Empirical Foundations
for Grammatical Description in the 21st Century

Marianne Mithun
University of California, Santa Barbara
mithun@linguistics.ucsb.edu

1. Traditional expectations
Description of sound system, word formation, sentence structures

Wider audiences
More kinds of linguists with more varied interests
More kinds of community members
  Documenters, teachers, curriculum planners, learners, afficionados
Merging of academic, non-academic communities of scholars

Linguistics as an evolving field
New knowledge brings new questions.
  Typological correlations
  Information structure
  Discourse structure
  Patterns of interaction
  Prosodic structure
  Processes of grammatical development
  Potential effects of language contact

New technologies bring greater capabilities and more questions.
  Quality audio and video recording
  Acoustic analysis
  Transcription software
  Database software
  Corpus software
  Online grammars

Evolving goals for grammars
Provide the foundation for a wider range of users and uses
  and greater understanding of what languages are like
But grammars should be more than a simple typological checklist.
  Capture what is special about the language.

As goals evolve, so too can the kinds of data that might help meet them.

2. Describing basic structures

2.1 Phonetics and phonology
Elicited translations of words: optimal for some tasks
Foundation for basic analysis of sounds, phonological processes
Raw material for acoustic analysis
Assembling clear and abundant examples for the grammar
But words can also be elicited by fields: ‘What kinds of animals do you know?’

Description: A picture can be worth many words
Mohawk contrastive tone

Figure 1: Rising versus falling tone contours in Mohawk

Phrase-medial and phrase-final effects: Tone continues to rise phrase medially.

Figure 2: Mohawk pitch in context
2.2. Morphology

Elicitation
Optimal for some tasks, with good elicitation techniques
- Discovering and exemplifying many categories and distinctions
- Filling out paradigms
- Filling out processes of allomorphy
- Assembling coherent examples
- Creating coherent sets of examples for grammars

Defectiveness
Should not exist
- Inflection is assumed to be fully productive
- Assumed by some to be produced online if regular
- But gaps do exist.
- Difficult to discover in spontaneous speech
  - Speakers simply do not say forms which do not exist.
  - But theoretically important and interesting
  - Tell us about speakers’ knowledge and processes of speech production

Mohawk kinship terms

1. **rak-hsótha**
   - M.SG/1SG-be.grandparent.to=DIMINUTIVE
   - ‘he is grandparent to me’ = ‘my grandfather’

2. **rii-aterè:’=a**
   - 1SG/1SG-have.as.grandchild=DIMINUTIVE
   - ‘I have him as grandchild’ = ‘my grandson’

Large pronominal paradigms: distinctions for humans
- Person: 1 (INCLUSIVE/EXCLUSIVE), 2, 3
- Number: SINGULAR, DUAL, PLURAL
- Gender: MASCULINE, FEMININE.ZOIC, FEMININE.INDEFINITE

(3) Verbs

<table>
<thead>
<tr>
<th>Feminine.Zoic</th>
<th>Feminine.Indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>ke-’kèn:’=a</td>
<td>khe-’kèn:’=a</td>
</tr>
<tr>
<td>1SG/1SG-have.as.younger.sibling=DIM</td>
<td>1SG/1SG-have.as.younger.sibling=DIM</td>
</tr>
<tr>
<td>‘I have her (EZ) as younger sibling’</td>
<td>‘I have her (FI) as younger sibling’</td>
</tr>
<tr>
<td>= ‘my little sister’</td>
<td>= ‘my little sister’</td>
</tr>
</tbody>
</table>

(4) Kinship terms

<table>
<thead>
<tr>
<th>ke-’kèn:’=a</th>
<th>khe-’kèn:’=a</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG/1SG-have.as.younger.sibling=DIM</td>
<td>1SG/1SG-have.as.younger.sibling=DIM</td>
</tr>
<tr>
<td>‘I have her (EZ) as younger sibling’</td>
<td>‘I have her (FI) as younger sibling’</td>
</tr>
<tr>
<td>= ‘my little sister’</td>
<td>= ‘my little sister’</td>
</tr>
</tbody>
</table>

(5) But gaps

<table>
<thead>
<tr>
<th>ak-hsótha</th>
<th>she (EZ) is grandparent to me = ‘my grandmother’</th>
</tr>
</thead>
<tbody>
<tr>
<td>no *ionk-hsótha</td>
<td>she (FI) is grandparent to me</td>
</tr>
<tr>
<td>ake-’nisténha</td>
<td>she (EZ) is mother to me = ‘my mother’</td>
</tr>
<tr>
<td>no *ionke-’nisténha</td>
<td>she (FI) is mother to me</td>
</tr>
</tbody>
</table>

Why?
- Feminine.Zoic was the only category originally.
- Feminine.Indefinite is an innovation.
  - Original Indefinite category ‘one, they’ came to be used as sign of respect.
  - Worked its way gradually through the verb paradigm
  - Still working its way through the kinship terms

So?
- Earliest learned, most often used terms are most resistant to change.
- Speakers really do know words, even inflected ones.

