Linguistics in the face of language endangerment

Marianne Mithun

To appear in:

In linguistics, as in most science, the current state of theory shapes our interests and observations. We are attracted to issues we feel we can shed light on. We value data that confirm or disconfirm current hypotheses. Theoretical awareness can help us to perceive important facets of our data, but it can also lead us to overlook data relevant to issues not yet addressed in the discipline. In the normal course of science we can fill in data as our theoretical models evolve: each new empirical finding can raise new questions, which in turn send us back for more data, pushing the discipline ahead step-by-step. Unfortunately, the current rate of language disappearance may not allow us the luxury of such step-by-step development. Much of our potential empirical foundation may be gone before we are in a position to be curious about it. We need to think now about how to provide answers to questions we do not yet know enough to wonder about.

The disappearance of a language is of course much more than a missed opportunity for science. It is a tragic loss for the descendants of the last speakers: a loss of intellectual, cultural, and social heritage. Once a language is gone, there is little that future generations can do to recover it. Now is the time for linguists, speakers, and communities to pool ideas about what descendants of speakers may want to know.

1. Empirical foundations for the future
If we are to provide a foundation for future discoveries in the discipline, as well as a meaningful record for local communities, we should consider the kinds of choices to be made in fieldwork when time is limited. Some of these possible decisions will be illustrated here with examples from Mohawk, a language of the Iroquoian family spoken in northeastern North America in Quebec, Ontario, and New York State. Half of the sentences in (1) originated in Mohawk; they were drawn from spontaneous conversation. The other half were elicited translations of English sentences, constructed to address specific theoretical issues in linguistics; they were drawn from publications by another linguist, but spellings and analyses have been regularized. The reader is encouraged to guess the origin of each sentence below, spontaneous or constructed, before reading further.

(1) Which sentences originated in Mohawk?

a. Owirà:'a wà:rate' ne o'wà:ron'.
   baby he ate the meat
   The baby ate the meat.

b. Sak iehiatónhkwà' wahshakó:ion' ne Onwá:ri.
   (name) pencil he gave her the (name)
   Sak gave Mary a pencil.
   OK you go and bug cause to lie
   OK, you go shake the bugs off.

d. Sak ñ:rehre’ Onwá:ri wa’ontia’tóhare’.
   (name) he thinks (name) she herself body washed
   Sak thinks Mary washed.

e. ñ:rehre’ kati’ken óksa’k ken
   he thinks perhaps just immediately INTERROGATIVE
   Well, what did he think,

   enthowaìr’othahse’ ne ohwîsta’?
   they will bag stand up for him the money
   they’d just give him a bag of money right up front?

f. Ka’ níká:ien’ kanhóha’ wahsate’níen:ten’ Sak ahanhoton:ko’?
   which door you tried (name) he would open
   Which door did you try for Sak to open?

g. Nè:’e ká:ri’ ne’ nè:’e tho ni:ioht tsi
   it is in fact this it is there so it is thus as
   And that is how

   wahshakotianerónhkwen’ ne Kaspé ne thó shikahá:wi’.
   he was footprint itchy to them the Gaspe the there as it carries
   Gaspe scared those fellows that time.

h. Kwâh ñ:ken tsi wahtentia’tshera:kson
   quite it is so it is a bad place to set out
   It’s so uneven that

   ne taiontwálk ki: sidewalks
   the one would walk this sidewalks
   when you walk on the sidewalks

   thsinekwen’takarenhrátie’s.
   you are ankle slanting along habitually
   your ankles are always at an angle.

It may not be difficult to recognize sentences a, b, d, and f as the constructed examples, even without familiarity with Mohawk. After some experience with typical Mohawk speech, the distinctly un-Mohawk character of these sentences is even starker. What is it that distinguishes such sentences?

Vocabulary is certainly one indication. It is not simply a question of European words. The constructed example in (1)b contains the borrowed names Sak and Onwá:ri
(originally from French Jacques and Marie), but the spontaneous sentence in (1)g, a fine example of good Mohawk speech, contains the English noun sidewalks and the verb walk. Languages spoken in contact situations often contain loanwords, words used even by monolingual speakers. A more telling comparison is the close, word-for-word match between the constructed sentences and their English stimuli: Owirà:'a wà:rake' ne o'wà:ron' Baby ate the meat. Notably absent from constructed sentences are many of the particles that can pepper both conversation and narrative, as in Nè:'e kà:ti' ne' ne:'e tho ni:ioht tsi ... it is in fact this it is there so it is thus as ... And that is how ... Perhaps even more striking is the way information is packaged in the content words. What linguist would have had the imagination to pinpoint the word thsinekwen'takarenhrátie's your ankles are slanting as you go along as a Mohawk lexical item to be elicited?

The kinds of meanings expressed in the constructed translations are also different from those expressed in the spontaneous Mohawk remarks. Who in the course of normal fieldwork would elicit Go shake the bugs off? (This was a command given by a harried mother to her children when it was noticed that insects were eating their plants.) At the same time, an idea such as Which door did you try for to open? would probably never be expressed at all. Speakers usually choose to say other things.

