Published by CU Scholar, 1989

Colorado Research in Linguistics 10

American Anthropological Association Conference on American Indian Languages Philadelphia, December 1986

MORPHOLOGY AND POLYSYNTHETIC LANGUAGES

DAVID S. ROOD

University of Colorado

O. Introduction

Linguists concerned with gathering and describing language data, and linguists concerned with discovering the universal principles of human language, have not always had much to say to each other, strange as that must seem to an outsider. For example, when I started to write my dissertation, Chomsky's <u>Aspects</u> (1965) was only a year or two old, and I found the model described in that book totally useless. During the past 10 years or so, however, the theoreticians, particularly those of the Government and Binding school, have rediscovered or revived the study of morphology, and what they are now investigating seems to be something to which my knowledge of particular languages can contribute. Today I would like to discuss with you a small part of the range of implications that polysynthetic languages, particularly Wichita, have for linguistic theory.

Wichita, as many of you know, is a North Caddoan language remembered (but seldom spoken) by one or two dozen people now living in central Oklahoma. It has a very small inventory of phonemes, but a complex system of morphophonemic rules, and its grammar is a prototypical example of polysynthesis. As is the case with such languages, much of the information in any utterance is expressed through bound morphemes on the verb. One thus has the impression that the language has mostly words (and often very long ones) rather than sentences.

The theoretical issue to which I wish to address myself today is the debate between those who advocate a grammar in which all word-formation is accomplished in the lexicon by means of lists, rules for compounding, and morphological rules, and those who argue that word-formation is often informed by and/or has effects on syntax. The former position, advocated by Arnoff (1976), Lapointe (1980, 1983), and Jensen and Jensen (1984) among many others, is often stated as a constraint on grammars something like (1):

(1) No syntactic rules can refer to elements of morphological structure.

This position is challenged by, among others, Anderson (1982), Sadock (1980, 1983, 1986) and Baker (1985). My own position is closest to that of Sadock and Baker. In arriving at their respective positions, Sadock has utilized data from Greenlandic, and Baker has employed Southern Tiwa and languages from the Iroquoian family, as well as languages from other parts of the world. I think that the Wichita data support their fundamental claims, but force some expansion of them.

I am setting out to prove that the grammar of Wichita requires the use of syntactic rules which refer to parts of words. To do this, I will first spend a little time demonstrating

that the structures I am describing are words. Then I will show that some of the rules which generate the core arguments of verbs are part of word formation, that parts of words must be accessible to syntactic rules to account for anaphora, that noun incorporation is most likely a syntactic process in this language, and that verbal morphemes also express non-head parts of both noun phrases and verb phrases.

The simplest possible Wichita verb is the intransitive structure in (2):

(2) tackwa:wa?as 'I am/was eating.'

It consists of four elements: a tense prefix, an aspect suffix, a first person subject marker, and a verb root. All the morphemes are bound and can occur in this sequence only, the morphophonemics (such as the change of /t/ to /c/ before a consonant, and the change from /w/ to /kw/ after /c/) are those of internal sandhi, and the structure is treated as an manalyzeable unit by native speakers. I do not believe that anyone could argue that this is

This structure can nevertheless be expanded considerably, although most such expansion takes place in the middle of the positions just illustrated. A somewhat longer word" is thus that in (3):

hiki:se⁹i:rikichi⁹ncki 'I guess I'll sleep a little while.'

Here we have supplemented the structure in (2) with two morphemes, one expressing abbreviated duration and one expressing diminution of the event.

with transitive verbs, we add information about the object, including features of person, humber, and (with some verbs) noun class; (4) is not exceptionally complex, and still epresents a single word:

hitaciya:rhi⁹nna⁹as 'we (incl) brought them (animate)'

Finally, we can expand these forms with morphemes which sometimes appear as ndependent words as well, but even these additional elements appear in the middle of things hich must otherwise be considered single words. I refer of course to incorporated nouns uch as 'knife' in (5):

(5) isi:ri''asta:hanna''a 'Bring them; their; knives'

Note that in (5) there are references to four separate noun phrases included in the verb: addressee, recipient, direct object, and possessor of direct object. The only potentially free morpheme in the complex is medial ta:ha 'knife', and even that is not the citation form of the noun (which would be ta:ha?a).

