The trees and the roots: Syntax in the larger ecology of language work

28th International Lexical Functional Grammar Conference

George Aaron Broadwell, University of Florida

broadwell@ufl.edu

July 2023
Doing syntax with an endangered language requires 'thinking big' about

- The role of syntactic documentation in larger language documentation projects
- How syntacticians can best interact with other linguists (phonologists, sociolinguists, etc.)
- How linguists can interact with Native language workers
- How linguists can interact with scholars from other disciplines.
Linguists and language workers

- Academic linguists and language workers and other professionals who work with language:
  - Native language teachers
  - Legal and medical translators
  - Tribal language staff

- As well as academics whose work relies crucially on careful language analysis:
  - Literary analysts, historians, anthropologists, librarians, digital humanists
My central metaphor:

To study a tree species in depth, it's not enough to focus only on the visible tree.

Much that is essential to the tree lies in the roots and their connections with the environment around them.
Necessity of corpora

To properly document a language for future generations of scholars and descendents, it is not enough to focus just on a list of elicited sentences. Much that is important about the grammar of the language needs to be grounded in a corpus of material, especially a naturalistic corpus.
Practical and theoretical uses of corpora

Language documentation and language workers need natural corpora for

- Example sentences in dictionaries
- Demonstrations of word senses
- Curriculum development
- Vocabulary frequency data
The corpus problem for small languages

- But a critical problem for smaller languages is the lack of any corpora.
- I have personally been convinced over the last fifteen years that frequency plays a crucial role in syntax, language learning, and historical change.
- But without a corpus, we do not know anything about word or construction frequency.
Endangered language issues

- For many Native American groups, there is a confluence of issues:
  - Diminishing numbers of speakers
  - Low literacy levels among speakers
  - Low status for the Native language
  - Pressure to assimilate to the larger state and its language (e.g. English, Spanish)
Endangered languages and history

● The majority of contemporary language documentation is based on
  ○ Elicitation/interviews with current speakers
  ○ Examination of prior documentation (e.g. older dictionaries, grammars, text collections)
● But this approach is largely ahistorical. It does not systematically engage with older texts.
Available historical sources

- In some Native communities, however, there are abundant sources of material written in the local language.
- Historical documents can strengthen documentation and be useful to speech communities.
- Today's examples:
  - Zapotec language material from the Colonial era (ca. 1521-1821)
  - Choctaw language material from the Indian Territory period (ca. 1835-1906)
Users of historical material

- Linguists (natural corpus, data for diachrony)
- Historians (Mexican, Oklahoma, and indigenous history, property transfers, legal system, race and citizenship)
- Native language workers
- Zapotec and Choctaw people (genealogical, language study, history)
Process in historical documentation

- Locate source material
  - Printed material (often in rare book sections of libraries)
  - Handwritten materials (often in archives)
  - Make digital copies of these when not available
- Transcription of material for study
- Interlinear glossing; lexical analysis
Zapotectec
Oaxaca

from Munro et al 2007
Zapotec languages

- 400,000 - 450,000 speakers of all Zapotec languages
- Communities are shifting to Spanish in more and more contexts
- All Zapotec varieties are endangered
- Zapotec has no de facto official status
- Most speakers of Zapotec do not write their language
- Discrimination against speakers of Zapotec (and other indigenous languages) has a long history and continues today
Zapotec writing
Zapotec alphabetic writing
Meta-linguistic texts
Religious texts
Example -- Feria Doctrina 1567
Legal texts

Wills / testaments
Bills of sale
Archives
Initial stages of our project

- Brook Lillehaugen and I began to collaborate on a corpus of Colonial Valley Zapotec texts around 2010. We jointly and separately published on CVZ morphosyntax along with students.
- For several years, this was primarily a scholarly database shared between us.
Adding material to the corpus

- Both Brook and I train undergraduates to transcribe documents.
- Handwritten material requires special training in paleography; Brook has been training students to read old handwriting.
- Our current corpus is about 165,000 words
Making a digital humanities project

● In about 2012, we began efforts to make our database available over the web. We named the project Ticha, which means 'language' or 'word'.
● Much of the programming and web infrastructure was provided by librarians at Haverford College. We launched the site in 2013.
● We also began to assemble an advisory board of Zapotec people for advice on how to make the site useful and useable.
Our initial interest in the texts was mostly scholarly. We were focused on Zapotec morphosyntax. Initiation of a DH project made us begin to think much more carefully about how our these documents could be useful to the community and to other scholars. In particular, we found that we needed to work with Zapotec communities and scholars outside linguistics.
Welcome to Ticha

A collection of Colonial Zapotec resources including transcribed texts, translations, and linguistic analyses.