2. Similarities and differences
But are such differences of theoretical importance? The discovery of general principles is certainly a central concern of most science. If our goal is the discovery of language universals, structural principles shared by all languages, the isolation of similarities among languages is key. The most efficient way to uncover similarities could be the elicitation of translations of sentences from a target language: How do you say ... ? Actually, if the formulation of universals is our ultimate goal, differences could be defined as theoretically uninteresting and subsequently ignored. Under such a view the death of individual languages would be of little consequence. But universals can be framed at a variety of levels of abstraction and in a variety of forms. Good generalizations often emerge best from an examination of arrays of differences. They may, for example, take the form of implicational universals or parameters: If a language has property A, it will also have property B. An issue that will surely be discussed throughout the development of our discipline is the relative value of similarities and differences among languages, and the relationships between the two. In any case, it seems clear that if we confine our documentation to immediately observable similarities, our understanding of language will necessarily be superficial.

It might be argued that elicited translations allow us to focus quickly and efficiently on specific points of pure syntactic structure. With elicitation, we need not wait for the crucial structure to appear haphazardly in speech, and the picture need not be clouded with the inconsequential idiosyncrasies of particular languages. Is there in fact anything of theoretical significance to be learned from the accidents of spontaneous speech? Consider the excerpt in (2) from a conversation among some friends, all excellent Mohawk speakers. The material is considerably less straightforward than the constructed sentences in (1) above and more difficult to process analytically.
(2) Mohawk in context: excerpt from a conversation

a. CB

Thó ken non: nihotenhenónhserote' ne:
tho ken non:we ni-ho-aten-hninon-hser-ot-e' ne:
there Q where PARTITIVE-M.SG.PATIENT-MIDDLE-buy-NMZ-stand-STATIVE it.is
there ? where so his selling place stands it is

Is that where the store was,

thikén: .. ratenohkwakéhtats rokstén:ha wáhi’?
thiken ra-ate-nohw-a-kehtat-s ro-ksen-’=ha wáhi'
that M.SG.AGENT-MIDDLE-bundle-JR-carry-HAB M.SG.PATIENT-be.old-STATIVE=DIM TAG
that he his bundle carries habitually he is old isn’t it
the one that belonged to that old peddler?

b. JD

A:, iáh ki' nè':e.
ah not just it.is

Mm, it wasn't that one.

karon ki’ nè':e na’-w-akwat-i
this side just it is PARTITIVE-NEUTER.AGENT-be.on.the.other.side.of-STATIVE
this side just it.is so it is on the side

That was on this side.

Né: kwi’ né’ thikén: ...
nè’:e ki’ wáhi’ ne’ thiken
it is just TAG it is that

Isn't it the one that's . . .

nahò:ten’ na’ ronwatí:iats ...
nà-’a ronwati-iats
PARTITIVE-NEUTER.AGENT-be.a.kind.of-STATIVE again M.PL/M.PL-call-HABITUAL
what again they habitually call them

what was their name . . .

o Konwákeri wi’ kénh niía:kahskwe’ wáhi’
oh Konwákeri wáhi’ kénh ni-iak-a-hskwe’ wáhi’
oh name TAG so PARTITIVE-F.SG.AGENT-be.a.size-PAST,HABITUAL TAG
oh Margaret isn’t it yea so she used to be big right

Oh, Margaret wasn’t it, . . . she was this big, wasn’t she.

c. DM

En.: Yes.
**d. JD**

Rón:tonhs  
kwi’  
ne’  
teiakotía’tanéká:ron.

ron-aton-hs  
ki’ wáhi’  
ne’  
te-iako-at-ia’-anekar-on

M.PL.ACENT-say-HABITUAL  
just TAG  
just  
DV-F.SG.PATIENT-MIDDLE-body-explode-STATIVE

They used to say her body had burst.

**e. DM**

Nia’té:kon

n-ia’-te-ka-on

PARTITIVE-TRANSLOCATIVE-DUPLICATIVE-NEUTER.AGENT-be.an.amount-STATIVE

it amounts to all kinds of things

That place used to

**tetkáhere’**

te-t-ka-her-e’

DUPLICATIVE-CISLOCATIVE-NEUTER.AGENT-be.on-STATIVE

it is setting on something there

be full of all kinds of things.

**Serónhké:nek**

se-ronhken-e=k

2.SG.AGT-struggle-STATIVE=just

you are just struggling

You had a hard time getting through.

