can be encoded into words, then we would have to expect that the full range of meanings available for constructing sentences would be available below the word level. If only some limited set of such sentence meanings can be reproduced morphologically then we have to ask just what the limitations of word-internal meanings (or means for combining them) are that lead to the constraints.

## II. Some examples.

In the remainder of the time I have here I will present some exhibits of complex words and their parts. A number of my examples come from the North Wakashan language Haisla, spoken in Kitamaat, B.C. I wish to acknowledge gratefully the help that I have had from a number of members of that community over the years. I take all responsibility for misunderstandings and mistakes. Here's a general scheme to show the structure of words in Haisla:

Haisla: EXT(ROOT) DerAff\* GrAff<sup>n</sup> (n = 0, 1, 2, 3, 4)

Here, EXT refers to some operation (including the identity operation) on a root to form an extended stem, as the basis for derivational suffixation. Items in the last place -- grammatical affixes -- might also be called clitics. To give some idea of the relative size of these various types of elements, there are about 1400 - 1500 roots, around 500 derivational affixes, and some dozens of grammatical affixes.

Haisla is syntactically VSO, exclusively suffixing, and is a prime candidate for a language with no lexical contrast between nouns, verbs, and adjectives. Boas (1947) remarks about the related Wakashan languages (Kwakw'ala, Westcoast) and several other languages of the area on the relatively large number of lexical or derivational affixes.

To give some idea of the kinds of words that exist in Haisla, let's look at a few examples:

7. 'i'ksduqwia (s.v. in LR) : 'ik- -sdu -qi -a bald eagle

The root /'yk (in LR's underlying form) means 'good, clean, bright' etc.' +  
 -[s]du 'eye, color, appearance' (Boas 47.343: cf. 239--) round  
     opening, eye, color) (/s/ absent after /lh/)  
 these two yield 'iksdu 'bright-colored, white' (one of a whole bunch of  
     color words based on -sdu) +  
 -qi(a) (morphophonemic rounding) means 'head'

Here are a couple more items with this suffix:

8. lhaxwmeç'uaqiala have a headache

(-c'ua means 'inside', the stem means 'hurt'), but the combination -c'uaqi(a) seems to also mean 'head' and not just interior of head (as we might expect with 'hurt'), cf. mexc'uaqia 'punch (/mex-) sb. on the head. Note: in the spelling I'm using here, e = ø, many of these schwas are epenthetic.)

9. Xuc'eqia close hair cut (mallet head) (Xuc' mallet)

10. 'um'aseqia (have a) big head (LR)

(the final element is the 'terminal' -a that often rounds off words with no apparent meaning and under mysterious circumstances, cf. Boas, 1947.308 f. and elsewhere).

There is a formal contrast between a word like the last example and a phrase like the following (this is 'um'as -s hixt'i) with a 'connective' --i.e. syntactic -- -s).

11. 'u'm'ac hixt'i 'big head'

Note that there is no connection between the independent word for 'head' and the suffix -qi(a). In general this is the case for the derivational suffixes. Haisla has no compounds. Note further, however, that the original word for bald eagle that we started from is a typical kind of compound (bahuvrihi) in its meaning.

Words and sentences:

The next example shows that there are words that can express an entire sentential meaning.

12. keta'tlnugwutl(a) 'I'm gonna shoot you'  
 ket- 'shoot' -(a)tl Future, -nugw(a) 'I' -utl 'you'

Note that the last two elements are grammatical affixes or clitics. In general, there is complementary distribution between these pronominal clitics and full NP's. I conclude that Haisla is not a pronominal argument language in the sense of Jelinek, 1984.

Relations among formatives?

Wordbuilding in Haisla is accretional, in general. In accordance with what we might call "Frege's projection principle" (functions need arguments) and in the categorial framework I'm assuming, affixes will be functors and will be interpreted as functions:

Function (affix), argument (stem)

By function composition, many affixes are conventionally combined into complex affixes that can and do then acquire noncompositional or idiomatic meanings (compare -c'uaqi(a) 'head' above).

There has been a great deal of discussion of "incorporation" (note scare quotes) in recent generative literature (Baker, 1988; Sadock, 1991). Accordingly, research on the more special semantic/syntactic relations among the elements of complex words has laid heavy stress on argument structures. In Haisla, there is no incorporation and in general the affixes never seem to encode the primary arguments of stems. There is no incorporation, perhaps because there is no compounding. (See Sapir, 1911.)

The basic semantic relation that is encoded by lexical suffixes is that of an adjunct or optional modifier to its argument. Here are some examples of two of the popular kinds of affixes, locational and somatic.

Adjunctive relationships:	[=SUF : softening, voicing]
	[-!SUF : hardening, glottalizing]
=ilh 'inside, in house'	
=is 'on field or beach'	
-Ṙs 'in boat or vehicle'	
-bt 'in(to) hole'	
-!Ṙd '(on) behind, buttocks'	
-[x]sis 'on foot, lower leg'	
-Ṙdia 'on back'	
-Ṙina 'on shoulder or upper arm'	

As I try to indicate in the glosses, these items don't act as arguments of stems with verbal meanings (with an apparent exception I'll mention in a minute). If they did, then we would expect them to yield items with one less argument place when added to a stem (e.g. an intransitive from a transitive, for example). Instead they yield items with a narrower, more specific meaning of the same general type: e.g. men'eṘd (< √mex- and -!Ṙd) means to punch someone on the behind). I believe the apparent exceptions to this, where a stem seems to impart an adjectival specification to a body-part for example, are in fact (like the 'eagle' word mentioned above) bahuvrihis, gel't'eṘd doesn't mean 'long ass' but 'long-assed, tall.' This fact is sometimes obscured, I think, by translations like 'Long River' for Gelda'la (i.e. Kildala) a local place-name, which follow English rather than Haisla habits of nominal place naming.

