Some More Impossible Words

EMMON BACH

University of Massachusetts, Amherst

As is well known, they [American Indian languages] are unusually
variant in form from each other, and it is perhaps not too much to say
that there is hardly a morphological type which is not illustrated in the
American field. [Edward Sapir Collected Works, V: 144, from a
previously unpublished note]

There is hardly a classificatory peculiarity which does not receive a
wealth of illumination from American Indian languages. It is safe to say
that no sound general treatment of language is possible without constant
recourse to these materials. [op.cit. p. 145]

"When I use a word." Humpty Dumpty said.... "It means just what
I choose it to mean -- neither more nor less."

1. Background
The title is ambiguous: 'some additional impossible words' -- 'some words that
are more impossible than some others.' I mean it as a blend of the two. The
'others' are the sort discussed in a recent exchange between Fodor and Lepore on
the one hand and Hale and Keyser on the other (iJ 30:3), which took off from
earlier papers by Hale and Keyser (see the references under these names). That
discussion circled around examples like these:

(1) *It cowed a calf. (= 'A cow had a calf.')
   *He shelved the books on.
   *The book shelved.

The former authors dispute and the latter defend the idea that such supposedly
impossible words or interpretations are evidence for complex derivations of some
words from syntactic or near-syntactic meaning structures.
The present paper is not in the first place about such examples, but the general question behind it is related:

(2) **What is the relation between the grammar of words and the grammar of phrases?**

(I hold that some of the kinds of meanings of word formations that are supposed to be excluded by explanations such as those of Hale and Keyser, can easily exist as meanings of sub word items and operations.)

In several papers (Bach, 1994, 1996) I have raised this question in the context of discussing languages with big words, ‘polysynthetic’ languages in one or the other older meanings of that term. The two older senses I have in mind are: ‘having unusually complex words’ and ‘expressing in a single word what can be only expressed by a sentence (in some other language of reference)’ (see Denny, 1989). My general point has been to try to undermine views that try to build word structures from phrasal syntax (or vice versa for all that). Let’s follow Swadesh’s happy terminology and speak of ‘internal’ and ‘external’ syntax (Swadesh, 1939).

One kind of argument goes to show that the two domains go their independent ways in various languages. A pair of languages can be similar in external syntax and very different in their internal syntax, or the reverse, with all combinations attested. In a previous paper (Bach, 1996), I cited pairs of languages that have very similar word structures and very different phrasal syntax: Haisla and other Wakashan languages compared with Yup’ik and other Eskimo-Aleut languages, as well as languages with globally similar phrasal syntax and quite different word structures: verb initial languages like Haisla and Nisga’a (Tsimshianic) or verb-final languages like Tlingit and other Na-Dene languages and Japanese or Turkish. These comparisons show that whatever else the languages might have in common, the details of how words are constructed and how phrases are constructed can vary quite independently.

Another kind of argument takes off from this question:

(3) **What is the relationship between the meanings of the elements of word-internal and word-external syntax?**

If the kinds of meanings that are found in the two domains are incomparable -- that is if there are types of meanings that are found in one domain but not in the other -- then that fact would serve further to undermine theories that hold that one domain mirrors the other, or that they are just two ends of a kind of continuum.

So to take one extreme view: if there are languages in which complex words are just sentences (on one of those views of ‘polysynthesis’), then we would expect the categories and their semantic values to be the same within words as outside of words.

Here I will pursue further this second question, and draw upon materials from some native languages of North America. I follow the sage advice of Edward
Sapir who admonished us that serious discussions of morphology could hardly ignore the rich data of American languages.

There are two subquestions that go into Question (3): one has to do with the meanings of the basic and derived elements in the domain in question, the other has to do with the kinds of relationships into which these meanings can enter in the two domains. (Another big set of questions has to do with the ways in which internal and external syntax are related. I won’t have very much to say about this last kind of question here.)

An even bigger question behind my musings is about language diversity:

(4) How come languages are as different as they are?

This question contrasts with its mate, which has been almost exclusively the center of attention in a lot of linguistics in the last decades:

(5) How come languages are as similar as they are?

Both questions presuppose that we have some sensible way of characterizing similarity and difference.

(6) General thesis: meanings of morphemes that make up parts of words are different from meanings of whole words (in general).

More precisely, the kinds of meanings that are associated with the elements of the phrasal syntax and those expressed by elements inside of words are incomparable (in the set-theoretic sense). There is overlap: some meanings can be expressed either in words or in parts of words, but there are residues in both directions.