**enwá:ton’**

en-w-at-on-

FUT-N.AGT-MIDDLE-be.possible-PRF

it will be possible

**enhsatóhetste’**.

en-hs-at-ohetst-e’

FUT-2.SG.AGT-MIDDLE-pass-PRF

you will pass through

**f. WS**

Io’nikonhrínekenht

io-’nikonhr-ineken-ht

NEUTER.PATIENT-mind-lead.out-CAUSATIVE.STATIVE

it makes the mind come out

It was overwhelming

**ienhsátáweia’te’**

i-en-hs-at-awei-a’-te’

TRANSLOCATIVE-FUT-2.SG.AGT-MIDDLE-enter-CAUS-PRAF

you will cause yourself to enter thither

when you went in there and saw all that stuff.

**g. CB**

Nahò:ten’

na’

tshitewana’tónhkhwahkwé’

na-w-o’ten-

PRT-N.AGT-be.a.kind.of-STATIVE  
again  
tshitewa-na’ton-hkw-hahkwé’

THI:KEN

What did we use to call that guy?
h. JD
Ro’tákeras.
ro-i’t-aker-as
MASCULINE.SG.PATIENT-excrement-stink-HABITUAL
he smells of excrement
“Ro’tákeras” [He smells of excrement].

i. WS

Ronohkwakéhte’
ro-nohk-wa-keht-e’
M.SG.PATIENT-bundle-JR-carry-STATIVE
he is bundle carrying
customarily
They used to call him “Ronohkwakéhte’” [He's carrying a bundle].

Enthatáweia’té
en-t-ha-ataweia’t-e’
FUTURE-CISLOCATIVE-M.SG.AGT-enter-CAUS-PRF
he will cause himself to enter hither
customarily
He’d come in with it tied up with cord.

O:
kwáh  shes  rotenohkwakéhte’.
kwáh  shes  ro-ate-nohkw-a-keht-e’
o:  kwáh  shes  kwáh  shes
oh quite customarily  M.SG.PATIENT-MIDDLE-bundle-JOINER-carry-STATIVE
he is his bundle carrying
Oh, he used to carry his bundle.

Kahón:tsi
ka-hon’ts-i
N.AGENT-be.black-STATIVE
it is black
ni:noht
ni-io-ht
PARTITIVE-N.PATIENT.be.so-STATIVE
so it is thus
thikén  ohnéhta’  wáhi’.
thiken  o-hneht-a’  wáhi’
that  N-pitch-NS  TAG
It was sort of a black gummy thing, you know.

j. CB
Ronohkwahón:tsi.
ro-nohkw-a-hon’ts-i
MASCULINE.PATIENT-bundle-JOINER-be.black-STATIVE
He had a black pouch.

Spontaneous material such as this can in fact teach us much of theoretical consequence about the structure of this language that we could easily miss in constructed translations. Here we will consider its relevance to several core areas of grammatical structure.

3. Lexical categories
Mohawk words fall into three formal categories on the basis of their morphological structure: particles, nouns, and verbs. Particles are by definition mono-morphemic, like thé there. Nouns and verbs are morphologically complex. Each is based on a stem accompanied by affixes. The set of stems that serve as the foundation of nouns is
completely distinct from the set that form the basis of verbs. The arrays of affixes that appear with each are also completely distinct. Mohawk nouns, like o-wir-a:' a baby in example (1)a above, contain a prefix specifying the gender of the referent (here o-Neuter) or its possessor, a noun stem (here -wir- baby), and a noun suffix (here -a'). Nouns may also be followed by various enclitics, like the Diminutive -a on baby. Verbs, like wahshakó:ion' he gave it to her in (1)b, have entirely different internal structure. They must contain a pronominal prefix referring to the core arguments of the clause (here -hshako1- he/her) a verb stem (here -on give), and an aspect suffix (here just the glottal stop -' Perfective). They may contain other affixes as well, like the Factual prefix wa- which here implies past tense. Some of the additional affixes that appear in verbs can be seen by glancing at glosses in the passage in example (2). Many Mohawk nouns correspond to English nouns in translation and syntactic function (owirdr ! a baby in (1)a)' and many Mohawk verbs correspond to English verbs (wahshak6:ion' he gave it to her in (1)b). But these two lexical categories differ strikingly in the two languages.

The ratio of nouns to verbs in English speech varies with the topic, genre, participants, etc., as it does in any language, but counts over conversational material comparable to that in (2) often result in noun/verb ratios of around 1/1. The ratios can be quite different in Mohawk. In the constructed examples in (1)a, b, d, and f above the noun/verb ratio is 3/2. This does not demonstrate a difference between English and Mohawk, however. The noun/verb ratio in these Mohawk sentences perfectly mirrors the ratio in their English models. In fact an unusually heavy concentration of nouns is quite typical of constructed examples in linguistics. The 3/2 ratio is characteristic of the genre of sentences constructed in isolation, not of Mohawk. If we compare the noun/verb ratio in the Mohawk conversational material in (2), a different pattern emerges. (This passage was not selected for its noun/verb ratio.) The reader is encouraged at this point to look back at (2) and try a noun/verb count. Mohawk nouns and verbs can be distinguished by their inflectional endings. Nouns contain in a Norm Suffix (NS). Verbs end in an aspectual suffix: Perfective (PRF), Habitual (HAB), or Stative (ST).