Good elicitation requires
- Typological awareness: Knowing what to probe for
- Sensitivity to
  - phenomena of constructions and lexicalization
  - the difference between filling in and creating structure
  - the difference between the actual and the possible
    - ‘Have you ever heard X?’ for ‘Can you say X?’
    - Learn to notice, cultivate, appreciate, and understand hesitation.

Elicitation can be insufficient alone.
- We might not know enough to ask about the most interesting morphology.
- Some is attached to specific lexical items.

(6) Mohawk ambulative: only with stative verbs containing patient prefixes

<table>
<thead>
<tr>
<th>roniären ‘ton:ne’</th>
<th>roniären ‘ton:ne’</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘his neck has fallen, his head is down’</td>
<td>‘he is sad’</td>
</tr>
<tr>
<td>tahoniaren ‘tön:ne’</td>
<td>tahoniaren ‘tön:ne’</td>
</tr>
<tr>
<td>‘he’s coming with his head hanging down’</td>
<td>‘he’s coming with his head hanging down’</td>
</tr>
</tbody>
</table>
2.3. Syntax

The larger the structural domain, the more important unplanned connected speech.
Patterns with smaller scope tend to be more routinized, less prone to error.

(7) Question/answer pair constructed by good speaker for pedagogical grammar

I:seks kenh ne kanà:taro?
‘Do you eat bread?’

Í:keks tiótkon ne kanà:taro.
‘I always eat bread.’

Phonologically, morphologically, lexically accurate.
But this speaker would not talk like this.

Mohawk word order: focus (newsworth information) first
Mohawk article ne: ‘the aforementioned’

(8) More idiomatic question/answer pair

Kanà:taro kenh i:seks?
bread Q you eat
‘Do you eat bread?’

Hén: tiótkon ne kanà:taro i:keks.
yes always the aforementioned bread I eat
‘Yes, I always eat bread.’

Importance of larger pragmatic context: often best from spontaneous speech

(9) A. Ónhka’ iáh teiakhotá:ton?
who not did one hear
‘Who didn’t agree?’

B. Né:ne ro’niha kí:ken.
it is he is father to him this
‘It was his father.’

Grammar users will look at all examples in the grammar
to learn how the language works.
Every example should be accurate from every point of view
phonological, morphological, lexical, syntactic, pragmatic
Full glossing and translation for all multi-word examples will be appreciated.

Larger context through translations?

(10) Daniel 1:11

Sok ki’ ne Tànier e’thò: niahà:re’ tsi ihrate’ ne shakónonhne
so then the Daniel there he went to he stands the one watches him
‘So Daniel went to the guard

tsi nihorihonte’ ne Ashpenaz ne ahshakoten’nikónhraren
to he matter attached him the Ashpenaz the one would mind him
whom Ashpenaz had placed in charge

ne raónha tãnon’ ne áhsen niah:ti ronten’ro’shòn:’a.
the he himself and the three so they number they are friends to each other
of him and his three friends.’

Perfect translation
All meanings of the original preserved, including style

1 sentence, perhaps 9 referring expressions
Daniel, the one standing, the one watching, the one assigned, Ashpenaz,
the minder, he himself, the three, friends

Mohawk speakers would not say the same things.
Complexity is of different kinds. (cf reference, particles)

(11) Mohawk Cosmology legend, comparable formal oratory
Seth Newhouse, speaker, 1896, Hewitt 1903:265.4-5

Né káti’ né nen sha’or:hen’ne’
that so then that now when it dawned
‘So then, when the next day came,

né: ó:ní’ né: nen sahatikhwén:ta’ne’
that also that then again they finished meal eating
and also when they had finished eating their morning meal,

e’thò:ne’ ne eià:tase’ wà:i’ron’,
at that time the she is new bodied she said
the young woman-being at this time said,

‘Nén ki’ enkahiten:ti’.
now in fact I will leave
‘Now I believe I will start out.’


**Information structure**

Translation can obscure choices speakers make in packaging information.

So should all examples of syntax be spontaneous? Spontaneous examples are often more interesting with more varied vocabulary and idiomatic turns of phrase.

But simple can be good in grammar examples. Spontaneous speech does contain simple constructions as well.

(12) Clause conjoining

\textit{Thò: nionsà:re’},

\begin{itemize}
  \item so he went back there
  \item ‘He went back there again
  \item \textit{sok are’ tahoië:na’},
  \item then again he grabbed him there
  \item he took him,
  \item \textit{ken’ roia’tenhâ:wi},
  \item there he bodily carried him
  \item he carried him,
  \item \textit{átste tahó:ti}
  \item outside there he threw him
  \item and threw him out.’
\end{itemize}

Spontaneous examples can be pruned.