Here is an initial example (from previous papers): Instrumental causatives, widespread in several North American language families: Haida (isolate), Siouan, Algonquian: these are affixes which form transitive verbs from intransitives, and add a notion of instrument or manner to a causative result: to cause to V by means of the hand, mouth, in a glancing manner, etc., as in these examples from Lakota:

(7) gniya ‘roll around’
kagmiya ‘roll (cause by striking)’
nagmiya ‘roll (cause with foot)’ (cf. Boas Deloria, 1941)

A sort of ugly paraphrase in broken English might be: strike-roll, kick-roll. These paraphrases pertain to the resultant words. Try to paraphrase the meaning that could be extracted from the prefixes themselves, na- and ka- and I think you will see what I mean. These meanings are even hard to render exactly, since we don’t have direct ways of expressing lambda-abstraction or the like in ordinary
language: ‘that function $f$ from predicates to relations such that ... More on such examples below.

What I will try to show is that, in derivational processes, any meaning that can be used to make an interesting class of words (or subwords) can be the meaning of a word internal morpheme or process (operation). This conclusion is not very interesting for theories of Constraints, but it is highly interesting for theories of Freedoms.

Preliminarily, it is necessary to sort out kinds of words, and kinds of formatives: e.g. open class words versus functional items, inflective versus derivational formatives.

For concreteness, I will assume a simple categorial framework with a standard structure of (model-theoretic) interpretations. The particular choice is not essential to my argument, but what is crucial is the assumption that there is a constrained mapping from elements, categories, rules, operations to their meanings. For a structure of meanings I will take a standard model theoretic interpretation: a domain of individuals, truth-values, worlds (and possibly times) and all functions built of these elements.

2. Exhibits

2.1. Instrumental Causatives
By ‘instrumental causative’ I mean this: an item or operation that takes as input (argument) an intransitive verb (or adjective) and produces an item which is a transitive verb and has (as defining characteristic) an instance of the following meaning rule:

(8) A. Let $a$ be the meaning of the intransitive verb that is the argument of the function, and $b$ be the particular instrument associated with the item or rule, then the resultant transitive verb denotes that function $f$ such that for any individual $x$ $f(x)$ denotes the set $M$ of individuals $y$ such that $y$ is a member of $M$ [at a time, world, etc.] iff by means of the instrument $b$, $y$ causes $x$ to be in the set $a$ [at that time, world, etc.].

In quasi-English: ‘$y$ uses $b$ to make $x$ (be) $a$’.

This is just a rough indication of the meaning of the item in question: an affix for making causatives with some particular instrument. A similar meaning rule would be needed for a causative for a manner (for example, ‘with a glancing blow’ or the like).

I’ll leave it as an exercise to the reader to try to derive these sorts of words from (quasi-) syntactic structures. Two candidate structures come to mind, corresponding to the paraphrases I’ve been using:

(9) $x$ causes $y$ to roll by means of $x$’s foot
$x$ uses $x$’s foot to make $y$ roll
Some More Impossible Words

These instrumental causatives pop up in a number of different languages, as noted already, including

(10) Haida Instrumental Causative Prefixes (Lawrence, 1977):

- kʼutʼahl die
- kʼakʼutʼahl kill by beating
- kikʼutʼahl kill by poking
- kukʼutʼahl kill by pushing or hitting with fist
- thakʼutʼahl let die (general causative)

Another interesting parallel is an affix in Misantla Totonac, which has the effect of composing a transitive verb with an instrument seeking function:

(11) Misantla Totonac Instrumental affix (MacKay, 1999):

- /maq-nii/ kill X
- /lii-maq-nii/ kill X using Y
- /kaak-spit/ peel X
- /lii-kaak-spit/ peel X with Y

Let’s note that there are perfectly ordinary lexical items that we can think of, in for example English, that have meanings corresponding to the resultant lexical items: my Webster’s Collegiate Dictionary includes this idefinition:

(12) strangle: to choke to death by compressing the throat with something (as a hand or a rope)

This meaning is quite complex, including the notions of dying, causation, performing an action on a certain body part or area, and an instrument. The body-part or body-area meaning, by the way, is quite reminiscent, in the way it is incorporated [!!] into the whole meaning of the word, of the well-known body-part suffixes of some American languages.

My intuition about the meaning of the instrumental causative affixes of Lakota and other languages, is that are not possible candidates for meanings of lexical words. My general point again is:

Any interesting component of the meaning of a word, is a candidate for the meaning of a an affix or operation that is used for making up lexical words or items. What is ‘interesting’ enough to generalize over will obviously be dependent not only on general cognitive properties of people but also on particulars of the lives that they lead. What is available for formal operations or relations will similarly be determined by general properties of Language and by parochial properties of individual languages.
Instrumental causatives occur in a number of American languages -- besides Siouan languages, unrelated Haida, and Algonquian are ones that I have encountered. The formations are sufficiently striking to have formed an important part of Chafe's arguments (1976) for the genetic relatedness of Siouan and Caddoan.

