Boundaries of Morphology and Syntax

EDITED BY
Lunella Mereu

Offprint
THE EFFECT OF NOUN INCORPORATION
ON ARGUMENT STRUCTURE

MARIANNE MITHUN & GREVILLE G. CORBETT
University of California, Santa Barbara - University of Surrey

0. Introduction
Important discussions of the boundaries between morphology and syntax have focused on cross-linguistic differences in the expression of similar ideas. What is expressed syntactically in one language is sometimes expressed morphologically in another. Morphological and syntactic expressions of possession are compared in Corbett (1987), for example, and passives in Börjars et al. (1997). Particularly frequent in such discussions has been noun incorporation, a kind of noun-verb compounding that yields new verb stems. An example can be seen in the sentence in (1) from Mohawk, an Iroquoian language of northeastern North America. The full Mohawk sentence consists of a verb based on the verb root -ienho “plant” compounded with the noun root -tsi ’tsi- “flower”.

(1) Mohawk noun incorporation: (Karihwénhawe Lazore, speaker)

\[
\begin{align*}
\text{Waklsi ‘tsiaiéthon} \\
\text{wak-tsi ’tsi-a-ienho-on} \\
\text{ISG.PAT.-flower-EP-plant-ST}
\end{align*}
\]

“I’ve planted flowers.”

* Mohawk examples are presented here in the community orthography adopted by the six Mohawk communities in Quebec, Ontario, and New York State (Kahnawá:ke, Kanehsatá:ke, Akwesásne, Thaientaná:ken, Ohswé:ken, and Wáhta). We are extremely grateful to the members of these communities who have generously shared their expertise on the language over the years, and especially to Kanerahtenhá:wi Hilda Nicolas of Kanehsatá:ke, Kaia’titákkhé’ Annette Jacobs of Kahnawá:ke, and Rokwáho Dan Thompson of Akwesásne, for their insightful discussion of issues addressed here. We are also grateful to Brian Butterworth for references to work on speakers’ knowledge of lexical frequency, and to Andrew Pawley for discussion about lexicalization. This research has been supported in part by the Ontario Training and Adjustment Board and the Academic Senate, University of California, Santa Barbara (MM), and by the Economic and Social Research Council, UK (grant R000236063) (GGC); this support is gratefully acknowledged.
Translations of single words like these into full English sentences have sometimes fostered an assumption that their differences are negligible, that whether an idea is expressed morphologically in a single word or syntactically in several is an uninteresting accident of form. Several lines of research have aimed at describing morphological structures in syntactic terms (Lees 1960, Postal 1962, Sadock 1980, 1985, Baker 1988, 1995, 1996, among others). This approach has not been uncontroversial. Over the past decade, most discussion of noun incorporation has focused on the issue of whether it should indeed be analyzed as a syntactic process (Di Sciullo & Williams 1987, Rosen 1989, Baker 1995, 1996, and others).

Describing incorporation in syntactic terms could have certain advantages. Already highly articulated models of syntax could be extended to a new domain. If established syntactic principles accounted for patterns of incorporation as well, the principles could be seen to be more general than originally thought and so, in a sense, predictive. The exercise of testing them on incorporation might, furthermore, prompt us to notice facets of the construction that could otherwise go unobserved. On the other hand, if noun incorporation shares important features with other morphological processes, analyzing it in terms of syntactic models alone could prevent us from capturing generalizations in that domain.

Probably the most ambitious treatments of incorporation as a syntactic process have been those of Baker (1988, 1995, 1996, 1997, and elsewhere), who takes Mohawk as his prime example. In his earlier work, Baker defended the view that incorporation in Mohawk was purely syntactic. More recently, he has retreated from this position somewhat to propose that some incorporating constructions should be considered syntactic but others lexical. Unfortunately, as he points out (1995), various criteria for distinguishing syntactic and lexical incorporation do not produce the same classifications of specific items. Furthermore, individual diagnostics do not

| Abbreviations in the text and in glosses are the following: |
|-------------------|-------------------|
| AGT | grammatical agent |
| CAUS | causative |
| CISLOC | cislocative |
| CONTR | contrastive |
| DIM | diminutive |
| DC | duplicative |
| DJ | dual number |
| EP | epenthetic |
| FACT | factual |
| FUT | future |
| IMPRF | imperfective |
| INDEF | (fem)/inddefinite gender |
| M | masculine |
| NOM | nominalizer |
| OPT | optative |
| PAT | grammatical patient |
| PL | plural |
| PRF | perfective |
| PRT | partitive |
| PURP | purposive |
| RFL | reflexive |
| REV | revesive |
| SG | singular |
| ST | stative |
| TRANSLOC | translocative |
always yield clear results for particular forms. The apparently conflicting results are more easily explained once we recognize the significance of the fact that incorporation creates lexical items.

1. **Background**

Several types of evidence indicate that the products of noun incorporation in Mohawk are different in kind from the products of prototypical syntactic processes. The differences are both formal and functional.

1.1 *The lexical integrity of Mohawk words*

It might be wondered whether constructions such as *waktsi’tsiaiéthon* “I’ve planted flowers” in (1) are truly single words, since they seem to convey the content of a full English sentence. Evidence of different types confirms that they are. Most important are speaker judgments, which are unequivocal on this point. Unlike words, individual morphemes are unintelligible to speakers in isolation. (The unconscious knowledge of morphological structure that speakers exploit to create new words is another matter.)

Other phonological and morphological patterns confirm the solidity of the word as an integral unit. Mohawk words have just one primary stress, usually on the penultimate syllable. Numerous phonological adjustments occur at morpheme boundaries within words but not between words. The order of elements contrasts sharply: morpheme order is rigid and invariant, while word order is almost fully fluid in Mohawk, governed by pragmatic considerations. Both nouns and verbs contain certain obligatory prefixes and suffixes which also serve to signal the beginnings and ends of words. Additional traditional criteria for wordhood are unequivocal as well. Speakers do not pause in the middle of words, searching for the next morpheme, then continue. If for any reason they stop mid-word, they go back and begin afresh, selecting the word as a whole rather than piece by piece.

