orary Linguistics

AUTOLEXICAL SYNTAX
A Theory of Parallel Grammatical Representations

Jerrold M. Sadock

The University of Chicago Press
Chicago & London

Joyce McDonough 1/91

4

Incorporation

4.1 Historical Background

For purposes of the present chapter, I will use the term "incorporation" to refer to phenomena in natural languages (other than cliticization as described in chapter 3), where a proper subpart of a word can be shown to have the function of a formative, i.e., an atomic element, in the syntax. The details of the morphological joining are immaterial to my definition. The incorporated item may be a stem, either compounded with another stem or used as a base for affixation. It may be an affix or even a symbolic process, and may, in either case, be classified as derivational morphology or as inflectional morphology.

In chapter 6, I will broaden the notion of incorporation to include examples of discrepancies between semantic organization and either morphological or syntactic structure, but for now my attention will be restricted to cases where it is clear that part of a morphological word is to be represented as a formative in syntax. In the absence of positive syntactic arguments to the contrary, I will assume that pieces of words are not to count as independent elements of the syntax. By itself, no purely semantic principle, such as Baker's 1988 Uniformity of Theta Assignment Hypothesis, will be taken as requiring an analysis in terms of noncoincident morphosyntactic bracketing, and thus the range of phenomena that will be discussed in this chapter will be considerably narrower than that of those phenomena labeled incorporation in Baker 1988.

By far the most discussed putative instance of nonclitic interpenetration of morphology and syntax is noun incorporation, so called because a noun stem representing a syntactic argument of a verb (typically the object) is found as a part of the morphological verb. This phenomenon is frequently encountered in North American languages, but it is not found either universally or exclusively there. Noun incorporation was introduced into the modern literature as evidence against the strict lexicalist hypothesis in work by Rischel (1971, 1972) and myself (Sadock 1980). In those languages where it has a clearly syntactic face, noun incorporation offers excellent evidence against the hierarchical model of the relation between morphology and syntax, and it was primarily as an attempt to handle this striking phenomenon that I originally suggested the model of autolexical syntax.

It has recently been reasserted, however, that noun incorporation is, after all, a strictly morphological process with no important syntactic ramifications (Mithun 1984; Hopper and Thompson 1984; Di Sciullo and Williams 1987; and most recently Rosen 1989), thus resurrecting a debate that took place early in this century. The term "incorporation" was first employed by Humboldt ([1836] 1988) who applied it to Nahuatl examples like *ni-naca-qua* 'I-meat-eat'. Humboldt's argument was the distinctly nonrelativistic one that languages like Nahuatl confuse syntax and morphology and are thus less perfect than those like Classical Greek or Sanscrit that clearly separate the two.

> [T]he METHOD OF INFLECTION in all its completeness, . . . alone imparts true inner fixity to the word for both mind and ear, and likewise separates with certainty the parts of the sentence, in keeping with the necessary ordering of thought. . . . Compared with the INCORPORATIVE PROCEDURE, and that of loose ADDITION without true word unity, the METHOD OF INFLECTION appears as a principle of genius, born of a true intuition of language. (Humboldt [1836] 1988, 145)

In apparent reaction to the by-then unpopular glottocentricity of the Humboldtian view, Kroeber (1911) attempted to ban the notion on general principles:

> In short, the term "incorporation" is a delusion, whether applied to pronoun or to noun. It must be relegated to the same category as other antiquated catch-words such as "agglutination," which have originated in the assumption that the languages of so-called uncivilized people must contain certain features of a kind totally different from those characteristic of Europeans—and incidentally features of an inferior order,—and which have found their chief vogue and employment not among serious painstaking students of language but among doctrinaires, compilers, and those false popularizers who think to diffuse knowledge by giving a phrase instead of an idea. (Pp. 582–83)

Kroeber (1909) offered the following definition of noun incorporation, and then proceeded to argue that the phenomenon did not (indeed, could not) exist:

> Noun incorporation is the combination into one word of the noun object of the verb and the verb functioning as the predicate of the sentence. It is essential that the resultant of incorporation is a single word, else the process is without limit and all syntactical relation may be construed as incorporation. (P. 37)

Note that Kroeber's straw-man definition is essentially antilexicalist; it assumes (as I do) that the phenomenon of incorporation is definitionally one that

straddles the border between syntax and morphology. Soon after Kroeber published the brief paper from which the above passage is quoted, Sapir (1911) criticized him precisely because of the transmodular nature of his definition. Displaying a strikingly modern prejudice, Sapir argued that general considerations required that the definition of noun incorporation be either morphological, or syntactosemantic, but not both at the same time.

> Examining this definition, we find that two things are required—a noun must combine with the verb-predicate into a word-unit, and the noun so combined must function as the object of the verb. The first requirement is morphologic in character, the second purely syntactic; in other words, the first calls for a certain type of word formation, while the second demands that a particular logical relation subsist between the two independent elements that enter into this word formation. Without denying the abstract right to set up such a definition, it would seem that the combining of a morphologic requirement with an independent syntactic one yields, on general principles, a definition of too narrow a scope for the discussion of as fundamental a problem as noun incorporation is felt to be. Noun incorporation is primarily either a morphologic or syntactic process; the attempt to put it under two rubrics at the same time necessarily leads to a certain amount of artificiality of treatment. (P. 255)

In denying that it was useful to talk directly about the "logical" (read: syntactic) function of pieces of words, Sapir was the first lexicalist. His own proposed definition of noun incorporation was purely morphological:

> It is this process of compounding a noun stem with a verb that it is here proposed to call noun incorporation, no matter what the syntactic function of the noun logically is. The type of verb, "to song-write," that Dr. Kroeber alone regards as illustrative of noun incorporation, is best considered a particular class of the more general type of noun-verb compound verb. (P. 257)

Since Sapir took the morphological means (compounding) by which the nominal and verbal element are joined as the criterion for noun incorporation, he was able to argue that it was well attested in American languages. In fact he argued that it exists covertly in what Bloomfield (1933) was to call synthetic compounds (Roeper and Siegel's (1978) verbal compounds), like *man-eater*. These Sapir took to be agentive nominals derived from fictitious compound verbs such as **man-eat*, itself an example of noun incorporation.

This entirely morphological definition of noun incorporation as compounding made Sapir unable to count the Eskimo phenomenon that now goes by that name as a genuine example of the type, since in Eskimo the morphological

facts are quite different. There we find straightforwardly derivational means being used to derive verb stems from noun stems. In fact, the Eskimo languages have no (original) compounding of any kind.

> Eskimo, a language particularly rich in suffixes that verbify nouns, has been termed polysynthetic, but has not been employed by serious students as a source of examples of noun incorporation. (P. 254)

Though Sapir spoke of the "syntactic function" of the incorporated noun, his actual idea was that a logical relation that could be expressed syntactically can also be expressed morphologically. The syntax is "sacrificed" in favor of the morphology, and only the logical relation remains. This is very nearly the same idea that recent work attempts to formalize in terms of theta-role assignment either in the syntax, or in the morphology, but not both (Rosen 1989).

Sapir won his theoretical point, and Kroeber quickly and completely capitulated (Kroeber 1911); noun incorporation was to be understood as nothing other than the morphological compounding of a noun with a verb. Nevertheless, throughout his article, Sapir seems to equivocate, speaking as if it is a necessary feature of noun incorporation that the equivalent of SOME syntactic relation is expressed. The last footnote in his article is telling; it reads:

> Since this was written (June, 1910) Mr. J. P. Harrington [Harrington 1910] has published sketches of two Tanoan dialects, Tiwa and Tewa. In Tiwa both direct and indirect objects may be incorporated in the verb complex, coming between the pronominal prefix and the verb stem; such incorporation is obligatory for singular direct objects. (P. 282)

The direct object relation is syntactic. It is therefore difficult to see how it makes any sense to say that the incorporation of a certain syntactic relation is obligatory in a theory in which incorporation is purely morphological. It is also hard to understand how Sapir could reconcile his claim that the logical relation of the incorporated element is of no consequence with the demonstration that a certain syntactic function (that of direct object) is treated specially. The explanation for this apparent inconsistency between Sapir's methodological pronouncements and his descriptive behavior lies, I believe, in the fact that noun incorporation in American languages like Tiwa often is of such a character as to demand a description that is simultaneously morphological and syntactic, theoretical prejudices to the contrary notwithstanding. The closing sentence of Sapir's important article in fact reveals that this really had been his opinion all along. Note that here he even relaxes the erstwhile morphological sine qua non, admitting that derivation can be an instance of the "process" (of noun incorporation).