My main demonstration that these complexes are single words, then, is that they are composed of bound morphemes. Even when nouns are incorporated into the verbs, their form is (usually) different from their citation form, and they are surrounded by bound elements.

Now to the question of how such complexes are generated by the grammar. My first argument is that the morpheme set which includes those glossed 'I', 'me', 'we' and 'you' in these examples are not agreement markers, but genuine pronouns. (This is similar to an argument about parallel structures in Breton made by Anderson (1981), though Jensen and Jensen (1984) challenge that analysis.) In the Wichita case, I think the strongest argument is a negative one: there is no element in the language with which these forms can agree, so an analysis of them as "agreement" would need to posit an abstract form which has no surface realization in the language. The so-called "pro-drop" languages, like Spanish and perhaps Breton, at least have independent pronouns which can occur in emphatic or topicalized constructions, but the equivalent forms in Wichita are the nominalizations of fully inflected verbe such as those in (6), no reference to a first or second account. verbs such as those in (6); no reference to a first or second person can be made unless a verb root is included in the construction.

(6) a. nac⁹ih tati:ci⁹is 'I was the one -- I did it' (CT 74:41) ta- t- ucna- t- 7i- h nom - I - be - subord indic - I - preverb - do - impf b. ha:ri: 'tac'i 'It's me' (answer to 'Who is it?')

ha:ri:9 ta- t-9i-Q that indic - I - be - perfective

It is even possible to argue that Wichita has no person agreement system at all. If the non-third person reference morphemes are pronouns, then the only candidate for "agreement" would be the third person singular form of some of the tense prefixes (there is no consistent third person morpheme), since it usually co-occurs with independently expressed nominal subjects. Yet there are text examples where overt nouns are cross-referenced anaphorically by non-third person verbs, and examples where non-third persons are crossreferenced with third person singular forms. Both phenomena are illustrated in the text passage in (7), while (8) provides a somewhat more convincing example of a first-person pronoun referring anaphorically to a noun:

(7) ti: acha: we: ass tackwe eskancira: ka: hi: k9i khi as te:

⁹ata:cira:kira:i:chann⁹istic kiya⁹a:cikichare⁹es kiye⁹e:rikice:we:s⁹i

'Now some of us old women, poor us, we old ones keep trying to do something for ourselves anyway; we plant a little, we have a little put away.' (CI 114-54-55)

(8) ka:hik²a ho:² natira:kwakha:r²ih 'Our (we women's) ways' (CT 114.38) ka:hi:k²a ha:wa² na- t- uy- ra:k- wakhahrwoman also nom - I - possessive - non.3rd pl - way -²i- h be - subord

In the first verb in (7) we find the construction 'we are old women', which could be analyzed as the verb 'be' with an incorporated predicate nominal, or simply as a first person verb in apposition with a noun. In any case, the referent of the noun and the referent of the pronoun are the same. In the last two verbs in the sequence, we have third person subject morphology but the translation, given by the Wichita speaker, and the logic of the sequences suggests that the semantic subject of these verbs is still 'we old women'. I do not insist on these analyses, but merely want to show that we might have here examples to show that "agreement" is not the function of the person-marking morphemes. In (8) it seems to me that we have a more straightforward example of first person "agreement" with an overt noun; it is difficult to provide a smooth English translation for that construction. If these examples are pointing in the right direction, then we would have to reanalyze the so-called third person morphology as simply the unmarked form of a finite verb.

We thus have two arguments against treating the pronominal morphemes in the verb as agreement morphemes. First, they cannot be agreement markers since the language has nothing for them to agree with, and secondly, they cannot be agreement markers if constructions occur where agreement is violated. If they are not agreement features, the only thing left for them to be is pronouns. And clearly, if these morphemes are pronouns, then whatever rules generate the subjects and objects of sentences must generate these, even though they are bound verbal morphemes; a syntactic rule must make reference to word structure.