Arte en Lengua Zapoteca
Examine Friar Juan de Cordova's 16th century grammar of Colonial Valley Zapotec.

Handwritten Manuscripts
Examine manuscripts in Colonial Valley Zapotec.

Bibliography
Look for related works on Colonial Valley Zapotec and its cultural context.
TITLE
Testament from San Sebastián Teitipac, 1744

PERCENT NEEDS REVIEW
20

SCRIBE
Manuel Valencia

DOCUMENT TYPE
Testament

ARCHIVE
Archivo Histórico de Tlacolula de Matamoros Oaxaca, Mexico
<table>
<thead>
<tr>
<th>Word</th>
<th>1-1</th>
<th>Laa</th>
<th>stissima</th>
<th>trinidad</th>
<th>Dios</th>
<th>bixoce</th>
<th>Dios</th>
<th>xini</th>
<th>Dios</th>
<th>espiritu</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Lex. Gloss:**
- Laa: name
- Dios: God
- bixoce: father
- xini: child
- espiritu: spirit

**Morphemes:**
- Laa
- Dios
- bixoce
- Dios
- xini
- Dios
- Espiritu

**Free:** In the name of the most holy Trinity, God the Father, God the Son, God the Holy Spirit. Three distinct persons and only one true God.

**En:** En el nombre de la Santissima Trinidad, Dios Padre: Dios Hijo: Dios Espiritu santo: tres Personas distintas y un solo Dios verdadero.

**Lit.** [In] the name of the most holy trinity God the father, God the child and God the Holy Spirit who are three persons but only one true God.
Laa stissima trinidad Dios bixoce Dios xini Dios espiritu

Santo ninacani chonna persona tobizica Dios nali Chela
Laa nayonna xiteni xonna xi too  Santa Maria xinani Amen
quienni quirra benni quecheri xiti chiaya sicanaa Manuel
da Cruz nacaya benn[i] huala chi tuari quechee Sans
bastian nagabaya barrio quieze rizacaya quelagui
chia xipella latiya laquelacani anna ricaviruya rinni
ya romniya goona Loo quira becogo medioga quizaa becogo
Chela rinnia cocachi leechle laya xipella latia missa vigilla
Chi Cue Dios xiquela na bannia ---
Chela rinnia yooohlichiya huane solar xitenia quea
Nane xini chapaya Manuela Agustina dela Cruz
Chela rinnia napaya tobi cue layoo teche taani quiebichi na
Chaga vizua Miguel ernandes nezaa sosilla sochi nachaga
Vizaas layoo Cumun Layoo ni betoo ohinelenia chinia biehanna
1-1. Laa stissima trinidad Dios bixoce Dios xini Dios espiritū
1
lāa s[an]tissima trinidad Dios bixoce Dios xini Dios espiritū
name most holy trinity God father God child God spirit name most holy trinity God father God child God spirit
[in] the name of the most holy trinity God the father, God the child, God the holy spirit

1-2. Santo ni nacani chonna persona tobizica Dios nali Chela
santo ni n-aca=nì chonna persona tobi=zi=ca Dios nali
holy REL STA-be=3s 3 person 1=only=EMPH God there
who are three persons but only one true God. And

1-3. Laa nayonna xiteni xonna xi too Santa Mariă xinani Amen
lāa na-yonna xiteni xonnaxi toŏ Santa Mariă xi-na=nì
name STA-sacred of Lady great Santa Mariă POSS-mother
[in] the sacred name of the Great Lady Saint Mary, his mother, Amen
But are texts on a website really accessible?