The passage in (2) contains just one morphological noun, ohnéhta' pitch, gum, and one borrowed proper name, Konwákeri Margaret [Marguerite]. There are 25 verbs. The noun/verb ratio here is thus at best 2/25. (The free English translation shows perhaps 21/23.) It is strikingly different from that seen in the elicited sentences in (1). Several factors might contribute to the overwhelming exploitation of verbs in Mohawk.

First we can consider the potential effect of pronominal prefixes. As noted, every Mohawk verb contains a prefix referring to the core arguments of the clause, as in tshitewa-na'tónhkhwahkwe' we called him by name. These pronominal prefixes are functionally equivalent to the unstressed pronouns of English, but they also appear when independent nominals are present in the clause. Independent pronouns occur in Mohawk only for contrast or emphasis; none appear in the passage in (2). If one were to count independent pronouns as nouns in tallies over English speech, this difference would significantly affect noun/verb ratios. The 1/1 and 21/23 ratios mentioned above for English did not include pronouns in the noun counts, however, so the difference between English and Mohawk cannot be attributed to the pronominal prefixes.

Running Mohawk speech typically contains fewer independent nouns than English for another reason. In Mohawk, noun stems can be incorporated into verbs, as in ronohkwakéhte' he is bundle carrying. As a result, some meanings conveyed by nouns
in English are not conveyed by separate words at all in Mohawk.

The noun/verb ratio is affected by another factor. Verb counts in Mohawk are
typically higher than in English because there is no adjective category. The information
expressed in English adjectives is typically expressed in Mohawk Static verbs, as in
kahón:tsí it is black. This verb consists of a Neuter pronominal prefix ka-, the verb
root -hon'ts- be black, and a Static suffix -i.

The noun/verb ratio is affected by still another factor. Morphological verbs in
Mohawk can function syntactically in several ways. They can serve as predicates, as in
English: Wahshakotianerónhkwen' ne Kaspé Gaspe scared them. Because the core
arguments are identified pronominally within the verb, verbs can serve as complete
clauses in themselves: Wahshakotianerónhkwen' He scared them. Morphological verbs
can also function syntactically in a third capacity. Without further overt derivation, they
can serve as nominals, as referring expressions. Many are in fact lexicalized as nominals,
like ratenohkwakéhtats he carries a bundle = peddler. The verbal morphology of
Mohawk is so rich that when speakers wish to coin new words, they usually use verbs.

Finally, many ideas that might be expressed with English nouns are simply expressed
as predications in Mohawk, as in ronwatitats they call them = their name (is).

Mohawk shows us that languages may differ not only in their inventories of lexical
categories, but also in the functions of the categories and speakers’ exploitation of them.
It shows us furthermore that categorization based on word-internal form (morphology)
may not match that based on syntactic use. These differences might be missed if our
database consisted entirely of elicited translations of English sentences.

4. Argument structure

It is generally assumed that basic clause structure in all languages consists of a predicate,
one or more core arguments, and possibly additional nominals identifying instruments,
companions, benefactors, locations, sources, goals, etc. (Various terms have been used
for these additional nominals in the literature, among them ‘obliques’ and ‘adjuncts’, but
the terminology is not our concern here.) Mohawk clause structure seems to comply,
particularly in sentences constructed on the basis of English models. Closer examination
of spontaneous speech reveals a surprising difference, however. There are almost no
oblique or adjunct nominals. Participants that might be introduced as obliques or adjuncts
in English are introduced in alternative structures.

In many languages, for example, instruments are typically introduced in oblique
nominals that carry an instrumental case marker or are associated with an adposition like
English with. Mohawk has no instrumental case for nouns and no instrumental adposition
comparable to with. Instruments are typically introduced as the grammatical Patient of
the verb use. To say We make soup with the same kind of corn, for example, a Mohawk
speaker said the equivalent of We use the same kind of corn when we make soup. (This
sentence originated in Mohawk in the course of a longer discussion. The free English
translation was later provided by the speaker herself.)
We make soup with the same kind of corn.

In other languages, companions and associates are often introduced in various comitative constructions, again with a case-marked nominal or an adpositional phrase. In Mohawk they are typically expressed as components of core arguments.

When you talk with a Mohawk, you can say all sorts of things.

In many languages, recipients and beneficiaries are typically introduced by nouns with
oblique case marking or in adpositional phrases like English to X or for X. In Mohawk, recipients and beneficiaries are typically introduced as the core argument of a simplex verb like give or feed, or the core argument of a derived applicative verb. The semantic recipient/beneficiary is a core argument of the verb soup-feed in (7).


*We would give them soup.*

**Aiakhinontará:nonte'**.

*a:-iaxhi-nontar-a-nont-e*

OPTATIVE-EXCLUSIVE.PL/3.PL-soup-JOINER-feed-PERFECTIVE

*we would soup feed them*

The benefactive applicative suffix -'s in the verb in (8) was added to the verb stem -atenhninon sell to derive a new verb stem atenhninon-'s sell-to.