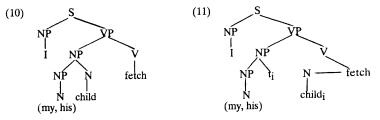
Before we leave the subject of pronominal markers, let us look a little at the rules which govern the case of the indicator of the possessor of an object noun. In particular, look at example (9):

(9) ita:ce⁹i:riye:s⁹akhi:wa 'let me go after my children first' (CT93:27)
 i- t- a:- uc - re⁹i:r- riye:s
 imperative - I - reflexive -dative - first - child

⁹ak- hi:wa- Q patient plural - fetch - perfective

(a) iti:c... 'let me...his (b) hiki:c...'let him...my'

In this verb we have co-reference between the subject and the possessor of the object, as indicated in Wichita by the presence of the reflexive. If either the subject or the possessor were third person, we would have the variants of the first part of the verb given in (9a) and (9b). Thus while the subject is marked, as expected, with subject case forms, the possessor is marked with object forms, including the reflexive if it is appropriate. The explanation for this case-marking, which is common in languages which allow noun incorporation, as given in Baker (1985) using the government and binding model, presumes an underlying structure such as (10) and an intermediate structure (after incorporation of the object) such as (11):



Briefly, the argument is that after the object noun has been incorporated into the verb, it loses its ability to assign case, and also its need to receive case from the verb. There is thus nothing left in the object NP which can assign Case to the possessor, and the verb has a Case-assigning property which is now unused. These two dilemmas solve each other if we look to the verb, which is now the governor of the possessor NP instead of the object NP, and allow it to assign its Case to the otherwise ungoverned possessor. Fortunately for Baker, the languages he looked at allow this kind of "passing along" of case functions only when the object is incorporated; unfortunately, Wichita works the same way whether the noun is incorporated or not; cf. example (12):

(12)ti[?]i 'aras hatakiccare:hih 'for me to put this meat of mine on top'
(CT93:37)

```
ti°i 'aras ha- t- a:- ki- uc-
this meat infin I - reflex -infin - dative -

ca- re:hi-
top - put pl. object - subord
```

I am not sure how GB theoreticians can account for such sentences, but this issue is peripheral to the point of this paper. For our purposes, we need to return to example (9) and recall that the representative of the object-case possessor must be a reflexive if the subject is identical with it. Thus the choice of the morpheme to represent the possessor is conditioned by the syntactic relationships among the parts of the verb. Again, we have what seems to be a clear instance of syntactic rules referencing pieces of words. (Note that the same argument would hold for simple reflexives, since after incorporation the grammar treats the possessor and the direct object alike.)

Having shown that the syntactic rules which generate subjects, objects, and reflexives must reference parts of words in Wichita, I now want to turn to another class of rules, namely those which deal with anaphora. Without going through the complex apparatus of government and binding with respect to PROs and pros and traces, or taking a position in the debate over whether anaphora is syntactic or guaranteed by various kinds of filters, let me simply indicate that such rules for Wichita must reference parts of words rather than whole words. In (13),

```
Al3) kiya iriwa:cte:s ic arasir is ha:s tahe e 'One made a big batch of parched corn. It always tasted good.'

(CT115:57)

kiya-
human subject - habitual - big - parched corn - make plural -

s ha:s ti-
impt always indic 3rd subject - preverb - taste - perfective
```

the 'it' of the second verb is clearly not the making of the big batch of parched corn, but just the corn itself. We thus have an instance of anaphoric reference to part of a word, namely the noun incorporated into the first verb. Sadock (1980:316-317) has pointed out similar behavior by when he calls incorporated nouns in Greenlandic, and compared various claims about "anaphoric islands" and "anaphoric peninsulas" with the Greenlandic data. In these cases, I do not see that the Wichita data are any different than the Greenlandic, and so I merely offer this as another case, from a language from another part of the world.

With respect to noun incorporation, however, I find that Wichita differs from the generalizations made by both Mithun (1984, 1986) and Sadock (1980, 1986). Mithun argues that the phenomenon is one of the word-formation, i.e., that it is not syntactically productive even though the number of combinations which speakers carry in their lexicons is very large. Sadock argues in favor of syntactic productivity, but he is describing a different phenomenon, namely, the derivation of verbs from bound roots by attaching nouns to those roots.