- Accessibility is more than making digital surrogates of texts available.
- Accessibility must include community guidance, capacity building, and community engagement.
- 2019-2021 ACLS Digital Extension Grant to create an open access educational resource on Colonial Valley Zapotec developed with Zapotec partners and in collaboration with the broader Zapotec community.
Caseidyneen Saen – Learning Together

Colonial Valley Zapotec Teaching Materials

Xóchitl Flores-Marcial, Moisés García Guzmán, Felipe H. Lopez, George Aaron Broadwell, Alejandra Dubcovsky, May Helena Plumb, Mike Zarafonetis, and Brook Danielle Lillehaugen

All Rights Reserved
Chapters

1. Ticha: an Introduction
2. Colonial Documents and Archives
3. Reading an Interlinear Analysis
4. Numbers
5. Language Shift
6. Twitter and Zapotec Language Activism
7. The Written and Intellectual Legacy of the Zapotecs
8. Reclaiming our Languages
The **Conversatorios 2019-2021**

- Original planned for collaboration with the broader Zapotec community through in person workshops in California and Oaxaca.
- These shifted to online *Conversatorios*, where a draft of each chapter of *Caseidyneën Saën* was work shopped and later revised based on feedback and additional contributions.
- The resulting online, open access resource is available in English and Spanish and developed with high school and undergraduate audiences in mind.
Open Source Digital Content for and by Diverse Audiences
Workshops on using Ticha in Mexico
Workshops on using Ticha in Mexico
Going beyond linguist expertise

- Understanding these documents requires various kinds of expertise that linguists usually do not have.
- For that reason, we need to consult with experts in colonial history, religious history, etc.
- These colleagues can often help us understand the content of the documents.
Community consultation

● Our workshops showed us that our website was not initially as easy to use as we hoped.
● Community members also had multiple requests that we tried to implement at Ticha.
● Thus the site became more collaborative over time.
Lessons learned

- Some of our scholarly goals (e.g. an analytic dictionary of CVZ; more analysis of morphosyntax) have been put on the back burner while we work on this project.
- We found much more interest among community members in historic documents than we had imagined.
- In general, the set of community members interested in language (for artistic, historic, education, cultural) reasons is much larger than the number interested in linguistics.
- Our current project is much more oriented toward documentation and promotion of written Zapotec from earlier centuries.
Lessons learned, part 2

- We found that many students and community members were eager to volunteer time for the project. Over the last few years, around a dozen students have worked as RAs on the project and 25+ people volunteered to help transcribe.
- Librarians have been an incredibly valuable part of our team. Their expertise in web design and digital humanities has allowed us to do an enormous amount.
- Students with web and/or programming skills worked on the back end to keep the site working.
Work still to be done

- Many of the texts at Ticha do not have an available interlinear analysis. We've found that it takes a fair amount of work to get these on the website, and we'd like to automate more of the process.
- We are not as well integrated with Mexican academics as we would like to be.
<table>
<thead>
<tr>
<th>Some linguistic topics of investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>● TAM morphology</td>
</tr>
<tr>
<td>● Pronominal systems</td>
</tr>
<tr>
<td>● Resumptive pronouns</td>
</tr>
<tr>
<td>● Locational verbs</td>
</tr>
<tr>
<td>● Valence alternation</td>
</tr>
</tbody>
</table>

| ● Negation                              |
| ● Toponyms                              |
| ● Word order and topicalization         |
| ● Conjunctions                          |
Some historical and anthropological issues

- Land boundaries
- Inheritance practices
- Linguistic deference
- *Guelaguetza* (reciprocal labor debt)

- Gender and property
- Kinship and marriage
Reflections

- Work on this project also led us to reflect on the different levels of privilege and power in our group.
- American professors have considerable power and access to resources that Zapotec people often do not have. We've tried to find ways to use these resources to the benefit of the community.
Choctaw
Choctaw Nation of Oklahoma today: 223,279 citizens
Mississippi Choctaws today, 11,000 Citizens
Choctaw history

- The Choctaw people were the largest Native group in Mississippi.
- They were missionized in the 1820s, and Choctaw converts were educated in Choctaw and English beginning in that decade.
- The Indian Removal Acts (1830) forced the majority of the tribe to relocate to Indian Territory (now Oklahoma)
Choctaws in Indian Territory

- After removal, the Choctaw Nation operated a national government with a school system, law enforcement, courts, etc.
- Much of the operation of the Choctaw Nation was recorded in written Choctaw.
- Many literate Choctaws wrote letters during this period and newspapers were published partly or completely in Choctaw.
Examples of printed material
A standing army of slaves.

Not a few years older, he has just a little more than a month old, when he was put in his cradle, and was left to the care of his parents. He was brought up in a little village near the sea, where he learned to know the world and to understand the language of the people.

His mother, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His father, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His mother, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His father, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His mother, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His father, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His mother, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.

His father, who was a slave, had him taken care of by a neighbor, who taught him to read and write. He soon learned to speak the language of the people and to understand the world.
Ans| Keyow Lewis Pstukcha ak illa tok.