*He sold it to his cousin.*

**Ronara'se':a**

*ron-ar-a'se'=a*

M.PL.AGENT-RECIPIRCA1-RECIPROCAL-be.cousin.to=DIM

*they are cousins to each other*

**wahotenhni:non'se'.**

*wa-ho-atenhninon-'s*

FACTUAL-MSG/MSG-sell-BENEFACTIVE-APPLICATIVE-PRF

*he sold to him*

In other languages, locations, sources, and goals are also typically introduced with nouns that identify a place and carry oblique case marking, or in adpositional phrases, like English at X, in X, from X to X, etc. Places are the one kind of referent that can be introduced as adjunct nominals in Mohawk. Such nominals carry no case marking and no indication of their precise semantic role. Their function is simply interpreted from the semantics of the clause. Even places, however, are often introduced in alternative grammatical structures in Mohawk. They may be evoked by a noun incorporated into the verb. The noun stem -hnek- liquid in (9) indicates the place from which the protagonist came.


*He came back up out of the water.*

**Tonsahahnekótka'we'.**

*t-onsa-ha-hnek-otka'w-e*

DUPLICATIVE-FACTUAL-REPETITIVE-MSG.AGENT-liquid-emerge-PERFECTIVE

he liquid-emerged

They may also be simply an element of the verbal semantics of the verb. The verb root -o-, for example, means be in water. There is no segmentable element meaning water.
The verb root -nhont- means *have in mouth*. There is no segmentable element referring to the mouth. (It does appear that this modern simplex verb root -nhont- may have been formed from a verb root -ont- be attached, which still persists in the language, and an incorporated noun root -nh-, which now appears in larger words involving doors and other openings. At present, however, -nhont- functions as a simplex verb root, even capable of incorporating a new noun, as below.) The speaker cited in (10) was describing some children’s efforts at saddling up an untrained horse.

(10) Mohawk location: verbal semantics: Joe Awenhráthen Deer

_They weren’t able to get the bit into her mouth._

Iáh wi’ tehatikwénie’s
iáh wahi’ te-hati-kweni-e’s
not TAG NEGATIVE-M.PL/AGENT-be.able-HABITUAL
not you see were they able

akonwaristanhónta’.
  a:-konwa-rist-nhont-’a’
  OPTATIVE-3.PL/ZOIC.SG-metal-have.in.mouth-CAUSATIVE-PERFECTIVE
they would metal-insert.in.mouth her

The differences between the Mohawk structures in (4) through (10) above, and their English counterparts with oblique noun phrases, can be described easily in structural terms. The English structures are dependent marking: the roles of the English noun phrases are specified by associated prepositions. The Mohawk structures are head marking: the roles of all participants are specified entirely within the predicate. If we can capture the difference with this generalization, have we exhausted its theoretical interest?

In fact the difference has implications beyond the simple locus of marking. In languages where case markers, adpositions, or constituent order identify roles, we assume full or nearly full productivity. An instrumental case suffix should attach to any semantically appropriate NP. A benefactive preposition should be combinable with any NP that could represent a beneficiary. In head-marking systems like Mohawk, by contrast, productivity is variable. Applicative affixes are derivational. Speakers know which complex applicative verbs exist in the language. They know, for example, that there is a derived verb stem meaning *sweep-for* but not one meaning *shovel-for*, though such a stem could be created. Furthermore, applicative verbs are derived for specific purposes, to name a concept the speaker has in mind, and they may or may not have meanings perfectly equivalent to the sum of their parts. The verb stem -en-hahs- *say-benefactive-*, for example, could mean various things, such as *say to* or *speak for*. It actually means to call someone (uncomplimentary) names. The speaker below used it jokingly as he joined a group and heard himself mentioned.
(11) Semantic transparency vs idiomaticity: applicatives: Joe Dove, speaker

Nahò:ten’ nòn:wa’ enhskwénahse’
naho’ten’ nonhwa’ en-hskw-en-hahs-e’
what now FUTURE-2.PL/1.SG-say-BENEFACTIVE-APPLICATIVE-PRF

What are you going to say about me?

Noun incorporation is also variable. Speakers know which compound verb stems currently exist in the language. They know that the stem string-tie is part of the language but *ribbon-tie is not, though again one could be created and, with sufficient use, become part of the lexicon. Incorporation, too, is done to meet a communicative need. Noun-verb combinations may or may not have transparent semantics. The stem metal-mouth.insert could mean various things, such as putting a spoon into the mouth, but speakers recognize it as the term for putting the bit into a horse’s mouth. Finally, the existence of semantically complex roots is also not predictable. Mohawk happens to have a simplex verb root meaning have.in.mouth but not one meaning have.in.pocket. The grammar is productive, and new forms can be created, but speakers recognize innovation when it occurs.