> The characteristic fact about the process is that certain syntactic relations are expressed by what in varying degrees may be called composition or derivation. (Sapir 1911, 282)

It is just this point that I tried to make in Sadock 1980 concerning noun incorporation in West Greenlandic. Though Greenlandic noun incorporation and Tiwa noun incorporation are morphologically dissimilar processes, the former derivational and the latter compositional, their transmodular nature binds them together and makes them different in kind from run-of-the-mill compounding, or run-of-the-mill derivation, where indeed, syntax is irrelevant. In an English form like *pan-fry*, for example, we have the same sort of morphological style that we find in Tiwa incorporation, and in the formation of denominal verbs like Sapir's neologism *verbify* we find the same morphological relation as in Eskimo incorporation. But we would not be tempted to call either of these English formations "incorporation," for neither has anything like the syntactic ramifications that noun incorporation in Tiwa or Greenlandic has, as we shall see.

Since a great deal more is known about the details of incorporation than was the case eighty years ago, I suggest that we reopen the Kroeber-Sapir debate. I believe it ought to be called a draw: Kroeber was right at the outset in providing a transmodular definition of noun incorporation, but wrong in claiming that it did not exist; Sapir was correct in his contention that noun incorporation is a fact of language, and also that various syntactic relations can be involved, but wrong in his claim that it can be understood in purely morphological terms.

4.2 The Nature of Noun Incorporation

We should not be at all surprised to find that in some particular languages some word-building process displays all of the characteristics that we have come to expect of lexical relationships: incomplete productivity, phonological and semantic unpredictability, syntactic identity with underived forms, and discourse opacity to the word-internal morphemes. We should not even be surprised if these traits show up in noun incorporation, a word-building process that would seem to be amenable to a syntactic treatment, if any is; for the null hypothesis clearly ought to be that any individual word-building process does not interact with the syntax.

In the past few years several authors have asserted that noun incorporation (and presumably then, all nonclitic morphological constructs) are properly treated without reference to syntax cf. Mithun 1984; Di Sciullo and Williams 1987; Rosen 1989). The copious citation of facts from various languages compelling a lexical treatment for noun incorporation IN THOSE LANGUAGES simply demonstrates that such a treatment is correct FOR THOSE LANGUAGES.

But the crucial issue for linguistic theory is whether there are ANY cases of noun incorporation that must be viewed as simultaneously morphological and syntactic. A single well-documented example will make the positive case, whereas the negative case can never be proved absolutely. It appears to me that there are languages, in particular Eskimo, the very same Tiwa mentioned by Sapir in 1911, and probably some South Munda languages such as Gta?, that present compelling evidence that their particular version of noun incorporation has a syntactic face. The ramifications of such a demonstration are considerable, for once the syntactic reality of some morphological entity is recognized and a constrained theory of its properties is set up, the door is open to the treatment in similar terms of other sorts of morphological operations, where the recognition of the syntactic nature of the phenomenon is not necessary but merely highly desirable from a descriptive point of view.

Following Sapir's lead, it might be assumed, since the morphological technique for incorporating noun stems in Eskimo is derivational rather than compositional, that the Eskimo morphological process called noun incorporation is just not relevant to the debate. But we must not lose sight of the larger issue—whether morphology is always irrelevant to syntax or not. The claim that internal morphology never has any syntactic value would be falsified by derivational incorporation, such as is found in Eskimo, every bit as easily as by an example of the compositional type. If anything, we should be more surprised to find syntactic interactivity in straightforward cross-categorial derivation than in compounding, since compounding is a morphological phenomenon that prima facie seems to straddle the border between two major components of grammar. At any rate, Southern Tiwa and Munda incorporate by compounding and still display characteristics that compel a transmodular analysis.

In what follows I will show that in some languages there are properties of incorporating structures that are inconsistent with a purely morphological view of the phenomenon, *pace* those who have claimed that such properties never exist. I will show, in other words, that there are languages where the syntactic relevance of noun incorporation cannot be denied.

4.2.1 Productivity

There are frozen syntactic constructions like *battle royal*, or *all of a sudden* that are usable with only a limited range of lexically specified words, but by and large, syntax is productive. A syntactic construction ordinarily has a form and a meaning that are entirely predictable on the basis of the form and the meaning of its elements and the manner in which they are joined. The only limits on whether some particular set of instantiations of the elements of a syntactic construction can be so combined are usually semantic and/or pragmatic.

However, as Chomsky (1970) pointed out, this stands in striking contrast to the situation in morphology, where there are often restrictions of an arbitrary sort, having nothing to do with meaning or import, making only certain combinations possible. There are frequently irregularities of form that arise when morphological elements are joined, and the meaning of the result often follows no general rule. Complex morphological forms are, in other words, frequently listed rather than constructed.

In her survey of noun incorporation, Mithun asserts that "N[oun] I[ncorporation] may be highly productive, but it is not free in the sense that syntactic operations can be" (1984, 889). She offers anecdotal evidence supporting the claim that in Mohawk, "speakers know not only whether a derivationally complex word is possible, but also whether they have heard it before. . . . They have no trouble understanding the new words, but they recognize that they are not part of their own (vast) lexicon" (p. 889). Thus according to Mithun, speakers of Mohawk (Greenlandic, Tiwa, etc.) must have much larger vocabularies than we do. Strangely, these very same languages are burdened with productive derivational processes that could serve to reduce the memorized vocabulary, but according to Mithun, this richness has just the opposite effect. One would think that such productivity would be made up for by a savings in storage space, rather than a requirement for more.

My own experience with a polysynthetic language, West Greenlandic, tells me that Mithun's claim is exaggerated. A great many fully transparent derivational forms are indeed lexicalized, bringing the lexicon up to roughly the same level as what we find in a less synthetic language (which after all will contain a relatively larger number of lexicalized phrases), but speakers do not have perfect memories with regard to what they have heard. The mathematical possibilities make it incredible that speakers could know every productive combination they have heard. There are approximately one hundred productive denominal verb-forming affixes in West Greenlandic. If there are (conservatively) five thousand lexical nouns in the language, there are half a million simple verbs formed from nouns. But there are about fifty fully productive noun-modifying suffixes in the language, and any of the nouns so formed can be incorporated. With just these two affix classes, then, there are twenty-five million possible forms. Taking into account the fact that verbs are also modifiable with the aid of approximately one hundred productive derivational suffixes we now have two and a half BILLION forms of a very common kind, involving only three derivational processes.

The fact is, though, that there are infinitely many possible forms involving noun incorporation, since verbs can be nominalized, and the resulting noun incorporated. Examples such as the following, demonstrating iterated incorporation, are by no means uncommon in spoken or written West Greenlandic.

(1) Apeqqutissa-qar-to-qar-poq
question-have-NOM-have-INDIC/3s
'There is someone with a question.'

The few truly productive morphological processes of English have much the same quality. Productive noun-noun compounding, for example, has yielded a huge number of transparent but lexicalized forms. From the realm of computer jargon, where such forms abound, there are examples like *disk drive*, *printer port*, *database management*, *word processor*, etc., etc. These are all part of my vocabulary and I know that I have heard them before. But what about *display shape*, *color flashing*, *printer resources*, *network chip set*, *non–von Neumann single-chip image processor*, *bar-code-reader port*, and *1-bit processor elements*, all of which (among others) I found on a single page of *Byte* (10, no. 1 [1985]: 9)? I have no idea whether I have heard any of these terms before, and even though I am writing them down and thinking about them, I seriously doubt if any is worth committing to memory.

The extreme of productivity is obligatoriness. Mithun (1984) therefore found it important to assert that noun incorporation never has this quality:

> [A]ll languages which exhibit such morphological structures [as noun incorporation] also have syntactic paraphrases. . . . The fact that productive morphological constructions of this type never exist in a language without syntactic analogs indicates that the morphologization itself must be functional. A comparison of the process across languages reveals that, in fact, speakers always incorporate for a purpose. (Pp. 847–48)

However it is clearly untrue that the productivity of morphology is always unlike that of syntax, as is shown by the fact that some morphology in some languages is the only means, or the only normal means, available for the expression of certain notions. In a great many languages, sentential negation can only be expressed by means of a verbal affix, for example. If some morphological processes are the obligatory means for the expression of certain concepts, why should noun incorporation be any different?

In fact, as was noted above, Harrington (1910) claimed obligatoriness of noun incorporation in Tiwa, a group of closely related Tanoan dialects of the American Southwest, a claim that has been recently confirmed in a revealing article by Allen, Gardiner, and Frantz (1984).

The fact of the obligatoriness of incorporation in certain languages immediately creates a problem for any theory that fails to recognize the syntactic reality of the process. It is difficult to imagine how one could even state that noun incorporation is obligatory if the morphological process is conceived of as having no connection with the syntax.