Mithun (1984) describes four kinds of noun incorporation, summarized as follows:

Type I: Noun and verb form a compound, with verb becoming intransitive; examples are somewhat like English mountain climbing or baby sitting.

Type II: Builds on Type I, but the new verb need not be intransitive; a formerly oblique argument may "advance" to core status. Examples look something like 'I face-washed him' or 'I back-broke the man.'

Type III: Incorporated nouns occur in stylistic variation with unincorporated equivalents, the former generally being old information, the latter new information.

Type IV: The incorporated element is a general, classificatory noun and another, more specific noun may occur in the construction. A Wichita example would be e:ckic^a nakic^ikih (CT 112:5) 'that which is milk', in which the element -kic- in the verb means 'liquid', and reinforces or "agrees with" the class of the subject noun, milk.

Sadock's (1980) arguments that the process he calls incorporation is syntactic include (1) the observation that "stranded" modifiers of incorporated nouns carry the same number and case agreement morphology that they would if the noun were free, even though the incorporated noun usually has no such marking; and (2) the observation that sometimes an incorporated noun can be possessed. Both of these facts argue that the noun must have been part of a noun phrase construction at some stage in its derivation.

Wichita incorporated nouns are not restricted in the way that Mithun's description of Class III implies. The first mention of a noun may find it in incorporated form, while subsequent mention leaves it unincorporated. Thus example (13) above is the first mention of the parched corn in the text, and is followed a few words later by (14), where the 'corn' is not incorporated, but it is marked as old information by a separate morpheme in the verb:

(14) te:s⁹ic⁹aras hakicire:re::skih 'to buy the parched corn'

```
haki- uci- re:r- re::hi- skih
infin - preverb - the - buy - subord impf.
```

Moreover, I do not think that there are any lexical restrictions on incorporation in this language, though of course such a claim is very difficult to prove. Note, however, example (15), taken from a text which was recorded when a bilingual speaker was addressing other bilinguals, and which included a considerable amount of English mixed with Wichita:

(15)icka: isa:cinn-mule-s?ice;risiki 'He just beat his mule up that way for

nothing'

```
icka: ' isa'- o- a:- uc- irir-
just thus -timeless - reflexive -dative - in vain
```

```
mule- s- rice:risiki-
mule - incorporated noun - beat up - perfective
```

In this form, the English word 'mule' has been treated like a Wichita noun, suffixed with the -s which accompanies most incorporated animates, and inserted into the verb. It

Sadock's arguments for the syntactic status of incorporation do apply to Wichita, however, so if they hold up for Greenlandic they will also hold up for Wichita. Examples of "stranded" pieces of noun phrases abound in the texts, and in fact, in Wichita, we find both heads and non-heads of such phrases inside and outside the verb. One example is the "definite" marker glossed "the" in example (14), and illustrated twice in (16):

```
(16)ni:ce:ri°iskih wiyasaks nare:r°ih 'the one that did it was the boy'
 (CT97:47)
```

```
re:r- ri?i-skih
                                           wivasaks
nom - preverb - the - do - subord impf boy
na- re:r- <sup>9</sup>i- h
nom - the - be - subord
```

It is necessarily a bound morpheme, usually found in a verb but occasionally suffixed to a noun instead. Sometimes the verb 'be' seems to be used only to carry this morpheme, as in

(17) a:ko:k'a ka:hi:ra:i:c'a nare:r'ih 'The old lady said' a:ko:k⁹a ka:hi:ra:i:c⁹a na- re:r- ⁹i- h 3rd said old lady nom - the - be - subord