5. Is language what speakers DO do or what they CAN do?

As noted, Mohawk speakers do not generally introduce instruments, companions, associates, or beneficiaries in oblique or adjunct nominals in spontaneous speech. On rare occasions, however, they have been observed to produce them. The circumstances were telling: the speakers were translating English sentences containing obliques into Mohawk. The rarity of such occurrences should be noted. Even when they are translating, speakers more often choose structures like those in (4) through (11). But the fact that adjuncts are used at all raises an important question. Is language what speakers DO do or is it what they CAN do? Is a language what speakers produce on their own, or is it what they are willing to produce in translations and judge acceptable when asked?

The context of language endangerment presents a special situation. For the most part, languages are endangered at present because their speakers are multilingual, competent in other languages that are replacing the target language in more and more contexts. It is natural for multilingual speakers to tap into their intuitions about their other languages, particularly the contact language, when asked for equivalents to structures in the contact language. Grammaticality judgments and translations may well reflect the structure of the encroaching language rather than the target language.

The implications of language contact have an important place in discussions about the documentation of endangered languages. When describing a language, most linguists would not hesitate to include words and grammatical structures that have come into the language through contact. Few would maintain that a grammar of English should contain no constructions originally modeled on French. If the structures in question are never used spontaneously by speakers, however, we should question whether those structures are really part of the language. Grammaticality judgments about constructed sentences are even more questionable as accurate reflections of the target language. The excerpt below was drawn from one linguist’s discussion of Mohawk. (Again, spelling and analysis have been regularized.)
(12) Speaker judgments

‘Although most or all of the textual examples of animate noun incorporation lack agreement, native speakers sometimes allow variation on this point. Thus, either an agreeing form or a nonagreeing form is accepted in the following.’

Onwá:ri ie(ronwa)-ksten-hser-enhawe’ ne rake’niha.
Onwá:ri ie(ronwa)-ksten-hser-enhawe’ ne rake’niha.
(name) F.sg.Agent(F.sg/M.sg)-be.old-NOMINALIZER-hold-STATIVE the my father
Mary she/(him) is old one holding

(13) Cited gaps

   I visited him because I was hungry the (name)
   I went to visit him because, Sak, I was hungry.

   while (name) I slept she cooked
   While, Mary, I slept, she cooked dinner.

There is no question that these sentences are not good Mohawk. They are not Mohawk for many reasons, some of them the same reasons why their English models are not good representatives of English. A problem with lists of rejected sentences like these is that we cannot be certain which of the reasons constituted the primary motivation for speakers to reject them. An even more serious problem is that such documentation leaves very little record of what Mohawk really is, either for linguists who will later be investigating other issues, or for descendants of speakers wanting to know more about their heritage.

Documentation that concentrates on non-occurring sentences also robs us of the opportunity to understand many structures in terms of the larger contexts in which they function. An important aspect of progress in any science is expanding the sphere of what can be explained. Looking at the spontaneous Mohawk speech in earlier examples, we see that much of what is expressed in separate words in English is expressed in parts of
words in Mohawk. The difference in information packaging could be dismissed as a formal typological detail. Closer study of Mohawk speech shows, however, that many of the meanings conveyed by morphemes inside of larger words can also be expressed by independent words. Speakers have choices. Not long ago, formal alternations in linguistic structure were often relegated to the inexplicable and thus ignored. When we follow the use of alternating constructions through speech, however, we see that it is systematic.

The packaging of ideas often shifts over the course of talk. Important entities and ideas, those the speaker deems worthy of special attention, may be introduced in a separate word that can carry special prosodic emphasis if necessary. Once such an idea has become an established element of the scene, however, it may be carried along simply as part of larger words, either as an incorporated noun or as an affix. Incidental information, not worthy of special attention, may never be expressed as a separate word at all. Such alternation in form can be illustrated by an account of a trip from the speaker’s home community in Quebec to another community in Ontario on September 12, 2001, when border patrols were on high alert. To get from one community to the other, it is necessary to cross the Canada-U.S. border twice, dipping down from Quebec into the United States, then back up into Ontario. The travelers had been worried about the first crossing, where they were to enter the United States, but they were surprised and frightened to find tight security when they crossed back into Canada. At this point the speaker used an independent word shë:kon again to emphasize her surprise at this second police presence. She noted that the police were positioned at the bridge, a noteworthy landmark, using an independent nominal for the bridge.

(14) Dynamic structural options: Watshenni:ne Sawyer, speaker p’c.
(Account of trip to another reserve on September 12, 2001)

a. Karihton.
Police.

Shë:kon karihto:n.    Independent particle again
again police

Police again.

Ahskwákta’  rón:nete’.   Independent nominal bridge
place near bridge they are standing
Standing near the bridge.

