1 *The problem of the lexical categories*

1.1 A theoretical lacuna

It is ironic that the first thing one learns can be the last thing one understands. The division of words into distinct categories or "parts of speech" is one of the oldest linguistic discoveries, with a continuous tradition going back at least to the *Téchnē grammatikē* of Dionysius Thrax (*c.* 100 BC) (Robins 1989: 39). Dionysius recognized that some words (*ónoma*, alias nouns) inflected for case, whereas others (*rhēma*, alias verbs) inflected for tense and person. This morphological distinction was correlated with the fact that the nouns signified "concrete or abstract entities" and the verbs signified "an activity or process performed or undergone." The historical precedence of this linguistic insight is often recapitulated in contemporary education: often when students enter their first linguistics class, one of the few things they know about grammar is that some words are nouns, others are verbs, and others are adjectives. Linguistics classes teach them many fascinating things that go far beyond these basic category distinctions. But when those classes are all over, students often know little more about what it means to be a noun, verb, or adjective than they did at first, or indeed than Dionysius did. At least that was true of my education, and of the way that I learned to educate others.

For many years, most of what the Principles and Parameters (P&P) tradition of Generative Syntax has had to say about the lexical categories is that they are distinguished by having different values for the two binary distinctive features +/−N and +/−V in the following way (Chomsky 1970).[1]

[1] Chomsky (1970) did not, in fact, include adpositions in his feature system at first. The gap was filled in by Jackendoff (1977), in light of his influential view (which I argue against in the appendix) that prepositions constitute a fourth lexical category.

More recent sources that use essentially this feature system include Stowell (1981), Fukui and Speas (1986), and Abney (1987). Fukui's innovation was to extend Chomsky's feature system from the lexical categories to the functional ones. Abney's goal is similar, except that he suppresses the feature +/−verbal, making it difficult to account for the difference between nouns and adjectives or between verbs and prepositions in languages where these are distinct. See section 1.3 below for Jackendoff's (1977) alternative system and others related to it.

(1) a +N, −V = noun
b −N, +V = verb
c +N, +V = adjective
d −N, −V = adposition (preposition and postposition)

But this theory is widely recognized to have almost no content in practice. The feature system is not well integrated into the framework as a whole, in that there are few or no principles that refer to these features or their values.[2] Indeed, it would go against the grain of the Minimalist trend in linguistic theory (Chomsky 1995) to introduce extrinsic conditions that depend on these features. All the features do is flag that there are (at least in English) four distinct lexical categories. Since 4 is 2^2, two independent binary features are enough to distinguish the four categories, but there is no compelling support for the particular way that they are cross-classified in (1). By parallelism with the use of distinctive features in generative phonology, one would expect the features to define natural classes of words that have similar distributions and linguistic behaviors. But of the six possible pairs of lexical categories, only two pairs do not constitute a natural class according to (1): {Noun, Verb} and {Adjective, Adposition}. Yet these pairs do, in fact, have syntactic similarities that might be construed as showing that they constitute a natural class. For example, both APs and PPs can be appended to a transitive clause to express the goal or result of the action, but NPs and VPs cannot:

(2) a John pounded the metal flat. (AP)
b John threw the ball into the barrel. (PP)
c *John pounded the metal a sword. (NP)
d *John polished the table shine. (VP)

In the same way, only adjectives and adpositions can modify nouns (*the man in the garden* and *the man responsible*) and only they can be preceded by measure phrases (*It is three yards long* and *He went three yards into the water*). All told, there is probably as much evidence that adjective and adposition form a natural class, as there is that noun and adposition do. The feature system in (1) is thus more or less arbitrary. Stuurman (1985: ch. 4) and Déchaine (1993: sec. 2.2) show that syntactic evidence can be found in favor of any logically possible claim that two particular lexical categories constitute a natural class.

[2] At one point, case theory was an exception to this. In the early 1980s, it was common to say that the −N categories could assign case, whereas the +N categories received case (Stowell 1981). That is not the current view however; rather, Ns and As license genitive case, which happens to be spelled out as *of* in English (Chomsky 1986b).

Stuurman goes on to conclude that the idea of decomposing syntactic categories into complexes of features is bankrupt.