4.2.2 Referentiality

Following Postal (1969), we do not expect pieces of words to have independent referential or discourse properties and it was for this reason that I pointed out in Sadock 1980 that incorporated nominals in Greenlandic have the same kind of semantic/pragmatic status as independent nominals would be expected to have. While it is frequently the case that noun incorporation is accompanied by lack of semantic/pragmatic autonomy of the incorporated nominal (particularly where there is no indication of the syntactic independence of the nominal), it is not always so, *pace* Hopper and Thompson (1984). Those authors say, "In a number of languages, a non-referring N is bound to the stem of a V, forming a compound with the V as its head" (p. 711). And later, concerning N-N compounding, "Just as with incorporation into V's, the compounded N in such examples is non-referring; it can play no further discourse role unless it is re-introduced with full categorial status. It is insulated from reference to syntactic processes or anaphoric rules" (p. 714).

The observation that noun incorporation is more frequent when there is no independent reference to be found in the object goes back to Sapir (1911), who remarked concerning Nahuatl and Paiute:

> [I]n both languages the objective relation is more often expressed by syntactic means than by noun incorporation, the latter method being employed, it would seem, in expressing "general" or "characteristic" acts as contrasted with "particular" or "accidental" acts. (P. 267)

But it is clear that this statement was not intended to cover all cases in all languages. Sapir observed that some of the Yana examples he mentioned "seem capable of being regarded as of the 'particular' type." Indeed, his Yana texts (1910) contain examples such as the following, where an incorporated noun refers to a particular thing on a particular (though here as yet unrealized) occasion:

(2) ba´i-rusi` k!u´nusik!u wê´tk'i^{c}i^{E}
deer-will.hunt and.I.shall fetch.it.home
'He will go to hunt deer, and I shall fetch it home.'
(181, line 12; translation on 183)

In Sadock 1980, I offered two Eskimo examples of incorporated nominals that serve to introduce new topics. The examples were purposely chosen from a children's story in order to drive home the point that this is possible even in very simple, unaffected styles. So as to make it perfectly clear that this was not an aberration of the text that I chose, I offered five more examples in Sadock 1986a, which were found within the space of 223 words of connected

text in a recent book of reminiscences. Here, I will provide one more, this from the fourth and fifth lines of a novella by my friend and Greenlandic consultant, Inooraq Olsen (1980). Such usages are extremely frequent in both written and spoken Greenlandic of all styles. They are, in fact part and parcel of the language.

(3) . . . kisiannimi usi nassata-qar-punga—
but in.fact baggage-have-INDIC/1s
'but I just remembered I have some luggage—'

(4) katersoriar-lugit ingerlaannarlunga . . .
collect-CONTEMP/3p I.just.go
'I'll just go and collect them.'

Here *nassataq* 'baggage' (a count noun in West Greenlandic) is incorporated in the verb *nassataqarpunga* 'I have baggage'. Nevertheless, it is referential and specific, and introduces a discourse topic, whose reference is then picked up anaphorically in the agreement on the transitive verb *katersoriarlugit*.

Certain of the noun-incorporating affixes of West Greenlandic even allow the incorporation of definite forms, in which case, of course, there is no question of their referentiality. Two such are *-leri(voq)* 'be involved with', as in *Amaalialerivoq* 'he is very much interested in Amaalia', and *-pallap(poq)* 'to strongly resemble, behave just like,' as in *Paaviarpallappoq* 'that's just like Paavia'.

The reason that examples of incorporated nouns with high referentiality and discourse salience are so frequent in West Greenlandic is that many of the most ordinary verbs that would be used for introducing a new topic—"to have," "to get," "to make," "for there to be," etc.—occur only as denominal suffixes. Any circumlocutory method of avoiding them would carry a heavy load of Gricean implicatures stemming from the abnormality of the construction. (Cf. the discussion of the difference between *pink* and *light red* in McCawley 1978.)

Thus we should not be surprised that in Tiwa, where incorporation is sometimes grammatically obligatory, the incorporated nominal can also introduce a discourse topic. Even in the very brief Tiwa text that Harrington recorded in 1910, there is one very clear instance of this, as I pointed out in Sadock 1986a. The internal nominal is an inanimate object, an argument type that is obligatorily incorporated. (See also Allen, Gardiner and Frantz 1984 and Sadock 1985b.)

The story concerns Old She-Wolf and Old She-Deer. She-Wolf goes to where She-Deer lives and suggests to her that they live together. One day, as they are out gathering wood, old She-Wolf playfully bites Old She-Deer, who

realizes that the reciprocal arrangement is fraught with danger. She tells her little ones that she might soon be killed by the wolf and warns them that if the wolf brings them some pieces of meat and if, when they roast it, it makes a sizzling sound, they should not eat it. This last part goes as follows, in Harrington's notation and gloss:

(5) Hu xu 'ăĭxä^{n}n hi ja
So then in case perhaps hither
'uma^{n}na^{n}m-tŭă-ka^{n}la^{n}
them.2+INAN.she.for.you.2+-meat-brings
n, ma^{n}n-xa-k'ŭĭ .t^{s}ɬɬi-m-ɬŏ ja-mä^{n},
when you.2-roast-put s-s-s.sound-sizzle. goes,
'ăĭta^{n}n xu ma^{n}n-na-k'al-pu^{n}
in.that.case then you.2-not-eat-shall

(Here reference to the meat, introduced as an incorporated stem, is continued in terms of zero-pronominal anaphors, the normal anaphoric device for definite third-person subject and object reference (Harrington 1910). Harrington lists the root *xa* as a verb, so the reference in the next two verbs must also be to the meat.)

Thus we must conclude that there is only a weak relation in general between incorporation of nominals and loss of referentiality. In those cases where noun incorporation gives evidence of its syntactic nature, for example by being obligatory under certain circumstances, there is very little in the way of semantic bleaching. Logic decrees that it could hardly be otherwise. If the conditions that make noun incorporation unavoidable had nothing to do with the referential properties of the incorporated nominal, and incorporated nominals were always lacking in reference, the language would be unable to express certain essential things. Since in Tiwa nonhuman objects must be incorporated, then one could not say the equivalent of "I bought a shirt," where the reference is supposed to be to a specific shirt!

4.2.3 Syntax

Neither the fact that noun incorporation is sometimes as productive as syntactic phrase formation, nor the fact that it does not necessarily create referential islands is really proof of its syntactic nature. There is no reason in principle why morphological operations must be limited, unless they are always restricted to the lexicon (as a list of occurring forms), a proposition that I think no one today would subscribe to. If some morphology is fully productive, why not noun incorporation? Similarly, no theory of grammar that I am aware of automatically requires that proper parts of words have less robust semantics than whole words or phrases. Thus Postal (1969), who first described the

anaphoric islandhood of words, had to add a special stipulation to his theory, which otherwise made exactly the wrong predictions.

The crucial evidence for the syntactic nature of certain varieties of noun incorporation would have to come from syntax itself. A variety of testable predictions flows from the assumption that a proper part of a word actually represents a syntactic formative, the most important of which are the following:

1. The presence of a word-internal syntactic formative should be inconsistent with the presence of an external syntactic formative in the same presumed syntactic role. For noun incorporation in particular, a word-internal noun stem that counts as, say, the direct object of the clause (or the head of the direct object) should be mutually exclusive with the presence of an external direct object (or one with an overt head N). Thus instances of noun incorporation with so-called "doubling" of the object in the form of an external NP clearly cannot be instances of syntactic noun incorporation, while those where doubling has not occurred might be, and those where doubling is impossible would seem to be.
2. The incorporation into a word of a piece of an external phrase ought to leave behind, or "strand," other possible constituents of the phrase. The stranded elements ought to have precisely the morphosyntactic form that they would have if the incorporated word were in the phrase, and to the extent that semantic organization mirrors syntax, the external phrase ought to be understood as having the incorporated element in it. In the case of noun incorporation, tell-tale remnants of the external NP ought to be left behind, including adjectives, relative clauses, possessors, articles, and so on.
3. Most important, the result of incorporation might yield a syntactic pattern that does not otherwise exist in the language. In other words, the word containing the incorporated element might have surface syntax distinct from any underived word in the language. As for noun incorporation, this could mean, for example, that the remnant left behind by incorporation might not be a phrase that occurs otherwise in the language in that particular syntactic configuration. The existence of otherwise impossible arguments of verbs that contain incorporated nouns is a sure sign of the syntactic relevance of the phenomenon.