(13) 

\begin{itemize}
  \item \textit{là:ia’k na ‘kahwistà:’eke’}
  \item six so it bell struck
  \item ‘We rode for six hours
  \item \textit{tiio’kehà:ke ionkwahonsi:sere’}
  \item train place we all are container dragged
  \item on the train
  \item \textit{tsi niio’re’ saiákawe’}
  \item to so it is far we all arrived back
  \item Montreal until we got back to Montreal.’
\end{itemize}

(13) is perfectly idiomatic without the last line.

And elicitation can provide just what one needs, if speakers value idiomaticity.

**3. Discourse and interaction**

Opening of story written by Mohawk teachers, excellent first-language speakers

(14) \textit{Tewakhwishenhé:jion}

I am tired

\begin{itemize}
  \item ‘I was tired
  \item \textit{sok iohsnó:re’ onkità:wha’},
  \item so then it is fast
  \item I went to sleep
  \item so I quickly went to bed.
  \item \textit{Sok wa’katà:swahte’}
  \item so then I extinguished
  \item Then I turned off my light
  \item \textit{tanon’ ia’kà:rate’},
  \item and I lay down there
  \item and lay down.’
\end{itemize}

Phonologically, morphologically, lexically, syntactically accurate. Appropriate word order Appropriate division of labor between nouns and verbs

Particles: \textit{sok ‘so then’, tanon’ ‘and’}

Co-constructed narrative in conversation

(15) \textit{Rorihwakwenie:nhs nen’ nè:’e;}

he is matter competent that that

‘He was respectful;

\begin{itemize}
  \item \textit{rorihwakwenie:nhs:ton:ne nek tsi}
  \item he had been matter competent the only that
  \item he used to be respectful but
  \item \textit{khere’ katì’ kenh tshitewana’kôn:nihskwe’ wàhi’}.
  \item I guess in fact Q we used to make him mad TAG
  \item I guess in fact we used to make him mad, didn’t we.’
\end{itemize}

Particles \textit{nen’ nè:’e, nek tsi, khere’, katì’, kenh, wahi’}.

Even texts constructed line-by-line by excellent speakers can be missing many particles with textual and interactional functions.

Issue for discussion: Should writing be the same as spontaneous speech?
Discovering and illustrating function: Tags

Speakers themselves are typically unconscious of larger structures and functions. Patterns may not show up in sentences elicited out of context. Different patterns appear in different genres of speech.

- Informal conversation, banter, serious discussion, procedural instructions, descriptions, anecdotes, reminiscences, narrative, legend, formal oratory, etc

Mohawk wáhi’ ‘isn’t it’ etc.

(16) **Epistemic function**: reduced certainty

A. Ró:ne, i:kehre’ wáhi’
   ‘It was his wife, I think, wasn’t it.’

B. Í:kehre’.
   ‘I think so.’

(17) **Epistemic and interactive**: joint plans

_Tówa’ nón:we nè:e aetewahiatónnion’ ka’ nón: teiotionhontsóhon Kaientarónkwen wáhi’_.

‘Maybe we should write the name Kaientarónkwen where the pieces fit, **don’t you think**?’

(18) **Primarily interactive**: little doubt

A. ... sótsi ka’nikonhráksen’s ótia’ke ki: nihotiiôn:sa wáhi’.
   ‘… because some of the youth these days have bad intentions, **don’t they**.’

B. Mmm.

(19) **Fully interactive**: Co-constructed narrative, with respect for audience

A. Wa’ka’rhé:nien’ne’ se’ wáhi’.
   ‘So it really did topple over, **didn’t it**.’

   ‘That’s right, **isn’t it**. On his head.’

Text structuring

(20) Setting the scene

A. Kí:ken atsa’ktontie’ wáhi’.
   ‘This place along the river, **right**?’

   _Thó: thonathéhtaien’, ónhka’k tsiok takwáh._
   ‘Somebody has a garden there.’

B. Mmm.

(21) Establishing a topic

A [‘There’s another matter I’d like to talk about.]

   _Kí:ken kën:, ioháhonte’ wáhi’._
   ‘This here, road, **you know**.

   _Malone highway shé:kon ratina’ tönkhwa’._
   ‘people still call it the Malone highway,

   _nió:re’ tsı iohwharaká: ronte’ wáhi’._
   ‘it’s outrageous how it’s potted with holes, **isn’t it**.’

B Mmm.

(22) Highlighting important points

A. Iáh tetkaié:ri tsi nítat si kakwatákwen ne--
   ‘It’s not right the way they fixed them,

   _tsı wa’aró:ton wáhi’._
   ‘the way the nets are set up all over, **you know**.’