1.2 *The functions of incorporation*

Languages with incorporating constructions generally contain analytic syntactic counterparts as well. In the Mohawk sentence in (1), the flowers were represented by the incorporated noun root *-tsi’tsi-*, but they could also be identified by an independent word instead, the noun *otsi:tsia’*, as in (2).

(2) Mohawk syntactic counterpart (Karihwénhawe Lazore, speaker)

*Otsi:tsia’*

*o-tsi’tsi-a’*

*NEUTER-FLower-Noun.SUFFIX*

*Wakiéthon*

*wakiéthon*

*1SG.PAT-Plant-ST*

“*I’ve planted flowers.*”
Most Mohawk noun roots cannot constitute words alone, but must be preceded by a prefix like the neuter o- in "flower" and followed by a noun suffix like -a'. This does not mean of course that every noun and verb root in the language occurs in both constructions, only that the parallel structures coexist as structures. Morphological and syntactic counterparts like (1) and (2) would not coexist so systematically if they were identical in function. Their semantic and pragmatic differences are well documented (Mithun 1984, 1986a, Evans [1991:264-292] among others). Essentially, incorporation allows speakers to package components of a concept together in a single word. Such packaging is done for two kinds of purposes, the creation of new labels and the regulation of information in discourse.

Incorporation, like all compounding, serves a basic word-formation function, as a device for creating terms for nameworthy concepts, such as te:wakhwishenhe:i'on (te-wak-lwish-enhei-on DC-1.PAT-strength-die-ST) "my strength has died" = "I'm tired". Often the incorporated noun serves to narrow the scope of the verb semantically, creating a verb stem for a more specific kind of activity or state than the verb root alone, as in ronwationhwa:nihen (ronwa-rnonhs-a-nih-en INDEF.AGT/ M.PL.PAT-house-EP-lend-ST) "one house-lent them" = "they were renting". Renting is a special kind of lending, house-lending for payment. Morphological verbs are used pervasively in Mohawk as syntactic nominals. Objects are often described in terms of their function or appearance, best expressed with verbs, as in iehwista'ekstha' (i-ehwista-ek-st-ha' INDEF.AGT-metal-EP-strike-INSTRUMENTAL-IMPRF) "one metal-strikes with it" = "bell".

Incorporation has a second kind of function in Mohawk. The alternation between incorporated and independent nouns is used pervasively to regulate the flow of information through discourse. Separate nouns tend to be used to focus individual attention on a newsworthy piece of information, such as a significant new participant or a contrast. Information that is already an established part of the scene, predictable, or incidental, may be carried along by an incorporated noun. During one conversation, a speaker was describing a community program in which seeds were distributed. Her grandson asked permission to plant on her land:

(3) General gardening

General gardening

Enwá:ton
en-w-at-on
FUT-NEUTER-be.possible-ST

ken
ken
Q

wahen:ron
wa-ha-ihron-
FACT-M.AGT-say-PRF

entié ntho’?
en-k-ié ntho’-
FUT-1SG.AGT-plant-PRF

"May I plant here?, he said."
She continued (in Mohawk): “I said of course you can, I guess. I was very happy that he would learn this right. What was so hard about that? All I did was talk to him, and he did all the work”. In the next line, she incorporated the noun -heht- “garden”.

(4) Incorporated noun: -heht- “garden”

Ahsatkáhtho’ tsi nihoheht:iio nón:wá.
a-hs-atkahtho’ tsi ni-ho-heht-iio n-onhwa
OPT-2SG.AGT-see-PRF that PRT-M.PAT-garden-nice.ST the-now
“You should see how nice his garden is now.”

Though the noun “garden” had not been used before, the idea of a garden had become an established part of the discussion and did not require special attention. It was accordingly backgrounded by incorporation. The speaker continued again, talking about other people she had seen planting.

(5) Independent noun: ö:nenhste’ “corn”

Ö:nenhste’ i:kehre’ rotiíenthon.
o-nenhste’ k-ehr-e’ roti-ientho-on
NEUTER-corn-NOUN.SUFFIX ISG.ACT-IHIINK-IMPRF M.PL.AGT-DIANST-ST
“I think they’ve planted corn.”

The corn was significant new information, worthy of a separate word.

Incorporation for discourse purposes can also be seen with presentative verbs, verbs with little semantic content of their own, such as “exist” or “have”. The entity and its existence together form a single unit of newsworthy information, so they are often packaged together in a single word, such as tkahéhtaïen (t-ka-heht-a-i-en CISLOC-NEUTER.AGT-garden-EP-lie-ST) “there.a.garden.lies” = “there is a garden there”.

The lexical and discourse functions of incorporation are not necessarily distinct. A given combination may both constitute a recognizable lexical item and regulate the flow of information.

1.3 Productivity

Incorporation can be highly productive in Mohawk. The pervasiveness of incorporation in the language has sometimes been taken as evidence that it is a syntactic process (Baker [1988:80] and others). The productivity is different in kind, however, from that of many syntactic constructions. It is like that of word-formation processes in general, tied to individual morphemes. The verb root -ianer- “be good”, for example, never incorporates. The verb -ken “see” almost never incorporates, though it frequently appears with an independent noun identifying the object seen. (Only one rare example has ever been encountered, a combination -ia’t-ken “body-see” = “be conspicuous”.) The verb -ientho “plant” appears often with an incorporated noun, as
The productivity of individual verb roots varies along a continuum. Noun stems show a similar range of incorporability: some are never incorporated, some rarely, some often, some usually, some always.