Let us consider first the possibility of doubling objects under noun incorporation. It is well known that this is sometimes allowed and sometimes not. West Greenlandic and Southern Tiwa are both languages in which an overt copy of the incorporated noun in the position of the argument that the noun is supposed to represent is ungrammatical. If the incorporated nominal is taken to BE the direct object, then this restriction is immediately explained, whereas

if it is not taken to be the syntactic object, some other sort of restriction must be invented to account for the lack of doubling. Rosen (1989), for example, suggests the following sketchy account for the failure of doubling of an incorporated N in certain languages:

> A second possible explanation of the lack of doubling within the lexical approach concerns the selectional restriction placed on the verb by the incorporated noun. . . . It is possible that the selectional restrictions placed on the verb in Southern Tiwa and West Greenlandic exemplify the other extreme: The head of the direct object must not duplicate any of the information in the incorporated noun. The restriction may be so strict as to rule out any instance of doubling. (P. 307)

This restriction, if actually formalizable at all, is first of all ad hoc, and second, quite unlike selectional restrictions of the familiar sort. Ordinary selectional restrictions that predicates impose are not such as to disallow the repetition in arguments of information that is predictable from the meaning of the verb. A sentence such as "My wife is pregnant" is more normal, in fact, than "My spouse is pregnant," even though the femaleness of the subject is selected for by the predicate. There are even verbs like *devein* and *diagonalize* that exert selectional pressures on their objects so strong that one can practically predict what head noun will occur. The first of these requires objects in the category of shrimp and prawns, the second requires mathematical objects called matrices. Yet they certainly allow overt expression of an object that just duplicates the information implied by the selectional restriction: *devein shrimp*, *diagonalize a matrix*. Selectional restrictions are just not the right sort of mechanism to appeal to in accounting for the lack of doubling of incorporated arguments in certain languages.

Rosen is, I believe, on the right track when she suggests that in some languages the incorporated noun merely restricts the meaning of the verb, but in Southern Tiwa and West Greenlandic "the incorporated noun carries with it a full specification of noun head features." This would follow if the noun occupied a syntactic position in the sentence. Rosen, however, gives no reason why this should be so in some languages and not in others. Furthermore, just saying that the noun is rich in nominal features is not enough to account for the lack of doubling, since there is no connection between this "richness" of the internal noun and the direct object slot outside the verb. There is obviously nothing wrong with having an argument in a DIFFERENT syntactic position that duplicates an incorporated object, e.g., Greenlandic *Meeqqat meeraqarput* 'Children have children'. So the lexical account requires some further statement (presumably involving either thematic or grammatical relations to be appended to the treatment in order to get the facts right. If, on the other hand, the incorporated nominal is taken actually to be the object, the mutual exclusivity of the appearance of an incorporated noun and a full direct object falls out directly.

A lexical account thus provides no real explanation for the failure of doubling in languages like Southern Tiwa and West Greenlandic, and must resort to ad hoc and unmotivated statements to account for it. Viewed autolexically, undoublable incorporated N's are (parts of) syntactic arguments, and doubled incorporated N's cannot be (parts of) syntactic arguments.

4.2.4 Stranded Elements

I turn next to the question of stranded elements. Grammatically or pragmatically obligatory noun incorporation virtually demands that modifiers of the incorporated nominal be allowed to occur outside of the verb form, lest the language's expressive power be intolerably diminished.

The first question bearing on the syntactic reality of the incorporated noun is whether the form of the external argument to an incorporating verb is precisely what one would expect if a single element, namely the head noun, were missing. In Greenlandic this is clearly borne out by several facts. For example, only the head noun of a Greenlandic NP may be inflected for possession; modifiers, though formally nominals, are inflected only for case and number.

(6) qatannguti-n-nik
sibling-1s-INST
'my sibling (INST)'

(7) qatanngutinnik arna-mik
sibling-1s-INST female-INST
'my sister (INST)'

(8) *qatanngutinnik arna-n-nik
sibling-1s-INST female-1s-INST

This restriction extends to modifiers of incorporated nouns. Just like modifiers of overt head nouns, they may not be possessed:

(9) Arna-mik qatanngu-seri-voq.
female-INST sibling-be.occupied.with-INDIC/3s
'He is occupied with (someone's) sister.'

(10) *Arna-n-nik qatanngu-seri-voq.
female-1s-INST sibling-be.occupied.with-INDIC/3s
presumably: 'He is occupied with my sister.'

Another indication of the status of the stranded elements, discussed in detail in Sadock 1980 and repeated in Sadock 1986a, is this: incorporating

verbs, and only these, may impose restrictions on the formal plurality of an external NP, which they do when the internal N is lexically specified as to formal plurality. Though some verbs may select a semantically plural complement (e.g., *katersor(pai)* 'collect (them)'), only incorporating verbs select for SYNTACTIC plurality. For example, the noun *qamutit* 'sled', is formally plural, being derived historically from a no longer extant root meaning 'sled runner'. Despite the fact that it is now semantically singular (one can talk about one sled), its formal plurality is reflected in the fact that it demands plural agreement. When incorporated, the stem still requires a formal plural outside of the verb, construed as a modifier of the incorporated noun:

(11) Ataatsinik qamuteqarpoq.
ataaseq-nik qamut-qar-poq
one-INST/p sled(p)-have-INDIC/3s
'He has one sled.'

(12) *Ataatsimik qamuteqarpoq.
ataaseq-mik qamut-qar-poq
one-INST/s sled(p)-have-INDIC/3s
'He has one sled.'

But even in languages where incorporation is not obligatory, it is sometimes the case that constituents outside the verb are understood as modifying the internal nominal.

This possibility is accounted for by proponents of a lexical account of all incorporation (e.g., Mithun 1984; Di Sciullo and Williams 1987; and Rosen 1989) by assuming that in some languages the incorporated noun serves as a sort of "classifier," restricting the range of sortal applicability of the verb, without actually specifying the object. On this analysis, an external argument to an incorporating verb represents a more specific entity of the kind selected for by the incorporated noun stem. The idea is that the incorporated noun can be taken as a qualifier of the activity expressed by the verb, and the external argument as the real argument, the effect of modification of the incorporated nominal being a product of the semantics rather than the syntax. Analogously, a sentence like *Three senators are pregnant* is understood as referring to a group of three women senators, but there is no need to postulate an NP containing "woman" in the syntax of the sentence. The predicate *pregnant* selects for female subjects, so whatever the subject properly refers to, it must also be understood as referring to females.

While such a semantic theory might be plausible in the case where the incorporated nominal is quite general and the external argument is necessarily a more specific entity of the same kind, as in English *compose a cantata*, *sing an aria*, *drink beer*, and so on, it is much less plausible where the two sets

have no necessary relationship to one another. This is the case in (13), a Mohawk example, taken from Mithun (1984, 870) and in general where the external nominal is a quantifier, as in Mithun's Caddo example (14), the Greenlandic example (11) above, or the Southern Tiwa example (15), taken from Allen, Gardiner, and Frantz (1984, 297). It is not at all clear that "polka-dotted" in (13) is more specific than "dress," and it is surely not the case that "a lot" in (14) and "two" in (15) are types of grass and cats, respectively. The classifier theory of incorporation in such cases requires a profoundly different theory of quantification from any that exists at present, whereas the syntactic theory can rely on ordinary quantification theory, since (14) and (15) would contain phrases meaning "much grass" and "two cats"—phrases that exist with the correct meanings in the languages in question anyway.

(13) Kanekwarúnyu wa'-k-akya'tawi'tsher-ú:ni.
it.dotted.DIST past-I-dress-make
'I made a polka-dotted dress.'

(14) Wayah hák-k'uht-'í'-sa'.
a.lot PROG-grass-be/grow-PROG
'There is a lot of grass.'

(15) Wisi ibi-musa-tuwi-ban.
two AGR-cat-buy-PAST
'They bought two cats.'

Mithun, Rosen, and others who wish to maintain a lexical account of noun incorporation analyze examples like (13)–(15) as containing headless, or empty-headed nominal objects. But empty headed arguments generally require a discourse antecedent in every language that I have researched.[1] Thus English *I sang three*, or German *Ich habe ein Schönes gesungen* are quite grammatical, but cannot begin discourses, even though we know from the meaning of the verb what sort of entity the object must be. The class of songs (or Lieder) must already have been introduced into the discourse for these sentences to occur appropriately.

Now in Greenlandic, this is also by and large true. Thus a sentence like (16) can only be used appropriately if the sort of thing that the object is supposed to refer to has already been established as a discourse topic.

(16) Marlunnik nerivunga.
marluk-nik neri-vunga
two-INST/p eat-INDIC/1s
'I ate two.'

The sole exception that I am aware of (other than that mentioned in note 1) is provided by noun incorporation. An example like (17) can be used to initi-

ate a discourse. It can be used, in other words, under exactly the same circumstances as (18), which has a free object.

(17) Marlunnik ammassattorpunga.
marluk-nik ammassak-tor-punga
two-INST/p sardine-eat-INDIC/3s
'I ate two sardines.'

(18) Ammassannik marlunnik nerivunga.
ammassak-nik marluk-nik neri-vunga
sardine-INST/p two-INST/p eat-INDIC/1s
'I ate two sardines.'

Mithun's 1984 glosses suggest that exactly the same collection of facts characterizes even Mohawk. Her gloss of (13) as 'I made a polka-dotted dress' suggests that no prior antecedent is required, while her gloss of the syntactically parallel (19), which includes *one* suggests the opposite.