There are functional (bound) roots, generally forming anchors for locational and somatic suffixes. We can think of these as type-lifted higher order functions taking fundamentally functional elements as arguments (just as in English, a word like treat wants to have a manner adverbial -- usually an adjunct -- as one of its arguments):

Functional (bound) roots:

√'u- locative, place where:  
'uṘwina' 'shoulder'  
'u'bta 'hole'

'u'qwia 'head of an inlet' (cf. -qi 'head' above)

One area of special interest in connection with the relations between the meanings of word-internal and word-external syntax is quantification. Since I want to deal with this question elsewhere, I'll confine myself here to a couple of observations.

Quantificational meanings are expressed in both roots and affixes, as these examples show:

y'ay'ugwau'lh rain constantly (y'u'gwa 'rain, be raining')  
 q'iutl to catch a lot (√q'i- much, many)

If words are just little sentences, we might expect the whole machinery of quantification to be expressible within word-grammar (including question formation and so on). It isn't, and I think this is generally true. (One way of looking at this within some theories would be to say that the maximal category in word grammar is VP or maybe IP, but never CP.) So there is no binding of variables within words (NB: I mean binder and bindee within a word), no WH- movement in words, no QR or whatever.

We have to be careful here: in general, words can contain all kinds of complex meanings. These meanings might include meanings that could be otherwise expressed by the formal machinery of quantification, but still we would not want to literally assimilate the words to this formal machinery (I assume). As an analogy: the word increase does not literally contain or involve the comparative constructions of English, even though it clearly involves ideas that could be expressed by comparative constructions: get, make bigger, more abundant.

I will close with some examples intended to show that the meanings of word-internal and word-external grammar are incomparable (although it's obvious that they overlap).

Here are some affixal meanings that are not possible meanings for phrasal constituents:

Lakhota etc. Instrumental causatives: 'cause to V by means of foot':  
 na- 'to cause with the foot x to V':  
 nawahomni 'I (wa-) turn it with my foot'  
 ka- '...by impact'  
 ya- ...'with the mouth' (cf. English etc. Passive)

Incorporated common nouns? (Type-lifted CN's combining by function composition with e.g. transitive verbs to make functions from determiners to intransitive verbs, cf. Baker, Sadock and others.)

What about the opposite case: are there supra-word meanings that are not possible as meanings of subword elements? Here's a possibility.

Speaking from a categorial viewpoint, I assume that all affixes are functors. It then follows that none of the real argument categories of natural language can be encoded into affixes: that is: t, e, and on some views whatever the correct semantic type of common nouns is. But this is a very theory-internal

constraint, and given the flexible sorts of categorial systems that seem to be indicated for natural languages, it's not clear that the constraint has much force. Clearly, there's a lot more to be done in the area of semantic research I have tried to exemplify in this paper.

### References.

- Baker, Mark C. 1988. Incorporation. Chicago: University of Chicago Press.
- Boas, Franz. 1947. Kwakiutl Grammar with a Glossary of the Suffixes. Edited by Helene Boas Yampolsky with the Collaboration of Zellig S. Harris. Transactions of the American Philosophical Society. N.S. Vol 37, Part 3, pp. 202[?]-377. Reprinted by AMS Press, New York.
- Bybee, Joan L. 1985. Morphology: a Study of the Relation between Meaning and Form. Amsterdam/Philadelphia: John Benjamins.
- Carlson, Greg. N. 1983. Marking constituents. In Frank Heny and Barry Richards, eds., Linguistic Categories: Auxiliaries and Related Puzzles (Dordrecht: Reidel).
- Hamp, Eric P. 1957/1963. A Glossary of American Technical Linguistic Usage 1925-1950. Utrecht/Antwerp: Spectrum.
- Di Sciullo, Anna-Marie and Edwin Williams. 1987. On the Definition of Word. Cambridge, Massachusetts: MIT Press.
- Jelinek, Eloise. 1984. Empty categories, case and configurationality. Natural Language and Linguistic Theory 2:39-76.
- Lincoln, Neville J. and John C. Rath. 1986. Phonology, Dictionary and Listing of Roots and Lexical Derivates of the Haisla Language of Kitlope and Kitimaat. (2 vols.) Ottawa: National Museum of Man Mercury Series. Canadian Ethnology Service: Paper No. 103. WAK HAI SLA
- Sadock, Jerrold M. 1991. Autolexical Syntax Chicago: University Chicago Press.
- Sapir, Edward. 1911. The problem of noun incorporation in American languages. American Anthropologist, n.s. 13: 250-282. [Collected Works V: 27-59.]
- Sapir, Edward. 1921. Language. (An Introduction to the Study of Speech) New York: Harcourt, Brace (Paperback reprint: Harvest books).
- Sapir, Edward and Morris Swadesh. 1939. Nootka Texts. Philadelphia: Linguistic Society of America.