Related to this is the fact that generative linguistics has been preoccupied with explaining the similarities that hold across the lexical categories, and has had little to say about their differences. X-bar theory, a central component of the theory (at least until recently), clearly had this goal. Chomsky (1970) introduced X-bar theory precisely to account for the observation that nouns take the same range of complements and form the same types of phrases as verbs do. From then till now, the job of X-bar theory has been to account for the sameness of the various categories, but not for their differences. This is also true of the extensive research on functional categories over the last two decades. A common theme in this work, as initiated by Abney (1987), has been to account for the structural parallels between clauses and nominals – for example, the similarity of complementizers and case markers, of tense and determiners, and of aspect and number. Much important insight has come from these two research thrusts. But when one is steeped in these lines of work, it is easy to forget that the various lexical categories also differ from one another, and the theory has almost nothing to say about these differences. In most contexts, one cannot swap a verb for a noun or an adjective and preserve grammaticality, and X-bar theory and the theory of functional categories by themselves can never tell us why. The time thus seems ripe to attend to the differences among the lexical categories for a while.

1.2 Unanswerable typological questions concerning categories

A serious consequence of the underdevelopment of this aspect of syntactic theory is that it leaves us ill equipped to do typology. The literature contains many claims that one language has a different stock of lexical categories from another. In many cases, these claims have caused controversy within the descriptive traditions of the language families in question. Since there is no substantive generative theory of lexical categories, we have no way to assess these claims or resolve these controversies. Nor do we make interesting predictions about what the consequences of having a different set of basic categories would be for the grammar of a language as a whole. Therefore, we cannot tell whether or not there is any significant parameterization in this aspect of language.

To illustrate this crucial issue in more detail, let us consider the actual and potential controversies that arise when trying to individuate the lexical categories

in the Mohawk language. For example, does Mohawk have adjectives? The traditional Iroquoianist answer is a unanimous no; Mohawk has only stative verbs, some of which are naturally translated as adjectives in English. The primary evidence for this is that putative adjectives take the same agreement prefixes and some of the same tense/aspect suffixes as uncontroversial intransitive verbs:

(3)

a	*ka*-hútsi	compare:	t-a'-*ka*-yá't-ʌ'-ne'
	NsS-black		CIS-FACT-NsS-body-fall-PUNC
	'it is black'		'it (e.g. a cat) fell'
b	*ra*-hútsi	compare:	t-a-*ha*-yá't-ʌ'-ne'
	MsS-black		CIS-FACT-MsS-body-fall-PUNC
	'he is black'		'he fell' (*ra* ⟶ *ha* when not word-initial)
c	ka-rák-ʌ	compare:	t-yo-ya't-ʌ'-ʌ
	NsS-white-STAT		CIS-NsO-body-fall-STAT
	'it is white'		'it has fallen'
d	ka-hutsí-(Ø)-*hne'*	compare:	t-yo-ya't-ʌ'-ʌ-*hne'*
	NsS-black-(STAT)-PAST		CIS-NsO-body-fall-STAT-PAST
	'it was black'		'it had fallen'

The tradition of considering inflectional evidence of this kind as central to judgments about category membership goes all the way back to Dionysius's *Téchnē*, and has been influential throughout the history of linguistics in the West (Robins 1989).

Putative adjectives are also like intransitive verbs in another way: they both allow noun incorporation, a process by which the head noun of an argument of the verb appears attached to the verb root to form a kind of compound (Mithun 1984; Baker 1996b):

(4)

a	Ka-wis-a-hútsi	thíkʌ.	
	NsS-glass-Ø-black	that	
	'That glass is black'		
b	T-a'-ka-wís-ʌ'-ne'		thíkʌ.
	CIS-FACT-NsS-glass-fall-PUNC		that
	'That glass fell.'		

This seems to corroborate the claim that words like *hutsi* 'black' are verbs in Mohawk.

Nevertheless, if "adjectives" are verbs in Mohawk, then they must be identified as a subclass that has some special properties. Adjectival roots cannot, for example, appear in the punctual or habitual aspects, but only in the stative aspect:

(5)

a	*wa'-ká-rak-e'	compare:	t-a'-ka-yá't-ʌ'-ne'
	FACT-NsS-white-PUNC		CIS-FACT-NsS-body-fall-PUNC
	'it whited'		'it fell'
b	*ká-rak-s	compare:	t-ka-yá't-ʌ'-s
	NsS-white-HAB		CIS-NsS-body-fall-HAB
	'it whites'		'it falls'