The productivity is ultimately a lexical matter, not fully predictable on grammatical or semantic grounds. The root -ianer-, which never incorporates, and -iio, which always incorporates, have approximately the same meaning: “be good/nice”. The variation is not entirely arbitrary: certain factors contributing to the patterns can be discerned (Mithun 1984). Noun and verb roots denoting recurring elements of name-worthy concepts often appear in names for them, for example. Among the most frequently incorporated nouns are those that are classificatory in useful ways. Mohawk terms for mental activities and states often incorporate the noun -nikonhr- “mind”; tho nikonhratihéthon “he is attracted to it” (“it/him-mind-pulls”); those with physical effects on living beings often incorporate the noun -ia’l- “body”; sakomwa’ia tenhá:wihte’ “they led them away (bodily)” (“again-they/body-carry-cause”); those pertaining to more abstract matters often incorporate the noun -rihw- “matter, affair, idea, word, fact...”: iah tehorihwató:ken “he wasn’t consistent” (“not was-he-idea-certain”).

It might be expected that incorporating constructions used for discourse purposes should show few constraints on productivity: in principle, all nouns that could represent established information should be incorporable. Even here, however, limitations can be seen. Relatively general terms typically stand in for large sets of more specific ones. There is no grammatical reason why nouns for domestic animals should not be incorporated, but they almost never are. In their place, the general root -nahskw- “domestic animal” is used. Similarly, there is no grammatical reason why the word for “bell” should not be incorporated. It is a deverbal nominal (“one metal-strikes with it”), but many deverbal nominals are incorporated (with a nominalizing suffix). In its place, the root -hwist- “metal” is incorporated.

(6) Incorporation of general terms

Wahatihwistantiion’tá:ko’
wa-hati-hwist-a-niion’t-ako’
FACT-M.PL.AGT-metal-EP-hang-CAUS-REV-PRF
“It (the bell) was taken down (from the church steeple) . . .”

The incorporated noun -hwist- is used in this way for a variety of metal objects, including stoves and money. If incorporation were a purely syntactic process, like clefting or topicalization, we would expect the full range of nouns to be incorporated. Incorporating constructions functioning on a discourse level are often highly productive, such as those involving presentative verbs, but it is not unrestricted. The
situation is not unlike that of English compounding. The English noun player is easily combined with almost any noun for a game, sport, or musical instrument: Parchesi player, tiddlywinks player. (Some may of course not be used because of existing formations: violin player.) Others are somewhat less productive, such as the English ache: speakers know that there is a noun headache but no noun footache, though it could exist and its meaning would be clear.

1.4 Transparency

A feature closely related to productivity is transparency, the degree to which the meaning and structure of a complex form correspond to those of its parts. The transparency of Mohawk incorporating structures has sometimes been taken as evidence of their syntactic status (Baker [1988:80] and elsewhere). A great many complex verbs in Mohawk are indeed fully transparent, like wakts'i'siai'énthon “I flower-planted” in (1). Their meanings are completely predictable from their parts, and the relation between the flowers and the planting is clear. Speakers may not be able to isolate individual morphemes on request, but they know which meanings are represented in the word. Other combinations have idiomatic meanings, but their literal meanings are still readily available to speakers, as in eniewirahni:non “she will baby-buy” = “she is expecting”. Still others show considerably less transparency, such as atekhwà:ra “it has food set on it” = “table”. Speakers have expressed surprise at discovering that the word for “table” contains the incorporated noun root -khw- “food”.

When Mohawk incorporating constructions are examined as a group, it becomes apparent that they show a cline of semantic and structural transparency, ranging from those that are fully transparent, like “flower-plant”, through those that are idiomatic but still transparent like “baby-buy”, through those whose original meanings can be recovered only with effort, and ultimately to many whose origins are inaccessible. Again, the situation is not unlike that of English compounds. Most speakers are aware of the rationale behind ice-cream or forthcoming, perhaps somewhat less of cockpit and handkerchief, and even less of foolscap or rigamarole.

In part because incorporated nouns can function in Mohawk discourse to keep reference clear, it has been suggested that they differ sharply in referential status from nouns in compounds like English head-ache. They are claimed to be syntactically equivalent to free nominals. Sentences like (7) from Baker (1997) have been presented as evidence for this view. (Spelling and morphological analysis have been regularized.)
Referential transparency test: (Baker 1997)

\[
\text{Wa'onkonhsóhare'}
\]
\[
\text{wa'-onk-kons-ohare-'}
\]

\[
\text{FACT-INDEF.AGT/ I SC. PAT-}\text{face-wash-PRF}
\]

“She washed my face.”

\[
\text{tanon kwa shé:kon ioná:na} \text{nawen}
\]

\[
\text{tanon kwa shé:kon io-na}'\text{naw-en}
\]

and even

\[
\text{NEUTER.PAT-be.wet-ST}
\]

and it (the face) is even still wet.”

When presented with (7) and asked “What is wet?”, speakers indeed have little trouble identifying the face. It is significant, however, that such structures are not usual in speech. The contrast points to the need for care in the data-gathering process. Particularly in a language like Mohawk, where so many grammatical choices are triggered by discourse factors, and lexicalization plays a pervasive role, grammatical patterns emerge more systematically in natural speech than in constructed examples. The sentence in (7) was apparently constructed on an English model then presented to bilingual Mohawk speakers for judgments about reference. Even out of context, if speakers construct the Mohawk translation themselves, they come up with a form that differs in a crucial way in the second line.

Speaker-generated Mohawk: (Kanerahtenhá:wi Nicholas, speaker)

\[
\text{... Shé: ki’ wakkonhsaná:wen}
\]

\[
\text{shé:kon ki’ wak-konhs-a-naw-en}
\]

still just

\[
\text{1SG.PAT-face-EP-wet-ST}
\]

“it’s still wet.” (literally “I’m still face-wet”)

In the speaker-generated translation, the incorporated noun is repeated in the second clause: “I’m still face-wet”. When presented with the alternative lacking the incorporated noun, some speakers have hesitated a moment, but all have responded that the incorporated noun should be included.