(19) Kanekwarúnyu wa'katkáhtho.
it.dotted.DIST PAST.I.see
'I saw a polka-dotted (one).'

The upshot of all of this is that it is simply not the case that existing semantic rules will account for the interpretation of sentences with external modifiers of incorporated nominals. We would minimally require a special rule of interpretation that gives a complete interpretation to free quantifiers and adjectives as complements of all and only verbs containing incorporated nominals.

A rather different case demonstrating the strict parallelism that can characterize the interpretation of incorporated nouns and free syntactic alternatives can also be found in West Greenlandic. A few affixes incorporate not the head noun, but specifically a quantifier. Thus *marloraarpoq* means 'He caught two', (<*marluk* 'two'). But if the quantificational expression contains more than one phonological word, e.g., *marluk affarlu* 'two and a half', part of the expression is stranded by incorporation:

(20) Marlo-raarpoq affar-mil-lu.
two-he.caught half-INST-and
'He caught two and a half.'

(Note the instrumental case of the stranded expression, for which see Sadock 1980.)

There are also a few nominal modifiers in West Greenlandic that have a different sense when used as modifiers than they do when used as head nouns. As Anthony Woodbury (1989) has observed, this includes the class of material terms like *qisuk* 'wood', *marraq* 'clay', and so on. When used directly as modifiers, these terms describe the material of which something is largely made: *illu qisuk* 'wooden house', *tiitorfik marraq* 'china cup'. When used as independent nominals, however, these terms designate an aggregation of the substance itself, not something made of it. *Qisummik takuvunga* means 'I saw a piece of wood', and cannot possibly be used to mean 'I saw something made of wood'. When used as modifiers of incorporated nouns, such expressions always have their modificational sense, not their independent nominal sense:

(21) Qisum-mik illo-qar-poq
wood-INST house-have-INDIC/3s
'He has a wooden house.'

Here again, the nonsyntactic account of the meaning of external modifiers that takes them to be ordinary arguments and intersects that meaning with the selectional restrictions imposed by the verb will not work. But the assumption that there is a syntactic phrase *illumik qisummik* in (21) gets the meaning right automatically, since that phrase means 'wooden house'. Furthermore, since any intransitive verb in West Greenlandic that means roughly what the suffix *-qar* means takes its object in the instrumental case, and since modifiers agree with their heads in case, the form of (21) is predicted as well. (See Sadock 1980; 1985a for details of this aspect of the syntax of West Greenlandic.)

The semantics of the classifier theory of the meaning of noun incorporation clearly needs to be complicated in various ways in order to account for examples like these. But that semantic theory collapses utterly, it seems to me, where the external modifier is what in the semantic literature is called an intensional adjective. Thus while a polka-dotted dress is a dress, counterfeit money is not money at all. Yet adjectival nouns with meanings like 'counterfeit' can perfectly well occur external to an incorporating verb in Greenlandic, with precisely the same semantic effect that they have when they occur as modifiers of independent nouns.

(22) peqquserluutinik aningaasiortoq[2]
peqquserluut-nik aningaasaq-lior-toq
false-INST/p money(p)-make-NOM
'one who makes false money, a counterfeiter'

In such a case, the suggestion that the incorporated modifier simply restricts the range of arguments of the verb fails. Finally, it is quite unclear how the classifier theory is supposed to work in the case of incorporated interrogatives, such as the sampling of Greenlandic examples below.

(23) Kinaavit?
'Who are you?'

(24) Sutorpit?
'What did you eat?'

(25) Susivit?
'What did you buy?'

(26) Susunnippa?
'What does it smell like?'

While the interpretation of stranded modifiers is strongly suggestive of syntax, it is not entirely convincing as a demonstration of the syntactic reality of the incorporated noun for the simple reason that interpretation is a matter of semantics. If, as I shall argue in chapter 6, semantic representation is profitably viewed as an autonomous level, parallel to, but not the same as syntax, then what has really been shown is that the incorporated N can, in certain languages, and under certain circumstances, be a SEMANTIC argument of the verb. The key fact that argues for the syntactic reality of some incorporated nouns is that the remnant of an argument phrase left by incorporation is sometimes not a possible argument in the language at all. In general, any unique syntactic property of verbs containing incorporated nouns, as compared with all other verbs in a language, counts as an extremely strong demonstration of the syntactic reality of incorporation, particularly if the unique structure can be reduced to ordinary syntax coupled with ordinary morphology, as in the cases discussed above.

In my 1980 article I laid great emphasis on the fact that the external syntax of Greenlandic verbs with incorporated nominals is not identical to that of basic verbs of the language. I repeated these arguments in Sadock 1986a, adding a new one. Because of their centrality to the argument at hand and to the general thesis of this book, I will rehash them here, too.

Perhaps the most important fact is that incorporating verbs (and no others) may have external possessors in the ergative case.

(27) kunngip panippassuaqarpoq[3]
kunngi-p panik-passuaq-qar-poq
king-ERG daughter-many-have-INDIC/3s
'There are many king's daughters (i.e., princesses).'

Here the external ergative NP *kunngip* is taken as a possessor of the incorporated nominal *panik*, exactly as in the NP *kunngip pania* 'princess'. What is important about this particular construction is that otherwise, ergative NPs CANNOT POSSIBLY SERVE AS ARGUMENTS OF INTRANSITIVE VERBS. Unlike English, and many other languages, a possessor can never be construed as such without an overt possessum. Therefore, the following example is completely

ungrammatical, and cannot mean, for example 'I saw the King's (something or other)'.

(28)) *kunngip takuvunga
kunngi-p taku-vunga
king-ERG see-INDIC/1s

There is one transitive noun-incorporating affix in West Greenlandic, *-gE*, an extremely frequent and characteristic morpheme which is usually translated 'have it as N'. A much better translation of a sentence of the form NP[ERG] NP[ABS] N+gE is "NP[ABS] is NP[ERG]'s N":

(29) Jerry-p Biinia erneraa.
erneq-gE-vaa.
Jerry-ERG Ben son-gE-INDIC/3s/3s
'Jerry has Ben as a son.' or
'Ben is Jerry's son.'

Now if the incorporated noun in such a sentence is possessed, stranding the ergative possessor, there results a sentence with TWO ERGATIVES, something that occurs nowhere in the language outside of incorporation constructions.

(30) Hansi-p puisi-p neqaati-gi-vaa.
Hans-ERG seal-ERG cache.of.meat-gE-INDIC/3s/3s
'It is Hans's cache of seal meat.'

If the incorporating verb is transitivized, say by the addition of a causative morpheme, it is also possible to produce clauses with two ergatives, one the subject, and the other the possessor of the incorporated noun. Once again, such examples are not even grossly like anything that occurs independently of noun incorporation in West Greenlandic.

(31) Hansi-p qimmi-p ame-qar-tip-paa.
H.-ERG dog-ERG skin-have-cause-INDIC/3s
'Hans let him have (gave him) a dog's skin.'

Impossible arguments understood as modifiers of incorporated nouns are also found in the South Munda language Gta?, which was brought to my attention by Norman Zide. According to Zide's field notes, the directional adjective *mb&?sia?* 'left' never stands alone, but is always combined with the word for 'hand'. All by itself it does not constitute a NP meaning 'the left one'. But the noun meaning 'hand' can be incorporated, stranding what is an unambiguous modifier in the language:

(32) n&N mb&?sia? gwe?-ti-ke[4]
I left wash-hand-PAST
'I washed (my) left hand.'

Returning to West Greenlandic, a second feature that gives rise to otherwise impossible sentence patterns (at least in the older language)[5] is found with verbal affixes that incorporate a predicate nominal (Sadock 1980; 1986a). Such verbs, and no others, may have an additional absolutive case NP associated with them that obligatorily follows the verb. Thus such sentences have two absolutives, one the subject, and the other a modifier of the incorporated predicate nominal.

(33) Joorut palasi-u-voq tusaamasoq.
Jorgen(ABS) priest-be-INDIC/3s famous(ABS)
'Jorgen is a famous priest.'

The final example of a unique configuration in which incorporating verbs of West Greenlandic are found is what we might call polysynthetic gapping.

(34) Maani amerlasuunik qaqqaqarpoq
maani amerlasuu-nik qaqqaq-qar-poq
here many-INST/p mountain-have-INDIC/3s
'Here there are many mountains

aamma tatsinik amerlasuunik.
aamma taseq-nik amerlasuu-nik
also lake-INST/p many-INST/p
as well as many lakes.'

Here the verb is understood as occurring in a second conjunct, just as in the case of gapping in more familiar langauges. What is remarkable here is that the gapped verb is morphologically a derivational suffix and not an independent word.