This restricted paradigm does not follow simply from the semantic stativity of words like *rakʌ* '(be) white' because transitive stative predicates like *nuhwe'* 'like' can easily appear in all three aspects. Even when both "adjectives" and verbs appear in the stative aspect, there are differences. Eventive verbs in stative aspect always show what looks like object agreement with their sole argument (see Ormston [1993] for an analysis consist with Baker [1996b]). In contrast, adjectival verbs in stative aspect often show subject agreement with their sole argument:

(6)

a	*ka*-rak-ʌ	(**yo-rak-v* NsO-white-STAT)
	NsS-white-STAT	
	'it is white'	
b	te-*yo*-hri'-u	
	DUP-NsS-shatter-STAT	
	'it has/is shattered'	

A more subtle difference between "adjectives" and (other) intransitive verbs is that only "adjectives" permit a kind of possessor raising. When a noun is incorporated into a word like *rak* 'white', that word can bear an animate object agreement marker that is understood as expressing the possessor of the incorporated noun (see (7a)). Comparable eventive verbs allow simple noun incorporation, but they do not allow a similar animate object agreement marker, as shown in (7b) (Baker 1996b: ch. 8.4).

(7)

a	*Ro*-nuhs-a-rák-ʌ	ne Shawátis.
	MsO-house-Ø-white-STAT	NE John
	'John's house is white.'	
b	*Sak wa'-t-*ho*-wis-á-hri'-ne'.	
	Jim FACT-DUP-*MsO*-glass-Ø-break-PUNC	
	'Jim's glass broke.'	

The unanswerable question, then, is this: do these differences justify positing a separate category of adjectives in Mohawk after all? Or do we continue to say that Mohawk has only verbs, but concede that there are two subtypes of verbs, intransitive stative verbs and other verbs? Generative syntactic theory gives no leverage on these questions, precisely because there are no

principles of the theory that mention verbs but not adjectives or vice versa. Therefore, the choice we make has no repercussions and makes no predictions. In essence, the decision comes down to a matter of taste or terminology (Schachter 1985).

Similar issues arise concerning whether Mohawk has a distinct category of adposition. Some Iroquoianists have argued that it does; others say that the putative adpositions are really stative verbs or derivational noun suffixes. The best candidates are four bound morphemes that have locative meanings: *-'ke/-hne* 'at,' *-ku* 'in,' *-oku* 'under,' and *-akta* 'near.' (8) shows the results of combining these elements with four representative nouns of Mohawk:

(8)	'bed'	'box'	'table'	'car'
Ø	ka-nákt-a'	o-'neróhkw-a'	atekhwára	ká-'sere-'
'at'	ka-nakt-á-'ke	o-'nerohkw-á-'ke	atekhwará-hne*	ka-'sere-ht-á-'ke
'in'	ka-nákt-a-ku	o-'neróhkw-a-ku	atekhwara-tsher-á-ku	ka-'seré-ht-a-ku
'under'	ka-nakt-óku	o-'nerohkw-óku	atekhwara-tsher-óku	ka-'sere-ht-óku
'near'	ka-nakt-ákta	o-'nerohkw-ákta	atekhwara-tsher-ákta	ka-'sere-ht-ákta

The attraction of saying that these locative morphemes are stative verbs comes from the combinations in (8) having some of the same morphological peculiarities as noun incorporation into verbs. Nouns that are historically derived from verbs must be augmented by a "nominalizer" morpheme when they are incorporated into a verb. Thus, *-tsher* is added to *atekhwara* 'table' in (9a), *-ht* is added to *'sere* 'car' in (9b), and nothing is added (9c).

(9) a ʌ-k-atekhwara-*tsher*-úni-'
FUT-1sS-table-NOML-make-PUNC
'I will make a table.'
b wa'-ke-'sere-*ht*-óhare-'
FACT-1sS-car-NOML-wash-PUNC
'I washed the car.'
c wa'-ke-'nerohkw-a-hninu-'
FACT-1sS-box-Ø-buy-PUNC
'I bought a box.'

The examples in (8) show that the same lexically idiosyncratic augments appear when combining the locative elements with the nouns. Furthermore, when the incorporated noun (plus augment, if any) ends in a consonant and the verb root begins in a consonant, a special joiner vowel /a/ is inserted between the two (e.g. (9c)); (8) shows that this rule also applies to locative elements. These idiosyncrasies do not take place when other, clearly derivational suffixes are added to nouns.