Even if sentences like (7) were to be taken as evidence that incorporated nouns function syntactically like full lexical noun phrases, it would first have to be demonstrated that the reference of the pronominal prefix io- “it” depends on a full antecedent noun phrase. A similar construction can be seen in (9), from a spontaneous conversation.
(9) Referential transparency in natural speech: (Awenhráthen Deer, speaker)

\[
\begin{align*}
\text{They were renting a house there,} \\
\text{not at all it wasn't expensive at all.}
\end{align*}
\]

There had been no previous mention of the house in this conversation, so it does appear that the antecedent of ka- “it” in “it wasn’t expensive” must have been the incorporated noun -nonhs- “house”. The speaker continued (in Mohawk): “That’s why, because it wasn’t expensive, it used to feel nice. And then they would cook, put a feed on once a month. It was this woman K. We used to go there with my mother to eat. And it was so close, such a short distance ...”

(10) Reference without incorporated noun: (Awenhráthen Deer, speaker)

\[
\begin{align*}
\text{And we just ate there,} \\
\text{not at all it wasn’t expensive at all.}
\end{align*}
\]

This time the same pronominal prefix ka- appears in “it was not expensive”, but there is no noun, incorporated or free, to serve as an antecedent. Was it the food, or perhaps the eating? Reference appears to be constructed by pragmatic inference. Incorporated nouns can certainly aid in inference, but they are not crucial. (Similar points have been made for other languages, as in Cornish 1986 among others.) The fact that incorporated nouns can assist speakers in determining reference is not surprising and does not in itself prove that they are different in kind from the components of compounds. Presented with the sequence I have a headache; it really hurts, few English speakers would have problems determining what hurts.

1.5 The product

Incorporation is indeed formed according to a regular pattern in Mohawk: a noun stem representing a semantic patient is compounded with a verb root to yield a
complex verb stem. But regularity of formation does not automatically mean that the formation is a syntactic process. Regularities are characteristic of productive grammatical processes in general, both syntactic and morphological. They differ in the nature of their products.

The essential difference between regular syntactic processes and Mohawk noun incorporation is that incorporation produces lexical items. It is not the usual mechanism for on-line production of connected speech. The resulting complex verb stem may then be stored and accessed as a unit on subsequent occasions. Where incorporation is highly productive, as it is in Mohawk, there are large numbers of such derived verb stems, and they appear pervasively in natural speech. There is ample evidence that speakers are generally aware of the fact that they are creating a new word when they use a new combination; such occasions are often accompanied by obvious delight. Fully fluent speakers vary widely in the frequency with which they innovate, and special talents for creating new forms are recognized and appreciated. Listeners generally know whether they have heard a particular combination before. They often describe the circumstances under which they first heard a combination, and they can identify the person who uses it. After visits to other Mohawk-speaking communities, they routinely remark on unfamiliar noun-verb combinations encountered, even when the novel forms are completely regular and transparent. Speakers helping to transcribe and translate tape-recorded speech from other communities constantly point to new combinations in the same way.

These lexical items are not exactly the same units as the phonological words described in section 1.1, but there is a close relationship. For the most part, they are the bases of the phonological words, the units that become phonological words in Mohawk with the addition of inflection. In general, speakers have clearer memories of lexical items than of phonological words. Thus they are likely to know whether or not they have heard the stem -tsi’tsiaientho- “flower, plant” of the word waktsi’tsiaienthon “I’ve planted flowers”, but not whether they have heard it with all possible pronominal prefixes. Speakers are typically aware of creating something novel when they form a new noun-verb combination via incorporation, but not when they use an existing combination with a different pronoun. Lexical items do not always correspond to stems, but this appears to be the most common state of affairs, particularly in Mohawk.

The fact that incorporation creates lexical items accounts for the continua of both productivity and transparency. The varying degrees of productivity are tied to individual morphemes, as in other word-formation processes. The gradation in transparency arises from two factors, one the initial creation of the forms, the other their subsequent use.

First, words are coined for a purpose, to allow speakers to package single
concepts into single words. The purpose may be to add new terms to the stock of vocabulary, or to regulate the flow of information in discourse. Combinations formed for discourse purposes often tend to be transparent, while those formed to enrich the vocabulary may be less so. A new vocabulary item may not be associated with its full range of possible literal meanings. The term *iehwista 'êkstha* ‘bell’, for example, is literally “one metal-strikes with it”, and speakers can generally retrieve its literal meaning on request. It is not used for just any object used to strike metal, however, such as the range of tools ones might find in a garage.

A second factor contributing to the cline of transparency is the fact that the products of the process are lexical items, with their own integrity. They are learned, stored, and accessed as units. As such, they may shift in meaning and function over time, without regard to their original components. These shifts are made possible by a side effect of the creation of lexical items, a reduction in the specification of internal structure.

2. **Grammatical relations**

The core arguments of Mohawk clauses are generally easy to identify. Every clause contains a finite verb with pronominal prefixes specifying its core arguments, whether independent noun phrases are present or not. The prefix on an intransitive verb refers to one participant, like *ra-* “he” in (11).

(11) Intransitive pronominal prefix: (Konwats'tsaien:ní Phillips, speaker)

\[
\begin{align*}
\text{Ratákhe'} \\
\text{ra-takh-ë'} \\
\text{M.AGT-run-IMPRF} \\
\text{“He is/was running”}
\end{align*}
\]

The pronominal prefix appears on the verb even if a coreferent nominal also appears in the clause, as in *shaia:ta ratákhe*’ “one boy was running”. A transitive pronominal prefix refers to two participants.