For reasons that I do not understand, this sort of gapping requires a quantifier in both conjuncts.[6] The enabling quantifier can itself be incorporated if the affix is one of those that specifically incorporates quantifiers (for which see the discussion surrounding (20) above).

(35) Marlo-riar-punga pingasu-nil-luunnit.
two-do-INDIC/3s three-INST/p-or
'I did it two or three times.'

In all three of these cases of a unique structure that occurs with incorporating verbs, the configuration we find arises naturally from the assumption that the incorporated noun is syntactically part of an argument phrase.

First of all, notice that removing the possessum from a phrase like *kunngip pania* 'king's daughter (princess)' would strand the ergative *kunngip*, giving rise to just the sort of sentence exemplified by (27).

Next, consider the construction represented by (36), with a postverbal absolutive understood as a modifier of an incorporated predicate nominal. Besides incorporated structures such as this, there are double-absolutive predicate-nominal sentences of West Greenlandic that contain nonverbal copulas, such as the demonstrative *tassa* or the cliticized pronoun *una*:

(36) Joorut tassa palasi tusaamasoq.
Jorgen(ABS) that priest(ABS) famous(ABS)
'Jorgen is the famous priest.'

(37) Joorun-una palasi tusaamasoq.
Jorgen(ABS)-it priest(ABS) famous(ABS)
'It's Jorgen who is a famous priest.'

The assumption that the incorporating suffix *-u* is the verbal analog to the nonverbal copular elements in (36) and (37) immediately explains the form of (33), including the fact (which is unexpected in this SOV language) that the modifier of the predicate element must follow the verb. In (36) and (37) the order is also fixed, with the predicate following the copular element.

Lastly, gapping of a full verb (to the right only, despite the word order of the language (cf. Ross 1970) is possible in West Greenlandic:

(38) Maani qaqqanik amerlasuunik takuvoq
maani amerlasuu-nik qaqqaq-qar-poq
here many-INST/p mountain-have-INDIC/3s
'Here he saw many mountains

aama tatsinik amerlasuunik.
aamma taseq-nik amerlasuu-nik
also lake-INST/p many-INST/p
as well as many lakes.'

Thus the unusual sentence pattern in (34) would likewise be explained if its syntax were the same as that of a comparable sentence with an independent verb, such as we have in (38).

4.2.5 Summary

Because the point seems to have been misunderstood (e.g., by Rosen 1989), let me first make clear what I have NOT done. I have not proved that everything that has been termed noun incorporation in the literature is always and only a morphological process that interacts with syntax. What I have shown is that surely in West Greenlandic, and most likely in Southern Tiwa and Gta?, a noun stem contained within a morphological verb sometimes has all the syntactic and semantic properties of an independent head noun. Even in these languages, though, there is nothing in the argument to suggest that every mor-

phologically included noun is syntactically real. Indeed, I know that is not the case in West Greenlandic, where there are a great many lexicalized incorporations that give no evidence whatsoever of interactivity with the syntax.

As an example, consider the lexicalized form *nunaqarpoq* 'to dwell in a certain locality', which is transparently *nuna* 'land' plus *-qar(poq)* 'have'. Despite the morphological composition of the word, doubling of the incorporated noun is perfectly allowable in this case:

(39) Nuna-mi tassa-ni nunaqarpoq.
land-LOC that-LOC he.lives
'He lives on that land.'

Note also that the argument of this verb is not in the instrumental case, which is what we would expect of a modifier of the incorporated noun, but rather in the locative, as befits the meaning of the verb. Furthermore, unlike the other examples of the use of *-qar(poq)* discussed above, the morphological noun in this case does not set up a discourse topic.

(40) *Kalaallit-Nunaanni nunaqarpoq, kisianni takunngilara.
Greenlanders-land.LOC he.lives but I.have.not.seen it.

The existence of items like this, with all the expected properties of syntactically inert derivational morphology makes it that much clearer that the syntactic reality of the productive examples has to be recognized.

What I have demonstrated beyond any doubt, I believe, is merely that there exists some morphological structures whose parts must be visible to the syntax. I use the word "merely" with hesitation, because theories of grammar allowing for the manipulation of syntactic elements by the morphology (and vice versa) will inevitably turn out to be very different from those that do not.

There is a precedent for this, of course, in cliticization, where at least the host to which the clitic attaches regularly counts as a syntactic formative, i.e., an element of the syntactic parsing. In chapter 3 I showed that a grammar allowing simultaneous and possibly discrepant syntactic and morphological bracketings yielded an elegant and empirically successful theory of cliticization. What I shall turn my attention to next is an account of incorporation along similar lines.

4.3 Incorporating Incorporation

I will use the word "incorporation" very generally to refer to any morphological process (other than cliticization) that produces morphological units in which some of the morphological constituents can be shown to have independent syntactic reality. The term is neutral as to the morphological technique that is employed: affixation, symbolic change, reduplication, and compounding could all potentially provide examples of incorporation. Likewise, it is of no consequence to this definition whether we would call the morphological operation in question derivation or inflection; either could provide examples of incorporation. The definition also does not imply anything about the categories involved, either in the morphology, or in the syntax. Nouns, verbs, adjectives, prepositions, and so on, could all participate in incorporation. The West Greenlandic phenomenon that was discussed in detail in the foregoing sections was just that particular instantiation of the general case where a syntactic nominal is the derivational base for a suffixal verb-forming morphological process.

According to this construal of the term, incorporation is, like cliticization, a phenomenon with both a syntactic and a morphological face, for in both, a proper subpart of a word-sized unit counts as an independent element of the syntax. Both phenomena, then, involve lexemes that are realized as morphology of some kind, and yet have the syntactic status of formatives. In all cases that I know of, both clitics and incorporating lexemes are syntactic functors, i.e., syntactic atoms that combine with phrases to form larger phrases, or more simply, lexical heads.

In the foregoing sections of this chapter, I have tried to provide arguments necessitating a syntactic treatment of certain nonclitic morphological operations by paying attention to the details of their behavior. An entirely different sort of argument, which is almost as persuasive, regards the relation of incorporative morphology to other kinds of morphological juncture, in particular, cliticization.

As I document in detail in Sadock 1988, there are no clear boundaries between the two phenomena. Bound morphemes that are syntactically active shade by degrees from those that display the constellation of properties that we associate with clitics through incorporators, to those that are typical of syntactically inert morphemes. In fact, all of the traditional classes of morphology—derivation, inflection, and cliticization—have unsharp boundaries on all sides, as several of the examples discussed in chapter 5 will serve to show.

One of the important lessons to be drawn from the existence of such clines is that it is impossible to consign the various traditional morphological types to different areas of the grammar. In that they shade into one another, all types of morphology must be dealt with in a single component. Their differences lie partly in their associations with other autonomous dimensions of grammatical organization. In answer to Stephen Anderson's (1982) question, "Where's morphology," I would not respond as did Jensen and Stong-Jensen (1984) that it is in the lexicon, but rather, "It's in the morphological component, where it belongs."

But a lesson of more immediate concern to us here is that there is no single

criterion that sets cliticization apart from other morphological operations. Thus the syntactic activity of clitics should not be viewed as in and of itself criterial. For an element to be comfortably labeled a clitic, it should display a whole host of phonological, morphological, semantic, AND SYNTACTIC properties that are, in principle at least, independent of one another, so where a grammatical phenomenon conforms to only part of this list, we should not demand a preexisting pigeonhole to put it in.

Nevertheless, there seem to be relatively stable associations of properties from various components of the grammar that correspond closely to the traditional categories. Indeed, there is a frequently observed set of behaviors from various modules that seems to characterize one particular intermodular phenomenon that we may call incorporation, and another that we may call cliticization. In chapter 5 I will survey a number of disparate examples from an assortment of languages that serve to illustrate both that the association of certain intermodular clusters of properties occurs with much greater than chance frequency, and also that the clustering is not necessary.

Incorporation and cliticization share a number of features, including the fact that both compose formatives of the syntax into morphological structures, that the pivotal element—the clitic or incorporator—is a lexical head in the syntax, and that it has robust semantics, including referentiality, where that is applicable. The two principal differences between canonical cliticization and canonical incorporation are these:

1. Clitics apply morphologically to full words, but incorporating lexemes attach morphologically to stems.
2. Clitics attach to syntactically nearby words, but incorporators attach to heads of syntactic phrases, regardless of proximity.

The attachment of clitics to nearby words was handled in chapter 3 by the Linearity Constraint, which requires that the elements of the morphological string and the elements of the syntactic string occur in the same order. As was demonstrated in that chapter, however, individual clitic morphemes, such as Latin *-que*, could be lexically exempt from the LC, and occur out of order in the string of morphemes with respect to their expected positioning in the syntax. But the fact that clitics that violate the LC attach as soon as they can, while still fulfilling their morphological obligations (i.e., as second position, or penultimate position clitics) indicates that the LC is still partially in effect.