I claim that no words or non-affixal lexical items have just the meaning of the instrumental causative affixes of Lakota or Haida. We can make up a syntactic constituent with such meanings only by using rather outre means, such as occur in weird right node raised phrases:

(13) **I soothed and used a pacifier to make calm -- my baby.**

Here the underlined portion of the sentence is categorially equivalent to the instrumental affixes we have been looking at.

2.2. **Phrasal and non-phrasal categories.**

By definition, the argument and resultant categories of the phrasal syntax are the categories that appear in the operation of the phrasal syntax. It is often assumed that these are chosen from a universal set, and include such categories as these: S, N, V, A, DP (NP), VP and so on. It is controversial whether all languages make the same choices, or whether such lists (under various theoretical frameworks) constitute a closed set of options. Again following a categorial framework, I will assume a definition of possible categories based on a primitive set, say **w, t, e** and all functor categories built up from these. The syntax to semantics map assigns the denotations:

(14) **Syntactic/Semantic assumptions:**

<table>
<thead>
<tr>
<th>w</th>
<th>worlds</th>
</tr>
</thead>
<tbody>
<tr>
<td>e</td>
<td>individuals</td>
</tr>
<tr>
<td>t</td>
<td>truth-values</td>
</tr>
<tr>
<td>b/a, a\b</td>
<td>(&lt;a, b&gt;) (functors: functions from a-type denotations to b-type denotations)</td>
</tr>
</tbody>
</table>

(NB: 'functor' does not mean 'functional category' in the sense of various non-categorial frameworks). In some version of X-bar theory we would have the X\(\theta\) lexical categories, the (possibly) non-lexical phrasal categories, and the various projections of the X\(\theta\) lexical categories. Note that in the categorial setup, the categories incorporate the sort of information given by subcategorization in X-bar theories. So, for example, a transitive verb is categorized in X-bar systems by the feature specifications [V -N] (say) and also by a subcategorization feature indicating that it takes a noun-phrase complement, in categorial terms the subcategorization is the primary property and the verb will be classed as (say) IV/NP or the like.
Some More Impossible Words

In addition to this general category and type theory, I assume that there is a
crosscutting distinction between basic (lexical) members of various categories and
(possibly) corresponding phrasal categories, as in the manner of Montague's PTQ
(Montague, 1973). And following Dowty (1979), I assume a sharp difference
between lexical rules, which extend the language, and rules of the grammar
proper. One implication or implementation of this difference is the principle that
the output categories of the grammar proper do not include the basic categories of
the lexicon. So we might at first blush perhaps identify these basic categories
(B_CAI) with the X^0 categories of various frameworks. This is not quite the whole
story, however, because of the existence of phrasal complex lexical items,
consisting of combinations of elements each of which is member of some X^0
item: look up, rely on..., and the like.

2.2.1. Resultant (output) Categories
So for a start, (morpho-)syntactically, any operation or item that is categorized as
taking items of some category as argument or input and yielding an item of a
(basic) lexical category as output will be a lexical category. This is hardly a new
idea. Traditionally, affixes and compounding operations are regularly classified as
noun-to-noun, adjective-to-noun, and so on. We want to ask this question:

Are there resultant categories for derivational processes that are not among the
syntactic categories of a language?

2.2.2. Argument (input) Categories
What the input or argument categories for lexical categories is is a more
controversial question. In particular:

(15) Can phrasal categories be input to lexical rules?

Watch out for a confusing factor here: the internal grammar of words is not to
be identified with lexical grammar (cf. Di Sciullo and Williams, 1987).

Now let's go back to look at some more real stuff.

2.3. Wakashan Roots and Affixes
First let's look at a few examples from Haisla:

All Haisla lexical words are built up on the basis of an initial root, and
sequence of (zero or more) derivational affixes. Here's the word for Haisla:

(16) 'ixsdq"nsa: \quad vl\text{-ik}\quad -edu\quad -qi\quad -a
    bald-headed eagle good \quad \text{color} \quad \text{head} \quad \text{completive}
guk"la \quad \text{vl}guk\"\quad -[gj]l\quad -a
    build a house \quad \text{dwell} \quad \text{-make} \quad \text{completive}
gux"(guk") \quad \text{house}
guk"ela \quad \text{dwell at}
Now, there are several things to say about these examples.