Once they had crossed the bridge, the party continued on into the Ontario community, where they spent the day. In the evening they set out again for home, crossing back over the bridge. The repetition of the crossing, not important in itself, was indicated only by the Repetitive prefix t-onsa- tucked into the verb cross. The bridge, now an established entity of peripheral importance, was incorporated.

b. Tonsaiakwahskwi:ia’ke’.
Repetitive prefix again

t-onsa-lakwa-ahskw-f:ia’k-e’.
Incorporated noun bridge
DUPLICATIVE-REPETITIVE: I.EXCL.PL.AGENT-bridge-cross-PREFOUR
We crossed back over the bridge.
The alternation between independent, stressable words, and bound elements of a larger word, is systematic. Such alternations are more than mechanical, however. They are powerful stylistic tools exploited masterfully by speakers. By the end of the conversational excerpt in (2) above, all of the participants had lost their composure completely; they were shrieking with laughter and rolling around in their seats, tears streaming down their faces. They later explained the source of the humor: “It’s the mental image of his black pouch, all through this part of the conversation.” An independent noun for bundle was never used in the conversation, but the noun root -nohkwy- bundle was slipped inside of several verbs: ratnohkwakéhtats he his bundle carries = peddler, ronohkwakéhté he is bundle-carrying, and ronohkwahón:tsi he is bundle black = he has a black bundle. The incorporation serves several functions here. It allows speakers to package what is viewed as a single concept into a single word like peddler and carry a bundle. These complex verbs were not invented by speakers on the spot. They are lexicalized: recognizable Mohawk vocabulary items. The incorporation served an additional function here. The repeated subtle references to the pouch or bundle with incorporated nouns allowed an ongoing string of double-entendres, where the pouch could be interpreted either as the peddler’s bundle or a male body part. The functions of this grammatical structure, so pervasive in Mohawk speech, could never be appreciated if one looked only at elicited translations of English sentences.

6. Structure and substance
Linguistic theory has traditionally focused on grammar over lexicon. General structural principles are typically more highly valued than lists of words. It is becoming increasingly clear, however, that the boundary between structure and substance may not be as clear as was once thought. When one examines corpora for examples of particular grammatical structures, for example, it often turns out that the structure in question occurs over and over but with a surprisingly restricted set of lexical items. This phenomenon may be particularly strong in polysynthetic languages, where so much of grammatical structure is morphological. Polysynthetic languages are disappearing particularly rapidly, however, so discussion of methods for documenting them is especially timely. Whatever the morphological profile of the language in question, if our descriptions of structure are based on elicited translations of invented English sentences, we will miss such dependencies.

The substance itself is not arbitrary. Most human beings who know more than one language recognize the fact that the vocabularies and idiomatic expressions of different languages are never isomorphic. They reflect the array of concepts that speakers have expressed so often over the course of development of the language that they have become conventionalized. They provide a glimpse of what speakers have deemed noteworthy. A record of grammatical constructions that includes the lexical material typically associated with them also provides us with a clearer idea of the exact nature and limits of the grammar. It is easy to assume that a grammatical structure that appears fairly frequently early in the course of fieldwork is fully productive, a possible template into which any lexical material may be inserted. In many cases, however, further work can show that the apparent productivity was illusory. The marker or construction in question may be limited by specific semantic or even lexical factors that will only be clear if we have a
rich body of spontaneous data, rather than translations of sentences of a contact language with lexicon selected by the linguist.

For speakers, vocabulary is usually the most obvious aspect of a language. One needs vocabulary in order to speak at all. Teaching and learning a language is often thought of in terms of words. Structure can be nearly invisible. For speakers and their descendants, a rich record of substance, of vocabulary and fixed expressions, can provide an important record of the culture, of what people have said and how they have chosen to say it. Learning such items can also provide a valuable sense of identity.

As speakers and their descendants delve more deeply into the language, structure can take on increasing importance for them as well. Teachers and learners realize that it is necessary if one is to learn to say more than isolated vocabulary and expressions. It can also be a wonderful source of pride, as it is realized that structure can reveal a different view of the world. It is, as well, a key to oratorical style and interpersonal interaction.

If we are committed to documenting both structure and substance, we need to consider the most effective methods for reaching the goals of both future linguists and communities. Elicitation is clearly a crucial tool for certain vocabulary. In many field situations one could spend considerable time waiting for the words for *dorsal fin* and *yellow-bellied sapsucker*, even if most speakers know them. It is much easier simply to ask. If this method is used exclusively, however, it is easy to end up with a record that is lopsided and does not reflect traditional cognitive categories. The process of elicitation does tend to reap many more nouns than verbs. We might exhaust the terms for body parts, wildflowers, medicines, trees, birds, animals, and diseases, long before those for mental attitudes and conditions. For Mohawk, a record of the nouns would be interesting but far from representative of the language. The elicitation of isolated sentences like the translations from English seen in example (1) tends to leave a record that is particularly impoverished lexically. We will learn relatively little about the tremendous lexical resources of Mohawk if our database consists predominantly of material like *Sak gave Mary a pencil*. When the record includes unscripted speech, we discover surprise after surprise, like *pick off the bugs* and *your ankles are slanting as you walk* seen in the spontaneous sentences in example (1). Much of the most interesting vocabulary could never be elicited because we could never predict its existence.