Locative elements differ from stative verbs and derivational suffixes in other respects however. For example, the inflectional prefix on the noun (usually *ka-* or *o-*) is lost when it is incorporated into a verb (see (9)), but not when it is combined with a locative element, as shown in (8). (10) shows that even a possessive prefix can show up on a noun-plus-locative form.

(10) Shawátis *rao-*'seré-ht-a-ku
John MsP-car-NOML-Ø-in
'in John's car'

This prefix *rao-* is phonologically distinct from any prefix that appears on true verbs.

Nouns that combine with locative elements also acquire new distributional possibilities. Nouns in Mohawk must normally be linked with a pronominal/agreement prefix on some verbal element in the clause. Thus (11b) is ungrammatical, in contrast with (11c). However, (11a) shows that this requirement does not hold of a noun plus a locative element.

(11) a Thík^ o-nut-á-'ke yó-hskats ne okwire'-shú'a.
that NsO-hill-Ø-at NsO-be.pretty NE tree-PLUR
'On that hill, the trees are pretty.'
b *Thík^ onúta', yó-hskats ne okwire'-shú'a.
That hill NsO-be.pretty NE tree-PLUR
'As for that hill, the trees are pretty.'
c Thík^ onúta' yó-hskats.
That hill NsO-be.pretty
'That hill is pretty.'

This difference in syntactic distribution is unexpected if the locative elements are merely derivational morphemes that form nouns from nouns.

Overall, then, nouns with the locative endings are not exactly like stative verbs, or simple nouns, or any other class of expressions in Mohawk. Again, the question arises whether these facts are enough to justify positing a distinct category of adposition for Mohawk. And again syntactic theory gives us little help in answering the question.

Finally, we can ask whether there is a category distinction between nouns and verbs in Mohawk. Most of the Iroquoianist literature says that there is, but there are potential grounds for doubting this, and Sasse (1988) argues against a distinction. Like verbs (and adjectives, if those are distinct), nouns can be used as the main predicate of a clause, as shown in (12).

(12) a Ka-núhs-a' thíkʌ o-'nerohkw-a'-kʌha.
NsS-house-NSF that NsO-box-NSF-former
'That old box is a house.' (a child's play house, or a street person's shelter)
b Ka-rák-ʌ thíkʌ o-'neróhkw-a'.
NsS-white-STAT that NsO-box-NSF
'That box is white.'

There are also inflectional similarities between nouns and other categories. Potential evidence for the standard view that nouns are a distinct category is the fact that no tense/aspect marker can be attached to nouns, not even the stative:

(13) a *wa'-ká-nuhs-*e'* punctual 'it housed'
b *ka-núhs-*ha'* habitual 'it always houses'
c *(y)o-núhs-*u* stative 'it is a house'
d *o-khwarí-(Ø)-*hne'* past 'it was a bear.'

Furthermore, the pronominal/agreement prefixes that attach to nouns are slightly different from the ones that attach to (adjectives and) verbs, as shown in (14).

(14) a *ka*-núhs-a' compare: *ka*-rák-ʌ
NsS-house-NSF NsS-white-STAT
'(it is a) house' 'it is white'
b *ó*-wis-e' compare: *yo*-hnír-u
NsO-glass-NSF NsO-hard-STAT
'(it is a) glass' 'it is hard'
c *rao*-núhs-a' compare: *ro*-nuhs-a-rák-ʌ
MsP-house-NSF MsO-house-Ø-white-STAT
'(it is) his house' 'his house is white'

The prefixes that appear on nouns are not *very* different from the prefixes that attach on verbs, however. The nominal prefixes are cognates of the verbal ones: they can be analyzed as having the same underlying form, the noun prefixes being derived from the verb prefixes by morphophonological rules that delete initial glides (as in (14b)) and that create diphthongs out of some simple vowels (as in (14c)).

There are also more subtle parallelisms between the prefixes on nouns and the prefixes on verbs. An unaccusative verb (a verb that takes only an internal, theme argument) takes a prefix that expresses the person–number–gender properties of its subject; typically the form is a "subject" agreement prefix ((15b)), although some verbs are lexically marked as taking "object"agreement. In a similar

way, a noun takes a prefix that expresses the person–number–gender properties of its referent, typically with a "subject" agreement (15b), but sometimes with an "object" agreement instead, depending on the particular noun root. A goal or affected object argument can also be added to almost any verb; this is always expressed as an "object" prefix (15a). In the same way, most nouns can take a possessor, and this too is expressed with the relevant "object" prefix ((15a)).