(12) Transitive pronominal prefixes: (Akwrâ:’es Natawe, speaker)

\[
\begin{align*}
\text{Tahshakóhsere'} \\
\text{ta-hshako-hser-ë'} \\
\text{CISLOC-M.AGT/3PL.PAT-chase-IMPRF} \\
\text{“He was chasing them.”}
\end{align*}
\]

Again, either or both participants may be further identified by independent nominals: *Kaspé tahshakóhsere’ ne ronón:kwe* “Kaspé was chasing the men.”

There are three basic pronominal prefix paradigms, two intransitive sets and one transitive set. The first intransitive set (I), usually termed the Agent set, typically represents participants who instigate and control intransitive events, like the prefix...
ra- “he” referring to the runner in (11). The second intransitive set (II), usually referred to as the Patient set, typically represents participants who are not in control but are affected, such as people who are ill or tired. The transitive set represents an Agent/Patient combination, like -hshako- “he/ them” in (12). Paradigm choice is categorical and lexicalized with individual verbs, though a semantic motivation can usually be discerned (Mithun 1991). Speakers do not select cases as they speak: some intransitive verbs are simply lexicalized with Set I pronouns, and others with Set II. The appropriate paradigm choice is learned with each verb. There are no independent pronouns equivalent in function to the unstressed pronouns of English. That function is served by the pronominal prefixes, which differ primarily in the fact that they are present in every clause, even when noun phrases are present as well.

Among the third person pronominal prefixes, three genders are distinguished: masculine, feminine/indefinite (“she, one”), and neuter/zoic. (Neuters are used for inanimate objects and zoics for animals, but in most contexts the two categories are expressed with the same forms.) Neuters are represented only when no other participant is present. Thus the prefix on ra-hsere’s “he’s chasing (it)” is simply the masculine singular agent ra- “he”, the same as the prefix on ra-tákhe’ “he is/was running”. As a result, there is often no formal distinction between transitive and intransitive verbs. The same verb is used for “I’ll plant” and “I’ll plant it”.

(13) Mohawk ambitransitivity

\[
\begin{align*}
\text{Enkïëntho'}
\end{align*}
\]

FUT-1SG.AGT-plant-PRF

“I’ll plant”, “I’ll plant it”

Spontaneous intransitive use of the verb “plant” was seen in (3) above “May I plant here?” No crop had been mentioned or even selected. In (14), the same verb is clearly used transitively with a nominal identifying the crop.

(14) Mohawk ambitransitivity

\[
\begin{align*}
\text{Ó:nenhste’} & \quad \text{enkïëntho'}
\end{align*}
\]

o-nenhst-e’ en-k-ientho’

NEUTER-corn-NOUN,SUFFIX FUT-1SG.AGT-plant-PRF

“I’ll plant corn.”

Thus apart from actions directed at persons, which require the specification of two parties in their pronominal prefixes, grammatical transitivity is not systematically distinguished overtly for verb stems, since neuters are not mentioned in the prefixes if any other participant is involved. Some stems never co-occur with a separate nominal identifying a patient, such as -e- “go”; such distributional evidence points toward intransitivity. Many others, like “plant”, occur sometimes with a separate
patient noun and sometimes without. When no such noun is present, the verb may sometimes be interpreted as transitive with specific patient ("I planted it"), and sometimes as intransitive ("I’ll do the planting"). The existence of a patient (crop) might of course be implied, just as with the English eat (food) or sing (a song).

In any case, it is clear that incorporated nouns are not treated grammatically as core arguments. Because the vast majority of incorporated nouns are neuter, and neuter participants would not be represented by pronominal prefixes in any case, the distinction is not usually evident. In the relatively rare instances where an incorporated noun represents a human being, the pronominal prefix does not refer to it as a core argument. The verb in (15) was used to describe the habits of a monster.

(15) Incorporation of an animate: (Konwati’tsaién:ni Phillips, speaker)

\[\text{Ratonkwe ti:saks} \]
\[\text{ra-at-onkwe ti-sak-s} \]
\[\text{M.AGT-MIDDLE-person-seek-IMPRF} \]

"He hunts people."

The pronominal prefix ra- is intransitive, referring only to "he". If the people eaten were core arguments, the pronominal prefix would be -hshako- "he/them", as in (12) above, but such a form is unacceptable. If the role of the incorporated noun is no longer overtly mentioned within the pronominal prefix, can we still say that it is specified? We might be able to maintain that even if it is not specified, it is recoverable, since only semantic patients are incorporated in Mohawk.

In many languages, all verbs containing incorporated nouns are intransitive. In Mohawk, verbs with incorporated nouns may be lexicalized as intransitive, transitive, or both, though, as we have seen, the distinction is not always clear. Often the incorporated noun is a general term that serves to narrow the scope of the verb, so that it denotes a kind of event or state appropriate for a class of entities. An independent nominal (consisting of a noun, descriptive term, demonstrative, or some combination of these) may then further identify the particular entity involved. Incorporation of this type has been termed ‘classificatory’ (Mithun 1984).

(16) Incorporation with separate nominal: (Konwati’tsaién:ni Phillips, speaker)

\[\text{Wahonhiákha'} \]
\[\text{wa-hon-ahi-ak-h-a'} \]
\[\text{FACT-M.PL.AGT-fruit-pick-PURP-PRF} \]

They went to fruit-pick

\[\text{sewahio:wane'} \]
\[\text{s-w-ahi-owan-e'} \]
\[\text{one-NEUTER-fruit-be.large-ST} \]

"They went to pick apples."

(The term sewahio:wane’ is a verb “it is the one that is fruit-large” that has been lexicalized as a nominal “apple”.) Structures like that in (16) could be taken as
confirmation of the grammatical relation of the incorporated noun, as a patient or object, since it could be analyzed as co-referent with the independent noun phrase which appears to be filling this function.

Yet the incorporated noun and independent noun phrase need not be coreferent. Constructions like that in (17) below are common. The incorporated noun is -hso 'kw-' "head", but the grammatical arguments of the clause are both masculine, as can be seen by the pronominal prefix ho- "he/him". The independent noun refers to the man, not his head.