Thus there are actually two versions of the LC—a strong one requiring absolute identity of ordering of matching elements in morphology and in syntax, and a weaker one, allowing minimal discrepancies of order, but otherwise requiring the order of elements in the two representations to be the same.

Now incorporation, as mentioned in several places above, and as will be documented below, involves heads of syntactic phrases (cf. Baker 1988; Marantz 1984; Hoeksema 1984), and is thus at odds with the Linearity Constraint,

since there is no guarantee in the abstract that a lexical head and the head of its complement will occur in adjacent positions in syntax. The Linearity Constraint might thus be overridden by the requirement of headship that governs cases of incorporation.

Recall from chapter 3 that there is a trade-off between the LC and the CIC in that both may not be violated at the same time. When what we now see to be the strong form of the LC is violated by a particular clitic, the CIC must be obeyed, and the clitic will attach to an element of its syntactic co-phrase.

Turning our attention to incorporation, we may notice that the balance between the two types of homomorphisms generalizes. The total relaxation of the Linearity Constraint that is typical of incorporation is counterbalanced by a strengthening of the CIC. With regard to clitics, the CIC requires only that they preferentially attach to some unspecified element of the phrase that they are in syntactic construction with; with regard to incorporators, it demands attachment to a uniquely identifiable, central element of that phrase, namely its head. Functionally, this trade-off between the two homomorphism conditions on the syntax-morphology interface assures that the syntactic function of a mismatched element is reconstructible.

So just as there are strong and weak conditions on the degree of correspondence of the linear order of elements between representations, there are strong and weak conditions on the degree of structural correspondence between syntax and morphology:

(41) Linearity Constraints
   a. Strong
      The associated elements of morphological and syntactic representations must occur in the same linear order.
   b. Weak
      The associated elements of morphological and syntactic representations must occur in as close to the same linear order as the morphological requirements of the lexemes allow.

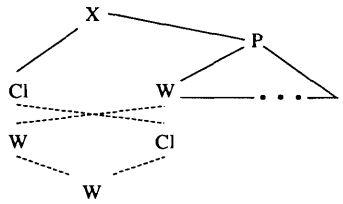
(42) Constructional Integrity Constraints
   a. Strong
      If a lexeme combines with a phrase P in the syntax and with a host in the morphology, then the morphological host must be associated with the head of the syntactic phrase P.
   b. Weak
      If a lexeme combines with a phrase P in the syntax and with a host in the morphology, then the morphological host must be associated with some element of the syntactic phrase P.

We may now think of the limitation on allowable discrepancies between syntactic and morphological representations in the following terms:[7]

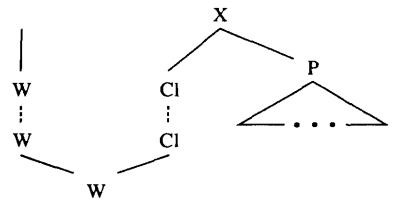
(43) Morphosyntactic Homomorphism Condition
Let each of the strong homomorphism constraints count as two degrees of similarity, and let the weak homomorphism constraints count as one. The total degree of similarity betwen autonomous syntactic and morphological representations must be at least two.

Thus if a particular morphosyntactic association violates the weak Linearity Constraint (and hence also the strong Linearity Constraint), it must obey the strong version of the CIC and count as incorporation as the term is used here. The types of clitics that violated the Linearity Constraint of chapter 3 still obeyed the weak Linearity Constraint, and the CIC which they obeyed was also the weak variety. Those clitics that did not obey the CIC at all, obeyed the strong form of the LC. These two varieties of cliticization are contrasted with incorporation in the following diagrams.

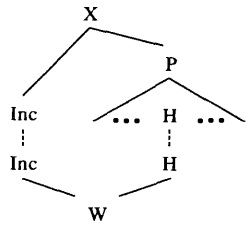
(44) Nonlinear Cliticization
Obeys weak LC and weak CIC.



(45) Nonconstructional (Anticipatory) Cliticization
Violates weak CIC and obeys strong LC.



(46) Incorporation
Violates weak LC and obeys strong CIC.



The formulation in (43) sets a lower bound on the degree of morphosyntactic similarity, but is quite consistent with the existence of associations of greater similarity than this minimum. The simple clitics of chapter 3, for example, obey the strong LC and the weak CIC. It is also possible for a lexeme to have both a syntactic and a morphological representation and to conform to both of the strong homomorphism conditions (41a) and (42a).

This situation in fact frequently occurs in languages that conform rather strictly to Greenberg's Type III class (Greenberg 1966a), i.e., verb-final languages in which the heads of nested phrases are adjacent. In such languages, postpositions will immediately follow the head noun of their object NPs, auxiliaries will immediately follow the main verb of their complement VPs, complementizers will immediately follow the finite verb of their clauses, and so on. Function lexemes in languages where lexical heads are uniformly final in their phrases strongly tend to be phonologically dependent suffixes on the preceding word, which is both the head of their own syntactic co-phrase, and immediately before them in syntax. Such bound morphemes thus obey both the strong LC and the strong CIC. They are simultaneously clitics and incorporators, at least morphosyntactically.

But as mentioned at the outset of this section, there is something else besides the question of whether or not the strong CIC is obeyed that sets incorporation apart from cliticization. In the data that I have examined, the great majority of lexemes that conform to the strong CIC combine morphologically with stems, rather than fully inflected words. Such lexemes are derivational or inflectional affixes (or morphological modifications) or pieces of compounds. On the other hand, the great majority of lexemes that conform to the weak (or the strong) LC, and not the CIC, combine morphologically with inflected words. Such lexemes are classical clitics.

These correlations are good, but as we shall see in chapter 5, they are not perfect. I will state them as morphosyntactic defaults which can be overridden by lexical information that imposes requirements that differ from what these principles would otherwise demand. In the absence of such overriding features, a lexeme will have some of its morphosyntactic properties determined on the basis of others according to the following two principles.

(47) Incorporation Principle
If a lexeme combines with a stem in the morphology and with a phrase in the syntax, its morphosyntactic association will conform to the strong CIC.

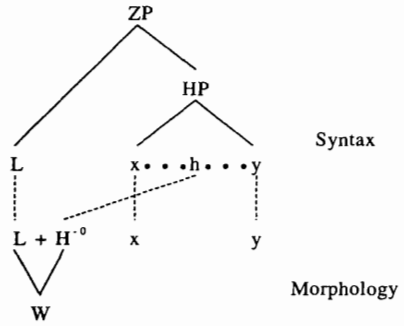
(48) Cliticization Principle
If a lexeme combines with an inflected word in the morphology and with a phrase in the syntax, its morphosyntactic association will conform to at least the weak LC.

If a language in which syntactic heads are regularly adjacent lacks clear inflectional morphology, as is the case, for example, in Japanese and Korean, then the distinction between cliticization and incorporation collapses, since the definitions above refer to inflection. In fact, it is sometimes not possible to distinguish sharply between either of these kinds of morphology and inflection, especially where the morpheme in question is syntactically obligatory. Inflectional morphology is distributed according to the Head Feature Convention (Gazdar et al. 1985), and will thus wind up in the same place as either a phrase-final clitic or a head-incorporating morpheme in a language of this type. The collapse of such distinctions is understandable in a model in which all morphology per se is attributed to a single component, but would be surprising in any system that segregated the various types of morphology into separate subcomponents.

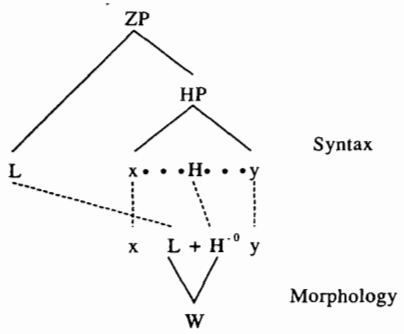
4.3.1 The Direction of Incorporation

Through the strong CIC, the Incorporation Principle requires that when a lexeme L combines syntactically with a phrase HP and morphologically with a stem H^{-0}, then H^{-0} is associated with the head of HP in the syntax. There are two ways that this could happen. Either the complex lexical expression $L+H^{-0}$ occurs in the position that L itself would be expected to occupy in the syntax, as in (49), or $L+H^{-0}$ occurs in the position that H, the syntactic correspondent to the stem H^{-0}, would be found in, as in (50).

(49)



(50)



In Sadock 1985a I attempted to approach the question of the direction of incorporation by counting the number of crossing association lines between the two representations. The idea was that a diagram with fewer crossing lines was to be preferred to one with more. I now think that this effort, which met with only limited success anyway, was misguided. Steven Lapointe (1987, 1988) has argued forcefully for an autolexical system in which association lines never cross, making the morphosyntactic interface obey a fundamental constraint of the allied discipline of autosegmental phonology (Goldsmith 1976). If Lapointe is right, then we must abandon the attempt to account for ordering in such terms. As it turns out, the facts concerning the direction of incorporation seem to yield easily to a straightforward and intuitively sensible principle that makes no appeal to line crossing.