First, the roots are not in general lexical items. In form, this is sometimes obvious: the item \( \text{bek}^\circ \) for example is unpronounceable by itself. With \( \text{guk}^\circ \) it is not quite so obvious, as there exists a word \( \text{guk}^\circ (\text{gux}^\circ) \) with the meaning ‘house’, but the meaning of this is not the same as that of the root, and it makes sense to think of the word meaning ‘house’ arising from zero-affixation.

Second, there is a question about the status of predicatives in Wakashan (Jacobsen, 1979): Are there nouns and verbs in these languages. I don’t think this issue is settled. Suppose there are. It still seems pretty clear that all predicatives in Wakashan (Northern at least) are available to be used as nominals or verbal in the syntax. Then the input and the output of the derivational processes are also of this category, a sublexical category that must have some additional dressing to emerge into the syntax. So roots and affixes and (possibly) stems are categorized in ways that are distinct from the ways in which lexical items and also phrasal constituents are categorized. What about their meanings? I am going to suggest that roots, and some affixes have resultant meanings that are also different from the meanings that lexical items have. Take first a root like \( \text{bek}^\circ \). It cannot take an argument, hence it cannot be a simple predicate of type \( <\epsilon,\epsilon> \), either nominal (as a base for generalized quantifier meanings) or verbal as an intransitive verb to combine (either as function or argument) with a subject. It contributes a component of the meaning of the words derived from it. A decompositional view of meanings might identify its meaning with some component of a complex meaning. Further, when it combines with a derivational affix, it forms an input category that cannot be identified with any lexical category. Moreover, the affixes that take roots and stems as arguments take and yield meanings that cannot be identified with the meaning of any lexical item. Hence, these items are truly ‘impossible words.’

Let’s now look briefly at another language (group) for some somewhat different kinds of exhibits: Algonquian. I’ll take my main examples from Western Abenaki, since it is one language I am beginning to get a little first hand knowledge about, but I draw freely from extensive literature on other languages, much of which is squarely based on Bloomfield’s work (on Menomini, Eastern Ojibway, Cree).

Unlike Wakashan, the Algonquian languages exhibit a very sharp difference, both morphologically and syntactically, between Nouns and Verbs. In the verbal
system, they all follow a fourway classification between intransitive and transitive verbs, crosscutting a classification according to the gender classification of the first argument of the verb, subjects for intransitives and objects for transitives, giving the wellknown categories:

(17) **Algonquian Verbs: AI, II, TA, TI**

that is Animate Intransitive, Inanimate Intransitive, Transitive Animate, and Transitive Inanimate, as in these examples from

(18) **Western Abenaki (Laurent, 1884):**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AI</td>
<td>mkuigo ‘he, she, it is red’</td>
</tr>
<tr>
<td>II</td>
<td>mkuigen ‘it (inanimate) is red’</td>
</tr>
<tr>
<td>TA</td>
<td>nami(h)ja s/he sees (a mink etc.) indefinite animate</td>
</tr>
<tr>
<td>TI</td>
<td>n’namito I see (a house etc. indefinite inanimate)</td>
</tr>
</tbody>
</table>

I want to mention here two interesting kinds of subword elements that go into complex words in Algonquian. The first are the so-called ‘finals.’ These are suffixes that go onto stems to yield items of the four categories just mentioned. Sometimes, they carry some kind of meaningful content, but often they only serve to make a verb of the various categories, often into animate/inanimate pairs, as in the examples just given. It would take us too far afield to try to give anything but the barest facts. The analyses are exceedingly complex as lots of stuff gets glommed together with a lot of edge effects, allomorphy and so on. I let me just cite a few examples from

(19) **Maliseet-Passamaquoddy** (Sherwood, 1986):

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AI - II pairs:</td>
<td></td>
</tr>
<tr>
<td>mceiko</td>
<td>he is bad</td>
</tr>
<tr>
<td>mceikan</td>
<td>it is bad</td>
</tr>
<tr>
<td>nisawak</td>
<td>they are two in number</td>
</tr>
<tr>
<td>nisanol</td>
<td>they (in) are two in number</td>
</tr>
<tr>
<td>talak*aso</td>
<td>it (an) is cooking</td>
</tr>
<tr>
<td>talak*ste</td>
<td>it (in) is cooking</td>
</tr>
</tbody>
</table>

(The suffixes here contribute the meaning ‘cook, burn’ as well as specifying the gender class of the subject.)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TA - TI pairs</td>
<td></td>
</tr>
<tr>
<td>namoyal</td>
<td>he sees him (obviative)</td>
</tr>
<tr>
<td>namihton</td>
<td>he sees it (compare the WA examples above)</td>
</tr>
<tr>
<td>mahal</td>
<td>he eats him (obviative)</td>
</tr>
<tr>
<td>micin</td>
<td>he eats it</td>
</tr>
</tbody>
</table>
Like the Wakashan examples, these finals are examples of impossible words, defined as taking stems that are unspecified as to their categorization and yielding items of definite lexical categories.