(17) Coreference not necessary: (Kaia'itáhkhe" Jacobs, speaker)

\[\text{TRANSLUC-Dc-M.AGT/M.PAT-head-hit-PRF} \]
\[\text{"He bopped the man who was laid out [in the casket] on the head."} \]

Such constructions, sometimes referred to as 'external possession', are common cross-linguistically. They serve an important function, permitting speakers to cast people rather than their parts as arguments (Mithun 1984).

We might try to salvage the retrievability of grammatical relations by stipulating that when a possessed entity is incorporated, the possessor is automatically registered as a grammatical argument of the clause. The incorporated possession remains the patient (or object). Incorporating a possessed noun does not necessarily entail a change in the argument structure of the clause, however. Possessors appear as core arguments only when the speaker expresses more interest in the effect of the situation on them than on their possessions. The sentence in (18) involves action directed at a possession, the door: "The two reached the door of an old couple". The verb contains the incorporated noun root -nhoh- "door", but the pronominal prefix is intransitive: -hni- "they two (masculine)". The old folks are the possessors of the door, but they are not sufficiently affected by the event, nor important enough to the narrative, to be portrayed as core arguments of the clause ("they/them-door-reached")?

(18) Incorporated possession, no argument shift: (Konwatsi'ts'aien:ní Phillips)

\[\text{TRANSLUC-FACT-M.DU.AGT-door-reach-PRF} \]
\[\text{"They (two boys) just barely reached the door of an old couple."} \]
It might be hoped that the recoverability of grammatical relations could be rescued by distinguishing alienable from inalienable possession, as proposed in Baker (1997). The man’s head in (17) is inalienable (inseparable), but the couple’s door in (18) is alienable. We might stipulate that possessors are coreferent with their inalienable possessions: if you hit his head you hit him. Both the head and its owner could then occupy the role of grammatical patient. The distinction between inalienable and alienable possession is in fact expressed formally elsewhere in Mohawk, in the choice of possessive prefixes on nouns (Mithun 1995). Furthermore, inalienable possessions are more often incorporated than alienable ones, and their possessors appear more often as core arguments of the clause.

Neither pattern is a mechanical consequence of alienability marking, however. Situations involving inalienable possessions like body parts simply significantly affect their owners more often than those involving alienable possessions. Alienable possessions may be incorporated as well, and their owners made core arguments, but only if they are substantially affected. Harnesses are classified grammatically in Mohawk as alienable, for example. Yet in (19) “he unharnessed the horses”, the horses still appear as core arguments, because the effect of the action on them was more important than its effect on the harness.

(19) Incorporation of alienable possession: (Tekaronhi:ken Jacobs, speaker)

Wahshakohkweniahrá:ko’
wa-hshako-akhwenni-hr-ako-
FACT-be/them-harness-be.on-REV-PRF
ne ako-aksatens.

“He unharnessed the horses.”

The grammatical relation of the incorporated noun in Mohawk is thus not specified formally, nor mechanically interpretable from form alone. Grammatically, incorporated nouns are simply modifiers of the verb root. In Mohawk, incorporated nouns typically represent semantic patients, entities centrally involved but not in control of events or states, but this pattern is not a necessary property of incorporated forms. Cross-linguistically, entities invoked by incorporated nouns can be seen to bear a variety of semantic relationships to their associated verbs, most often semantic patients, instruments, and/or locations (Mithun 1984).

The relation of the entity invoked by the incorporated noun to the predication is understood by pragmatic inference, in much the same way as the relation of horses to horseshoes and alligators to alligator shoes in English. Incorporated nouns indicate the involvement of an entity that qualifies the verb in some unspecified way. There is no internal syntactic structure specifying grammatical relations.

The lexical status of incorporating constructions can have important grammatical consequences. Since they are lexical items, their meanings are interpreted as
unitary concepts, without formal articulation of the specific roles of the elements to each other. Particularly when they are first formed, the role of the incorporated noun may still be deduced, in part because in Mohawk it is semantic patients that are incorporated. Once a word has been formed, however, it has a semantic life of its own, without regard to its internal structure. As seen in (16) and (17) above, Mohawk transitive verbs containing incorporated nouns may co-occur with independent noun phrases further identifying either the referent evoked by an incorporated noun ("fruit-picking apples") or its possessor ("head-hit the man").

When we examine the range of use of incorporating constructions over longer stretches of speech, however, we see that the relation between an incorporated noun root and external nominal are not limited to these types. The incorporated noun -renn- "song" in (20) is neither coreferential with nor the possessor of o'nó:wa' "guitar". The incorporated noun -renn- is simply an element of the complex verb -renn-ot- "sing, play", appropriate for music. Singing/playing can be used intransitively or transitively.

(20) Mohawk incorporation: (Awenhrathen Deer, speaker)

O'nó:wa'
ne: ' thaIeren nótha'.
o- 'now-a'
ne: ' e t-ha-ate-remn-ot-ha'

NEUTER-guitar-NOUN.SUFFIX it.is DC-M.AGT-MIDDLE-song-stand-IMPRF
guitar it.is he song-stood

"He played the guitar" ("He song-stood the guitar"???)

It is not difficult to see how the structure in (20) might have come into being. The verb stem -renn-ot- "song-stand" = "sing, play" was transparent when first formed, and soon came to co-occur with independent nominals identifying particular songs. With use, speakers became less and less conscious of its original internal structure, and it was simply associated with a concept "sing/play", and easily extended to contexts like that in (20).

A similar situation can be seen with the verb -nia'kwakwarihsi "drink straight down" in (21), literally "neck-straighten"). The relation between the incorporated -nia'kw- "neck" and the independent nominal "nice liquid" is neither classification nor possession.