Returning to the problem at hand, note that if the language in question is one in which L and H both occur on the same periphery of the phrases they form (as is quite often the case with incorporative phenomena), it will be impossible to tell the difference between these two cases, at least on the basis of word order alone. But linear order ought to reflect the direction of incorporation if L and H are not adjacent, giving rise to either (49), (50), or their mirror images.

Now any of these morphosyntactic associations have a certain air of implausibility about them. The configuration in (49) (or its mirror image) corresponds to a raising analysis under a transformational treatment such as Baker's (1988); the element H is found in a higher position in the tree, namely in association with the higher element L. Such a structure would be licit on Baker's account in that the trace of movement of H would be properly governed by the element that has been moved to a c-commanding position. But such a process could also easily lead to the production of new superficial structures, not found outside of incorporation, since there is no guarantee that the remnant of HP, namely x-y, is an otherwise occurring phrase type in the language. Even if x-y is a possible phrase in the language, it must be a possible complement of L, in other words, a null-headed HP, for the association to be structure preserving (assuming that the complex morphological entity L+H counts as an L).

The sort of association we find in (50) is also less than plausible looking, for the lowering of L to H would (in transformational terms) leave an ungoverned trace. On the other hand, this direction of association would seem to have an easier time producing a surface order of morphemes that resembles what otherwise occurs in the language. For it to do so, the sequence x-L+H-y would have to resemble an HP, and thus all that would have to be true is that L+H is an H in the language.

But despite their initial implausibility, there seem to be good examples of the type of association shown in both (49) and (50) in the data surveyed in chapter 5. Therefore, structure preservation is not a necessary characteristic

of incorporation. Indeed, if it were, some of the most important evidence for the syntactic reality of the process would not have existed. Furthermore, government does not seem to play a prominent role in licensing incorporation of the kind I am dealing with, though it might in the case of analogous mismatches between semantic and syntactic representations, which form the bulk of the cases that Baker treated in his work.

All examples of noun incorporation that I know of are either compatible with (49), or must be analyzed in that way. In West Greenlandic, for example, the normal order of a stranded modifier with respect to an incorporating verb matches the normal order of an instrumental case object with respect to an intransitive verb:[8]

(51) Angisuu-nik qimme-qar-poq
big-INST/p dog-have-INDIC/3s
'He has big dogs.'

(52) Meeqqa-nik asannip-poq
child-INST/p love-INDIC/3s
'He loves children.'

Since the order of elements in a West Greenlandic NP is N+modifier, this clearly shows that the incorporated N-stem is attracted to the verb (as in the mirror image of (49)), rather than the verbal affix being attracted to the N-stem.

As a case of the opposite kind, consider the incorporating determiners of Icelandic, as analyzed in chapter 5, section 1.1. If, as I suggest there, the determiner is "lowered" from the position that ordinary demonstratives occupy (example (54)) to the head noun, then we have the situation diagrammed in (50).

(53) rauði hestur-inn sem týndast
red horse-DET which got.lost

(54) sá rauði hestur sem týndast
DEM red horse which got.lost

The survey in chapter 5 contains a near minimal pair of phenomena with respect to the direction of incorporation which is instructive to consider in more detail. Both Hungarian (chapter 5, section 4.2) and Crow (chapter 5, section 4.3) contain incorporating postpositions, but those of Hungarian descend to the head of their NP complements (if the derivational metaphor may be excused), while those of Crow rise to combine with the head of the phrase that governs them.

(55) János bement a nagy ház-ba.
J. entered the big house-into
'Janos went into the big house.'

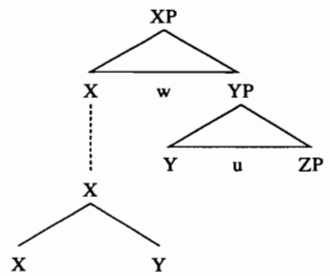
(56) Húuleesh Jerry-sh Chichúche ku-ss-dée-k.
yesterday J.-DEF Hardin it-GOAL-go-DECL
'Jerry went to Hardin yesterday.'

One might think that the difference has something to do with what item is responsible for the incorporation, but this notion is falsified by the pair of cases at hand, since in both languages the incorporating power is a lexical feature of the adposition, some of which do not participate in incorporation at all. We may also compare the Eskimo situation, where the lexeme that must be specifically marked as triggering the incorporation is clearly the suffixal verb, with Gta? (and other south Munda languages such as Sora) where most nouns have special incorporating forms, but those that do not cannot incorporate (Ramamurti 1931). In either case, the noun-verb combination occurs in the position where the simple verb would be found.

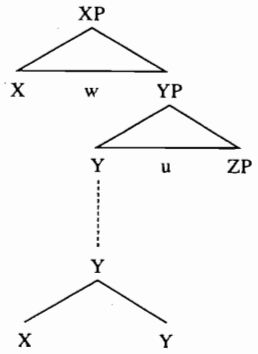
The actual difference resides in the relationship between the morphological and syntactic rules that the autolexical element participates in. In both Crow and Hungarian we are dealing with an element that combines with NPs in the syntax to form adpositional phrases. But the morphological combinatorics of the postpositions differ greatly between the two languages. In Hungarian they participate in a morphological rule that combines a noun and a postposition to form a noun,[9] whereas in Crow they combine with verb stems to form verb stems. What we can now observe is that the complex noun of Hungarian occurs where the noun would, and the complex verb of Crow occurs where the verb would. In general, then, the complex morphological item occurs where its head occurs in syntax. The reason that all noun incorporation involves raising is simply that noun incorporation, among other things, is the morphological combination of a nominal and verbal form to form a verb, which by the principle just adumbrated will occur in the position of the syntactic verb.

The situation can be understood by examining the following diagrams, which resemble those argued for by Lapointe (1987; 1988) in that nonterminal nodes of the morphological tree are associated with the syntactic tree.

(57)



(58)



The diagram in (57) corresponds to the Crow situation, where X = V, and Y = P; The diagram in (58) analyzes the Hungarian facts, where X = P, and Y = N.

Though it is quite possible, and certainly desirable, for the principle governing the direction of incorporation to follow from the architecture of the system, I do not see that it does at present, and so I will state it as an observational law to be explained by further research:

(59) Direction of Incorporation
The syntactic position of a complex morphological expression is the same as the syntactic position of its morphological head.

Thus in (57), the morphological expression [X,Y] is headed by X, and so it occurs in the syntactic position of X, in the syntactic tree, namely before w-u-ZP. In (58), the morphological phrase [X,Y] is headed by Y, and so the complex occurs in the syntax where Y does, namely between w and u-ZP.

5

A Survey of Morphosyntactic Mismatches

The interface principles discussed so far greatly limit what kinds of discrepancies are allowed to exist between syntactic and morphological bracketings, and therefore the possible types of incorporative and clitic-like phenomena that can exist, but they still allow quite a range of them. As described above, incorporation involves morphology to a stem, but it is neutral as to whether this will be inflectional or derivational morphology, compounding, affixation, or modification. Clitics generally manifest themselves as affixes, but in at least one case, that of Tongan discussed below, an argument can be made for the existence of a clitic that shows up as an effect on stress.

Saussure (1959) and Sapir ([1921] 1949) emphasized that the same notional category can be expressed in terms of a variety of formal means. Such categories as tense and aspect are manifested in some languages (e.g., various creole languages) in the form of a verb that takes a verb-phrase complement, but in others as an inflection on main verbs (Latin), and in others (e.g., Eskimo), as derivational morphology. We might take Chinese to be a language in which aspect is realized as a clitic. If there is evidence that tense or aspect is to be represented in the syntax of a language as a V with a VP complement, and the formal expression of this category is in terms of some morphological operation, then we would have a case of incorporation, regardless of the specific morphological means that are employed. The Kirundi example discussed in section 3.2 below is an inflectional example of incorporated tense.

Another dimension of variation for morphosyntactic mismatches concerns the syntactic substance of the incorporating element. For each syntactic rule in natural language that introduces a lexical category and a complement, the IP and CP allow for there to be incorporative and clitic instantiations of the category of the head. We will be able to recognize such morphemes by the fact that they leave behind what is left of an ordinary complement phrase after its lexical head has been removed, have the semantics that we would ordinarily expect of syntactic formatives, and are mutually exclusive with the actual occurrence, in the form of an independent formative, of a duplicate of the syntactic category they instantiate. If such a morpheme is expressed as morphology to a stem, it should obey the IP and be found in morphological association with the head of its complement phrase, and if it is realized as