(15)

a	*akó*-wis-e'	compare:	t-a'-*akó*-hs-ʌ'-s-e'.
	FsP-glass-NSF		CIS-FACT-FsO-Ø-fall-BEN-PUNC
	'her glass'		'it fell on her; she dropped it'
b	*ra*-ksá'-a	compare:	t-a-*ha*-yá't-ʌ'-ne'.
	MsS-child-NSF		CIS-FACT-MsS-body-fall-PUNC
	'boy'		'he fell'
c	**shako*-ksá'-a	compare:	*t-a-*shako*-yá't-ʌ'-s-e'.
	MsS/FsO-child-NSF		CIS-FACT-MsS/FsO-body-fall-BEN-PUNC
	'her boy'		'he fell on her; she dropped him'

Given these generalizations, one would think that nouns and unaccusative verbs should also be able to bear explicitly transitive agreement prefixes, with the subject factor of the prefix expressing the referent of the noun or the theme of the verb, and the object factor expressing the possessor of the noun or the affected object of the verb. But this is not so: transitive prefixes are impossible on both nouns and unaccusative verbs, as shown in (15c). There is a rather striking overall parallel between the inflection of nouns and the inflection of unaccusative verbs in Mohawk, with the referent of the noun being analogous to the theme of the verb, and the possessor of the noun being analogous to the goal/affected object of the verb. This parallelism led me to propose that nouns in Mohawk form the same kinds of syntactic structures as unaccusative verbs (Baker 1996b: ch. 6). One could then take this one step further, and claim that nouns actually *are* unaccusative verbs. In this view (roughly that of Sasse 1988) there would be no distinction between the two categories in Mohawk syntax, but only at a superficial level of morphophonology.

This radical conclusion would be premature, however, since there are also differences between nouns and unaccusative verbs. As mentioned above, an important property of unaccusative verbs (including "adjectives") in Mohawk is that they allow their theme argument to be incorporated. In contrast, the referent argument of a noun can never be incorporated into the noun, as shown in (16).

(16) a *Ka-'nerohkw-a-núhs-a' (thíkʌ). (compare (12a))
NsS-box-Ø-house-NSF that
'That box is a house.'
b Ka-'nerohkw-a-rák-ʌ (thíkʌ)
NsS-box-Ø-white-STAT that
'That box is white.'

In Baker (1996b), I had no explanation for this difference between nouns and unaccusative verbs. Yet it does not seem to be an accidental difference; there are quite a few languages that allow noun incorporation into verbs (Mithun 1984), but no known languages that allow noun incorporation into nouns. Such a difference should ideally follow from a proper understanding of what it is to be a noun as opposed to a verb. It does not, however, follow from a theory that merely says that nouns are +N, −V and verbs are +V, −N. Nor does this theory give any firm basis for deciding whether nouns are a distinct class of heads from verbs in Mohawk or not.

I have lingered over the lexical category system of Mohawk because I believe that the issues it raises are entirely typical of those presented by other languages. Many languages are said not to distinguish certain adjectives from stative intransitive verbs, including other Native American languages (Choctaw, Slave, Mojave, Hopi, etc.) and some African languages (such as Edo and Yoruba) (Dixon 1982; Schachter 1985). Other languages are said not to distinguish adjectives from nouns, including Quechua, Nahuatl, Greenlandic Eskimo, and various Australian languages (Dixon 1982; Schachter 1985). But even in these languages writers of dictionaries and grammars are often led to distinguish "adjectival nouns" from other nouns or "adjectival verbs" from other verbs because of some subtle phenomena. There is also a great deal of uncertainty across languages over what counts as an adposition as opposed to a noun suffix or dependent verb form. Even the existence of a noun–verb contrast is controversial in a few language families, most notoriously the Wakashan and Salish families of the Pacific Northwest and some Austronesian languages (Schachter 1985). These controversies typically hinge on disagreements about what importance to assign to different kinds of evidence, such as inflectional paradigms, derivational possibilities, syntactic distribution, and semantically oriented factors. The general problem of distinguishing categories from subcategories in a principled way has been observed by typologists like Schachter (1985: 5–6) and Croft (1991), among others. Since generative theory offers no decisive way to resolve these questions, we are left not knowing whether there is significant crosslinguistic variation in this respect or not, and if so what its repercussions are. This is a fault that I wish to remedy.