The other kind of morphemes I want to cite from Algonquian are the so called ‘theme signs’ which comprise part of the inflectional system of verbs. Let me go back to

(20) **Western Abenaki Theme Signs**

knami\(i\) ‘you (sg) see me’
k- nami- -i
2- see -you-to-me
knamiol I see you
k- nami- -ol (-\(\square\)\(l\))
2 see me-to-you

Here, the theme signs are -i and -\(\square\)l, which specify the orientation of the arguments as indicated in the glosses, and there are two more theme signs for local (1st, 2nd) subjects and objects and third person arguments. (Sometimes people put all these together into a system of ‘direct’ and ‘inverse’ marking.)

We might call these elements ‘submanteaux’ or ‘partmanteaux’: converses of the items called ‘portmanteaux’ that combine several disparate elements into a single affix: many languages show such forms in the inflections of transitive verbs, where we see subject-object combinations as in the following examples from Yup’ik.

2.4. **Portmanteaux**

(21) **Yup’ik transitive inflections**

tangrr- see

3
ISG-IISG: tangramken SG-IIDU: tangramtek
ISG-IPL: tangrameci SG-IISG: tangrraka
ISG-IIIDU: tangrragka ISG-IPL: tangrranka
IDU-IISG: tangramegten IDU-IIDU: tangramegtek
etc. for all combinations...

| B Meaning rule for a ‘me-you’ portmanteau: a function f from transitive verb meanings to sentence meanings such that for all transitive verb... | 10 |
Some More Impossible Words

meanings $x \quad f(y)$ is true iff the speaker stands in the relation $x$ to the hearer ($y$ (you')(me')).

Patterns like these are familiar also from the Iroquoian languages as well. The point is that it is pretty useless to analyze such affixes into combinations of independent morphemes for the two participants.

Claim: no lexical word denotes such a function. Is it impossible to derive according to your favorite theory of deriving words? Great, then we have another morpheme that is an impossible word.

(In discussion, Richard Rhodes gave a counterexample: Klamath has three words that are portmanteaux for you-and-me subject/object meanings, cf. Barker, 1964: 242-3. I take this to show that my main claim has to do with open lexical classes, so that there will be no general category of Nom-Nom combinations like Japanese oyako ‘parent-child’ that can function as subject-cum-object with a transitive verb: OYAKO love(s) = ‘parents love children’.)

3. Final Discussion

Let me return to the two questions I asked earlier in this talk: How come languages are so similar to each other? How come they are so different? There is a kind of practical puzzle that reflects these questions. On the one hand, serious linguists can claim that human languages are so similar that a Martian scientist would conclude that they are all the same (Chomsky, 1995; Pinker, 1994, actually, this is not exactly what Chomsky says). On the other hand, try to tell an unspoiled -- that is, nonlinguist -- speaker of any two mildly different languages that they are basically identical and you will be laughed out of court. How can this be?

Even writers like Chomsky will add a proviso: “up to the lexicon” or the like. The kinds of facts I have touched on here today show that differences in the grammar of words and in the resources and patterns that people draw on to make up new words and other lexical items are a far from trivial matter. They make up a major part of the texture of a language as it lives in the minds and hearts of people. I suspect that when we are impressed with the extreme diversity of languages, we are basing our judgments primarily on differences among languages in the realm of word-grammar, and on the relationships between word-grammar and phrase-grammar. These domains lie behind the old and venerable typologies that use terms like: analytic, synthetic, polysynthetic, isolating, fusional, agglutinative.

I guess what it comes down to, when you reflect on how similar or different languages are, is: Compared to what? Compared to cabbages or kings, languages are very similar, practically the same, we might say. But compared to styles of walking? to modes of expression and decoration in clothes? cooking? art?

Maybe I can end as I started with Sapir and Humpty-Dumpty.
Sapir: there is a section in *Language* where Sapir argues for the psychological reality, one might say immediacy of the word, there is another where he discusses differences among languages in quite esthetic terms, and his examples go primarily to the domains I have been discussing here. It is hard to talk about the word structures of a polysynthetic language like Nuuchahnulth or Haisla without using words like “exuberant.”

Really finally, Humpty again:

“*The question is,*” said Humpty Dumpty, “‘which is to be the master — that’s all!'”

References

Some More Impossible Words