An example of an adjective outside the verb while its modified noun is inside is found in (7) above, 'old women'. Examples of adjectives inside the verb are found in (7) and (13), though in (7) there is no noun as head of the phrase, and in (13) the head is also inside the verb. I should note that there are apparently only three adjectives which behave this way, these two forms for 'little' and 'big', and an unincorporated noun:

```
(18) k<sup>9</sup>ita:ks ka:<sup>9</sup>a:<sup>6</sup>a:kiriya:s<sup>9</sup>a 'along came old Coyote'
        k<sup>9</sup>ita:ks ka:9-
        coyote new topic - quotative - preverb -
```

riya:s- ºa- 🛇 aorist -old - come - perfective

It is possible that "adjective" is the wrong analysis for these morphemes, and that they are instead some kind of modifier of the nature of the action. Both 'big' and 'little' occur this way--see example (2)--but it is difficult for me to imagine "old" as having adverbial force. Moreover, Wichita speakers always translate these forms as if they were noun modifiers when there are nouns around to modify.

There thus seems to be considerable disruption of the structures which form noun phrases in the English translations of these Wichita examples: the definite article and certain adjectives are frequently separated from the nouns they supposedly modify, with one or another of the pieces expressed as part of the verb. There are no agreement phenomena such as the Greenlandic case and number suffixes to indicate that the pieces were once all parts of

the same phrase, but unless we conceive of Wichita underlying structure as very different from English (which, as Chomsky says, would imply problems in learning one language or the other), we have to believe that they do belong together at some level of structure. Whatever that abstract level is, it must describe English syntax and Wichita morphology at the same time.

There is much more to be said about the relationship between word formation and syntax in Wichita. We have not mentioned the treatment of dative objects, even when saying something about possessor marking, though those two case structures are intertwined. We have not discussed the adverb-like morphemes illustrated in some of the examples ('for a while' in (3), 'first' in (9), 'in vain' in (15)); some of these are manner expressions, and thus analyzed in GB as part of the verb phrase, but some are time, traditionally outside the VP. Nor have we discussed the embarrassing fact that all this verb structure comes with inviolable, rigid sequencing of elements, but sequencing that totally defies any kind of binary branching tree structure that maintains the integrity of phrases; I have not tried to formulate the movement rules that would put the pieces of phrases into the right position classes, but I know that they would be complex. And finally, we have said nothing about the Wichita lexicon, meaning the derivational apparatus that generates complex verb stems and compound nouns, although such processes are important (but different in kind from the syntactic operations we have been looking at).

So what can we do to describe Wichita (or other polysynthetic languages) in a way that captures their similarity to other languages without doing violence to their own special character? Sadock (1983:210) has phrased the dilemma as follows: "If we treat the productive processes as syntactic, then polysynthetic languages end up being described as having little or no morphology-where in fact they have more than other types. But if we choose to remove the redundant power from the syntax (the common procedure) then we end up describing polysynthetic languages as virtually lacking syntax, such processes as passivization, raising, comparison, and complementation becoming rules of the morphology only."

It seems that word-formation takes place in different places in the grammars of different languages. Much of Semitic surface structure is apparently best generated in the phonology, if I understand McCarthy (1979) correctly. It may well be possible for descriptions of English and similar languages to exclude syntactic rules from referencing pieces of words. Whether or not such a move is feasible for Breton or Georgian, which display considerable co-indexing between verbs and arguments, seems controversial; that is the subject of the exchange between Anderson (1982) and Jensen and Jensen (1984). But such a move is clearly impossible for Greenlandic, Koyukon (Axelrod 1986), or Wichita. These languages must allow the syntax to generate a considerable chunk of their word structure, while nevertheless maintaining rules for morpheme order which are likely not the same as syntactic transformations.

It is exciting to read theoretical work like that of Baker (1985), which accounts for a variety of structures in numerous languages from all over the world with a few modificationsin fact, simplifications--of GB theory; but before we can claim to know what is really universal about human language, there is more work to be done.