(21) Mohawk incorporation: (Margaret Edwards, speaker)

Kwah thiahania kwakwarihsia' te'
kwah th-i-a-ha-nia'kw-kwari-hsi- 'i-

just CONTR-TRANSLOC-FACT-M.AGT-neck-be.crooked-REV-CAUS-PRF
just he neck-straightened
The complex verb stem was apparently originally formed by incorporating the semantic patient "neck" into the verb "straighten" ("cause to be uncrooked"). The combination was used in contexts where someone drank something in a single draught, without separate swallows. Once the verb stem was formed, however, its meaning "drink straight down" took over, and its internal structure faded in the minds of speakers. At that point, it could be extended to contexts like that in (21). It should be noted that such an extension does not entail that speakers cannot retrieve the literal meaning of the expression on request.

3. The status of incorporating processes and constructions

At first glance, it is easy to see the appeal of analyzing incorporation as a syntactic process. First, in many languages, incorporation can seem to produce sentences, like Mohawk wa'tsi'tsiái'énthon "I’ve planted flowers". In such languages, however, verbs without incorporated nouns can also constitute sentences: wakiéntho "I’ve planted, done my planting".

Second, incorporation operates on morphemes with meanings comparable to those of words in many languages, like -tsi 'tsi- "flower" and -ientho "plant". If such comparisons were sufficient, however, we should consider the formation of the English un-happy to be a syntactic process as well, since negation in English is usually accomplished with the word not.

Third, incorporation can have syntactic implications, most clearly in its capacity to alter argument structure. It can convert a verb like Mohawk -ohare "wash", whose core arguments are the washer and the object washed, to an intransitive verb like -nonhs-ohare "houseclean", whose only core argument is the housecleaner (wa'kenonhsóhare "I did the cleaning"), or to a different transitive verb like -konhs-ohare "face-wash", whose arguments are the washer and person affected by the washing: wa'khkonhsóhare "I face-washed her". Such effects can be produced by derivation as well, with antipassives, causatives and applicatives. In fact even lexical choice can affect argument structure, such as substituting "feed" for "eat", or "enter" for "go". If effect on argument structure is taken as evidence of a syntactic process, then lexical choice should be considered a syntactic process as well.

Specific arguments put forth for viewing incorporation as a syntactic process have depended to some extent on the model of syntax assumed. In a generative
framework, noun incorporation has been viewed as the result of syntactic movement (Move-Alpha), which leaves behind a trace coreferent with the incorporated noun (Baker 1988, 1995, 1996). Because of the coreference, the trace is expected to show the anaphoric properties of a lexical noun. The derived complex verb is assumed to be intransitive; any separate nominals that occur with it are adjuncts outside the clause. Arguments for this view contrast it with a different movement process, identified as lexical, that would leave behind an empty pro, an invisible referential pronoun. This process is said to derive a transitive verb. Within this theory, then, arguments for the analysis of incorporation as a syntactic process have focused on anaphora and transitivity.

Unfortunately, the two syntactic tests proposed for anaphora, involving Condition C of the Binding Theory and licensing and agreement, are applicable only to a narrowly restricted set of forms, those whose incorporated nouns are animate. But animate nouns are almost never incorporated in Mohawk, so of the few constructed combinations for which the test might be valid, nearly all are highly unnatural, unacceptable on other grounds, making grammaticality judgments difficult to assess.

A third test proposed for determining the syntactic status of the process focuses on transitivity by examining question formation. It is argued that sentences like (22) cannot be transitive, because there is no object position for the question word (Baker 1995, 1996). The intransitivity is taken as proof of the syntactic status of the incorporation process. (Transcription and morphological analysis have been regularized.)

(22) Picking up: (Baker 1995:18, 1996:323)

?*Ônhka’ tenhsewi:rahkwe’?
ônhka’ t-en-lse-wir-a-hkw-e’
who DC-FUT-2SG,AGT-baby-EP-pick.up-PRF

**“Who are you going to pick up (a baby)?”**

The sentence is actually problematic for other reasons. Picking up a baby is not expressed by incorporating the noun root -wir- “baby” into “pick up”. It is an action physically affecting an animate being, so the derived stem -ia’t-a-hkw- “body-pick.up” is used.

(23) Picking up: (Kanerahtenh6:wi Nicholas, speaker)

Tenhonwaï:a:tahkwe’
t-en-honwa-ia’t-a-hkw-e’
DC-FUT-INDEF,AGT/M,PAT-body-EP-pick.up-PRF

“She’ll pick him up.” (“She will pick him up bodily.”)

This verb is transitive, as can be seen from the pronominal prefix -honwa- “she/him”; it may co-occur with a separate nominal identifying the child. It can also be
used in "who" questions. Asked whether the noun -wir- "child" could be incorporated into the verb -hkw- "pick up", speakers consulted are categorical in their judgment: the form would not be Mohawk.

Even if the verb stem -wir-a-hkw- "baby-pick.up" did exist, the structure in (22) would be unacceptable for other reasons. If the noun "baby" were incorporated, it would be for discourse purposes, to background established or secondary information. For this reason, it is incompatible with a "who" question, where the focus is on discovering the identity of the one picked up. Nouns incorporated for discourse purposes can of course appear with questions formed with ka’ nika:i'en "which one". "Which" questions presuppose the previous identification of a relevant set of possibilities, just the conditions conducive to incorporation.

The same question test was used to classify the combination -ks-ohare- "dish-wash" as syntactic (Baker 1995, 1996). Incompatibility with the question word nah:ten "what" was viewed as a sign of intransitivity, which was taken in turn as proof of the syntactic status of incorporation.


*Nah:ten wahseksohare'?
*naho'ten wa-hse-ks-ohare-’?
what FACT-2SG.AGT-dish-wash-PRF
*"What did you dish-wash?’’