Similarly, we could wait a long time for speakers to produce certain grammatical forms and structures. One such form might be a Mohawk transitive verb based on a stem beginning in the vowel o, with a first person exclusive dual agent acting on a zoic plural patient (*He and I covered up the cows*). Another might be a relative clause construction in which the shared referent functions as a possessor in the main clause and as an instrument in the dependent clause (*I sharpened the knife’s edge that you cut it with*). Here we should sensitize ourselves to the subtle difference between filling in gaps in known structures (like the shape of the transitive pronominal prefix before o) and creating structure (like the relative clause).

7. Conclusions
The tragedy of language loss will be compounded if we as linguists fail to do all we can to lay a foundation for future discoveries in the discipline and to collaborate with communities to provide the richest record we can of their linguistic heritage. We know already that we will not be able to do enough, but we can share ideas on where to put our
energy and attention at a time when resources are limited.

Because we do not have the luxury of time, we need to sharpen our skills and those of
our students in observation, organization, and analysis. If we are aware of the patterns
that are already known to occur in languages, we can recognize similar patterns in the
language under study on the basis of less data, and recognize surprises more quickly. At
the same time, we will be neglectful if we fail to prepare a foundation for work beyond
what is currently known, beyond current theoretical paradigms. We need to leave more
than lists of currently recognized typological features with check marks. We should
refine our understanding of the sometimes subtle difference between documenting
structure and creating it. We must record as much spontaneous speech as possible in
natural settings and in a variety of genres, registers, topics, etc. We must also foster
sensitivity to the actual versus the potential in our consultants, so that they can specify,
for the record, the difference between what speakers could say and what they do say.

It is also useful to recognize that the needs and goals of communities vary across both
space and time. They will continue to evolve as the world changes for their members. For
some communities and community members, the traditional language is most important,
at present, for a sense of identity. For them, the most valuable linguistic material may be
emblematic phrases that can be used on a daily basis. For others, ceremonial language is
of primary significance, perhaps the political or religious oratory handed down through
generations, or more modern speeches that can be used at intertribal gatherings. For many
communities, a detailed record of language structure and vocabulary is a crucial resource
for language teachers and curriculum planners who are working to keep the language
alive. For some, a record of the full intricacies of the special architecture of the language
can serve as the basis of an appreciation of their intellectual heritage, even if those
learning about the language will never become fluent speakers. A rich record can provide
a picture of traditional styles of interaction, rhetorical artistry, and humor. In most of
these situations, a catalogue of the similarities between the traditional language and all
other languages will not meet the needs of those to come. Descendants of speakers will
not be learning the language so that they can order a meal in a restaurant or ask directions
to the railway station. They will want to know what is special about their heritage.

In the end, documenting spontaneous speech may not yield such random data after all.
It can provide a record of what speakers actually talk about and how they choose to do it.
It can offer clues to what speakers have said over generations. A record of what exists
can also provide a foundation for future discoveries of the principles that shape grammar
and lexicon as well as the culture out of which they come.

Are the two endeavors at odds, documenting a language for future linguists and for
descendants of speakers? Some aspects of the record will inevitably be of more interest to
one group and some to the other. All, however, will be best served if a substantial portion
of the record is open-ended: spontaneous language in a variety of genres. Let us not
compound the tragedy of the loss of linguistic diversity by constraining future discoveries
in our discipline and obscuring rich traditions, all because the poverty of our imagination.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGT</td>
<td>Agent</td>
</tr>
<tr>
<td>CAUS</td>
<td>Causative</td>
</tr>
<tr>
<td>DIM</td>
<td>Diminutive</td>
</tr>
<tr>
<td>DV</td>
<td>Duplicative</td>
</tr>
<tr>
<td>EXCL</td>
<td>Exclusive</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>HAB</td>
<td>Habitual Aspect</td>
</tr>
<tr>
<td>INSTR.APPLIC</td>
<td>Instrumental Applicative</td>
</tr>
<tr>
<td>JR</td>
<td>Stem Joiner</td>
</tr>
<tr>
<td>M</td>
<td>Masculine</td>
</tr>
<tr>
<td>MID</td>
<td>Middle</td>
</tr>
<tr>
<td>N</td>
<td>Neuter</td>
</tr>
<tr>
<td>NMZR</td>
<td>Nominalizer</td>
</tr>
<tr>
<td>PAT</td>
<td>Patient</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PRF</td>
<td>Perfective Aspect</td>
</tr>
<tr>
<td>PRT</td>
<td>Partitive</td>
</tr>
<tr>
<td>Q</td>
<td>Question Particle</td>
</tr>
<tr>
<td>RECIP</td>
<td>Reciprocal</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
<tr>
<td>ST</td>
<td>Stative Aspect</td>
</tr>
<tr>
<td>TLOC</td>
<td>Translocative</td>
</tr>
</tbody>
</table>