B. En:.
   ‘Yes.’

(23) Explanation or justification

[‘When a chicken finishes a meal, it puts it in here. After awhile it will keep bobbing its head up and down and down it goes.]

   _Enwatatién:kahse’ wáhi’ naotenà:tshera’._
   ‘It will save its food, **you see**.’
4. Prosody

Traditional grammars
Little attention to intonation beyond the word
But significant aspect of structure for speakers and listeners
Important information for teachers and learners

4.1. Questions

Yes/no questions

![Figure 3: Mohawk and English prosody](image)

Question-word questions

![Figure 4: Mohawk ‘Where are you from?’](image)

Tag questions

![Figure 5: Falling pitch in tag construction](image)

4.2. Complexity

Do all languages have syntactic complexity?
(clauses within clauses)

Mohawk sentence

(24) *Sok iá:ken’ tahatáhsawen’ wahentsiahserón:ni’*. so one says he began (it) he fish cleaned
‘So then he apparently started [cleaning the fish].’

English translation is a complex sentence
Mohawk literal glossing looks like a string of simple sentences.

![Figure 6: Mohawk complex sentence](image)

(25) Contrast: two independent sentences

so one says it started it current picked up this water place
‘So then, they say, it started. The water swirled around.’

![Figure 7: Mohawk sequence of two sentences](image)
4.3. Information structure

(26) Mohawk complex sentence

Né: ki’ a: wà:kehre’, enkewanòhetste’ ki’ ne ronónha’ aotirihwa’shön:’a. it is actually I wanted I will word pass in.ft the it is theirs their various words ‘It’s just that, ah, I wanted [to pass along their message].’

The clause ‘to pass on their message’ is syntactically subordinate.

But it is prosodically prominent.
It carries the main information.

Elicitation or spontaneous speech?
Accurate prosodic patterns not produced reliably on demand, out of context.
Consider challenges facing actors.

5. The dynamic side of language

Now more than ever, we recognize that languages are constantly evolving.

Speakers are working to
make sense out of perceived patterns
repair apparent exceptions
extend existing structures to new contexts for new purposes
routinize frequent constructions
reinforcing faded constructions for greater force

Constructions often begin in specific lexical contexts with narrow meanings
then are extended to more lexical items and more contexts
acquire more with more general meanings
increase in frequency
major lexical items lose categoriality
ultimately fuse phonologically and lose substance

Language change typically involves variation.
Allophones may become distinctive sounds.
A new construction may compete with an older one, then ultimately win out.
Earlier informal registers may become standard.

Elicitation alone can miss this dynamism.
Pairing of substance and structure (words and grammatical patterns)
is set by the interviewer rather than by the speakers.
Gradual progress of a construction through the lexicon difficult to spot.

Variability may not be documented.
There may be no record of informal speech.
Speakers strive to produce ‘proper’ language.
No documentation of reduced forms
No record of transitional stages from major lexical items (full verbs, nouns)
to minor items and grammatical and discourse markers
Elicited examples are typically poor in such particles.

Mohawk íā:ken’

Verb íā:ken’
  iak-en’
  INDEFINITE-AGENT-say-STATIVE
  ‘one says’

Can be matrix verb.

Much more often hearsay evidential.
Example (24): integrated syntactically into the sentence: after sok ‘so then’
Phonological reduction: íā:ken’ > iaken’

(24) Sok íā:ken’ tahatàhsawen’ wahentsiahserón:ni’.
  so one says he began (it) he fish cleaned
‘So then, they say, he started [cleaning the fish].’

Prominent discourse role in structuring narrative.
Rarely appears under sentence-by-sentence elicitation.
6. Language contact

Traditional grammarians often took pains to include only native material.
But contact can play a major role in shaping language.

We now want to know how contact can affect a language
what features can be transferred
in what sequence
under what linguistic and social circumstances

Lexicon: borrowing may be easy to spot
Spanish in Latin America, Portuguese in Brazil, Russian in Siberia, etc.

Structure: not so obvious
Replica grammaticalization: Heine and Kuteva 2006 etc.
Bilinguals replicate a structure from one language in the other
using only native lexical items

From minor to major pattern
Frequency of existing pattern may increase
on the model of its counterpart in another language.

Frequency can have consequences.
Increased use of passives can result in ergative systems
Increased use of antipassives can result in accusative systems.
Increased use of lexical constructions can speed grammaticalization.

Elicited translations
Speakers may consciously avoid borrowed words or phrases.
But speakers typically strive to replicate target sentences as closely as possible.
How do we know whether a structure is actually now part of the language?
How do we calibrate frequencies?

7. Conclusion

As linguistic theory and technology progress
so can our ideas about contents of a good grammar
and the data necessary for analysis and exemplification

No grammarian can do it all, but it can be useful to be aware of possibilities.