The question in (24) is indeed ungrammatical and the verb intransitive. There is no reason, however, that the intransitivity should be taken as proof of the syntactic status of the incorporating process, apart from Baker’s own premise that lexical incorporation must create transitive stems. If, however, a lexical process is identified simply as a process that creates lexical items, to be stored and accessed in speakers’ minds as units, it is not surprising that it should be able to create intransitive, transitive, and ambitransitive stems, mirroring the range of underived stems. The stem -ks-ohare- "dish-wash", for example, was coined as a label for a recurring cultural activity, like ironing or cooking. It does not generally appear with an object nominal because none is necessary. The term -ksa’t-iio- “child-be.good” = “be a good child” is also intransitive. The stem -ia’t-a-hkw- "body-pick.up" for picking up a person is typically transitive, occurring with an agent (the one picking up) and a patient (the one picked up). The stem -hnek-ihr- “liquid-consume” = “drink” can function either way: Enhshneki:ra’ ken? “Will you have a drink?” Nah:ten enhshneki:ra’? “What will you drink?” The transitivity of lexical items derived by incorporation in Mohawk is not fully predictable by general rule: it depends on the concept to be named.
As Baker realized (1995, 1996), his tests for the syntactic status of incorporation do not yield clear results. Some incorporating constructions would be classified as syntactic in his scheme ("dish-wash", "wood-cut", "baby-pick.up"), some would be classified as lexical ("wood-get", "liquid-drink", "liquid-throw"), and some are un-classifiable because speakers are too uncomfortable with the diagnostic constructions ("meat-buy", "baby-lose", "fruit-pick"). As can be seen, these results do not coincide with any of the usual views of the lexicon. For Baker, the primary issue in the study of incorporation is "the interesting question of how a language learner decides whether a given morphologically complex form is derived lexically or syntactically" (1995:4). If we simply assume that children learn the words they encounter, then deduce patterns for forming more as needed, this problem disappears. The fact that all of the combinations follow a clear pattern simply makes learning them easier.

What it means to say that a process should be viewed as lexical depends of course on the view of the lexicon assumed. One view has been that the lexicon is only the repository of irregularities. A more comprehensive view, and that adopted here, is that the lexicon is the set of lexical items that speakers know. New items may be added to the lexicon at any time. The new items may be borrowed, or they may be constructed according to regular rule, a common phenomenon in a language with morphological processes as rich and productive as those of Mohawk. These processes differ crucially from syntactic processes in that they are not the primary mechanism for producing on-line speech. They create lexical items to be learned, stored, and accessed as units.

Recognition of the fact that incorporation is primarily a word-formation process has several advantages. It permits the uniform treatment of all noun incorporating constructions in Mohawk, constructions which show the same formal characteristics. All are created according to the same word-formation process, by which a noun stem is compounded with a verb root. In all formations, the noun, which invokes a semantic patient, precedes the verb. The noun must be a formal morphological noun stem, with an overt nominalizing suffix if it is deverbal: wa'ke'serehtahni:non "I drag-thing-bought" = "I bought a car". If the incorporation would result in a consonant cluster, a non-stress-bearing stem joiner -a- is added: -'sereht-g-hninon- "car-STEM JOINER-buy".

At the same time, it explains certain rare forms that do not follow the pattern exactly. Incorporation in Mohawk involves the combination of one noun stem with one verb stem. Yet a very few complex verb stems appear to contain two incorporated nouns, as in (25).
(25) Double incorporation? (Niionhi: ’a Montour, speaker)

Enkonia’thahónni
en-kon-ia’t-hah-omni-en’
FUT-1SG.AGT/2SG.PAT-body-path-make-BENEFACTIVE-PRF
“I shall make a path for you” = “I shall guide you”

This construction could be formed only in several diachronic stages. One stage involved the creation of the complex verb stem -hah-onni “path-make”, a useful vocabulary item. Once it had become an established lexical item, it was reinterpreted as a simplex verb, and again underwent the process, this time incorporating the noun root -la’t- “body” to signal an action with a physical effect on another living being.

The lexical status of incorporation also explains its variable productivity. As is typical of word-formation processes, the productivity varies with the individual morphemes combined. It also accords well with what speakers report knowing about their language. The possibility of a mental lexicon containing all the words (stems) of a language has on occasion been rejected, on the grounds that 1) speakers could not possibly keep track of the words they have heard and 2) that such a model would mean that different speakers of the same language must have different lexicons (Baker 1995). The first objection has been laid to rest by numerous studies showing that speakers not only know which words exist, they also have surprisingly accurate knowledge of their relative frequencies (Underwood 1966, Shapiro 1969, Carroll 1971, Gordon 1985, Connine et al. 1990). Work with speakers of Mohawk indicates clearly that they certainly have detailed knowledge about which stems exist in the language and which could but happen not to. Furthermore, we find just the lexical variation among speakers that such a theory would predict. Mohawk is currently spoken in six communities in Quebec, Ontario, and New York State. The majority of the lexicon is shared, but speakers are keenly aware of differences among communities, generations, and even individuals.

Viewing incorporation as a word-formation process also accounts for the variation we find in semantic and grammatical transparency. Many incorporating constructions are fairly transparent semantically and grammatically. Speakers are aware of their components, and the combinations yield predictable results. Many others have idiomatic meanings like eniewrahninon “she will baby-buy” = “she is expecting”, or unpredictable argument structures like thihania’kw-akwarihsia’te’ “he neck-straightened” = “he drank it straight down.” They may be intransitive or transitive. Moreover, these different features do not covary systematically: a combination with a highly idiomatic meaning may still be fully transparent structurally, and its components easily accessible to speakers. Alternatively, a combination may be regular in every way except for its argument structure, or the accessibility of its components to speakers. The independent variation of these factors poses a problem
for models that differentiate two types of incorporation, syntactic and lexical. Different criteria for classification give conflicting results, as discovered by Baker (1995, 1996).

In the end, when all incorporation is understood as a process of word-formation, the various idiosyncrasies are not surprising. Lexical items are created for specific purposes, and are subsequently free to change over time as units, without regard to their original internal structures.

REFERENCES