S

THREE-TERM SPACE DEIXIS

Colorado Research in Linguistics 10

Kumiko Takahara

Abstract. Space Deixis is a linguistic device to locate objects in space relative to speakers and hearers. These deictic systems differ cross-linguistically in creating spatial divisions from two to nearly a dozen ways. This paper is concerned with three-term deictic systems which are commonly found among Altaic languages with special attention to Korean, Japanese, and Turkish. Semantic characteristics of these systems are generally explained as locating objects (1) close to speakers, (2) close to hearers, and (3) remote from both speakers and hearers (Lyons 1977). Various occurrences of space deixis are analyzed with the application of this classificatory scheme, and it will be pointed out that complexity of semantics involved in space deixis of these languages requires an alternative description. First, all three divisions are modified with the notion of psychological shift in point of orientation. Secondly, the contextual features of direct witness or absence of witness need to be considered in determining the function of space deixis. In non-witness speech situations, objects are localized distinctly in space for the reason that the tertiary divisions by the space deixis are functionally reduced to identify speech event either (1) close to speakers and hearers, or (2) remote from speakers and hearers.

All languages have linguistic devices for making reference to an entity in the context of speech events by placing it at a certain point of time or location or in a certain status relationship with a speaker, addressees and other participants. These devices are typically grammaticalized into time, space, and person deixis in the languages with differing degrees of complexity. Take space deixis, for example. Although most languages tend to locate and identify an object or an event at proximal or non-proximal distance relative to where a speaker or the addressees are, some languages elaborate their deictic systems further by encoding different space dimensions and different point of spatial orientation. Consequently, according to the size of combined variables space deixis may consist of any number of deictic terms. Out of these varieties the present paper is concerned with the semantic characteristic of the three-term space deixis which is commonly found among the Altaic or peripheral Altaic languages such as Korean and Japanese. Furthermore, the deictic expressions which are examined consist of lexical space terms, not the phrasal or more structurally complex types.

Initially, morphological constitution of the three-term space deixis of our interest is looked at in comparison with a prototypical two term deixis in English. While English space deixis has two locative adverbs and two demonstrative phrases, the three-term deixis consists of three base morphemes from which the deictic terms in all categories can be derived.

English	<u>Korean</u>	<u>Japanese</u>	<u>Turkish</u>	Tungus
Proximal: here/this place	yo-gi ko-gi	ko-ko so-ko	b-r(d)a s-r(d)a	e-ne
Distal: there/that place	cho-gi	a-so-ko	o-r(d)a	te-re

The Altaic deictic terms combine locational indicator of close to and further from the speakers and/or hearers and a prime morpheme for point in space, -gi, -ko, or -ra. In this system two points of reference are used in addition to two-way divisions of space. A referred entity is, therefore, identified in terms of whether it is close to a speaker or a hearer, or is remote from both speaker and hearer. On the other hand, in the two-term deictic system,

References

Anderson, Stephen R. 1982. Where's Morphology? Linguistic Inquiry, 13:571-612. Aronoff, Mark. 1976. Word Formation in Generative Grammar. Cambridge: MIT Press. Axelrod, Melissa. 1986. Word-based Morphology: Some Problems from a Polysynthetic Language. (In this volume).

Baker, Mark. 1985. <u>Incorporation: A Theory of Grammatical Function Changing.</u>
Unpublished doctoral dissertation, MIT, Cambridge, MA

ensen, John T. and Margaret Stong-Jensen. 1984. Morphology is in the Lexicon! Linguistic

apointe, Steven G. 1980. A Theory of Grammatical Agreement. Unpublished doctoral dissertation, U. Mass., Amherst, MA.

1983. A Comparison of Two Recent Theories of Agreement. In Richardson,

Marks, and Chukerman, pp. 122-134.

McCarthy, John J. 19 . Formal Problems in Semitic Phonology and Morphology. Indiana University Linguistics Club.

Mithun, Marianne. 1984. The Evolution of Noun Incorporation. <u>Language</u>. 60:687-894. 1986. On the Nature of Noun Incorporation. <u>Language</u>, 62:32-37. arks, Douglas R., ed. 1977. Caddoan Texts. NATS 2.1.

Richardson, John F., Mitchell Marks, and Amy Chukerman (eds.) 1983. Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax. Chicago: Chicago Linguistic Society.

adock, Terrold M. 1980. Noun incorporation in Greenlandic. Language, 56:300-319.

3983. The Necessary Overlapping of Grammatical Components. In Richardson, Marks, and Chukerman, pp. 198-221.

1986. Some Notes on Noun Incorporation. Language, 62:19-31.