Que| Joseph uto ona mut issuba akkowa cha pih onah?

Ans| Ah yummuto mishihma ho chukka ut talaya tuk u yummako ant hili chi cha ona tuk achine chamo?

Ques| Akka aya hosh?

Ans| Ah akka aya hosh. Lewis ak osh oka hume sapita ho ishkoli hosh sahaksi tuk achine achit am anoli tok Joseph ut.

Ques| Afumme hulhti naa ash attoko.
| 18 Word | Que | Joseph uto | ona | mut | morphemes | ymmako | ymmak -q | that | acc | come and stand:pl | causative | srs | arrive | past | to seem | so | past |
| 19 Word | Ans | Ah yummuto mishlima ho chikka ur talaya nuk y | yummako | hili | chi | cha | ona | tuk | achine | chamo | ? | Free: When Joseph, reached the place he got off of his horse and came to the meeting? 
Lit: When Joseph arrived, did he dismount his horse and come [on foot]? 
Free: Yes, There is a house at some distance from the Church. He left his horse there.
The Oklahoma Historical Society has 483 volumes of material related to the Choctaw Nation -- more than half is probably in the Choctaw language.
Opportunities and beginnings

- As we document the Choctaw language spoken today we are compiling a corpus of spoken Choctaw, transcribed from interviews.
- But we think Choctaw language research could benefit from a similarly "thick" and historically rich corpus.
A new cross-disciplinary research group

- Using Ticha as a model, we formed the *Choctaw Language and History Group* in 2020, composed of four linguists and five historians. Our group includes three Native historians.
- We've transcribed 625,000 words of 19th century Choctaw text.
- Each week our group meets to discuss one of the case, focusing on both the linguistic and historical problems raised by the documents.
Early stages and eventual public resources

● We're now in the 'scholarly research' stage as we try to understand the kinds of documents present in these archives.
● Our next steps are to work on a way to share these results with interested people.
● A point of sensitivity and discussion for us --
  ○ Because these are legal texts, some of them discuss crimes committed in the Choctaw Nation, probably by people who have living descendents.
Because we formed as an interdisciplinary group, we have tried to be sensitive to both linguistic and non-linguistic issues from the start.

We are nevertheless a group of scholars, and our next steps involve connecting our work with the Choctaw community.
Ticha as a model

- We've learned a lot about how to work with historical material in indigenous language through Ticha.
- We are attempting to apply this model in other cases. A key component in our current understanding is the inclusion of Native people and historians in the creation and curation of resources.
Conclusions -- Trees and roots

- Training in morphosyntactic analysis is an excellent start for a linguist.
- But today's realities require significant collaboration with others. Those collaborations might include -- lexicography, paleography, digital humanities, history, corpus linguistics, and language teaching.
- Our field does best when we think widely about how our interests fit into the larger scheme of scholarly and community language work.
Ticha team:

George Aaron Broadwell  
University of Florida

Janet Chávez Santiago  
Teotitlán del Valle

Laura Curiel  
California State University, Northridge, San Lucas Quiavíní & Oaxacalifornia

Xóchitl M. Flores-Marcial  
California State University, Tlacolula de Matamoros & Oaxacalifornia

Moisés García Guzmán  
San Jerónimo Tlacochoahuaya
Ticha team:

Rogelio Hernández Sernas
Santiago Matatlán

Eloise Kadlecek
Bryn Mawr College ‘22

Collin Kawan-Hemler
Haverford College ‘22

Brook Danielle Lillehaugen
Haverford College

Felipe H. Lopez
Haverford College Libraries & San Lucas Quiaviní
Ticha team:

Edith Matías
Diaz Ordaz

Yaneth Molina
San Blas Atempa

May Helena Plumb
University of Texas at Austin

Ignacio Santiago-Marcial
San Felipe Güilá

Maria Velasco-Vasquez
California State University, Northridge, San Francisco Yatee & Oaxacalifornia
Ticha team:

Michael Zarafoonetis
Haverford Libraries

Alejandra Dubcovsky
University of California, Riverside
Awards for Ticha

2023  Canadian Social Knowledge Institute
      Honorable Mention – Open Scholarship Award

2021  Hispanic Digital Humanities Award – Best Teaching Project

2021  Latin American Studies Association Section Award – Best Digital Humanities Project

2021  Library Company of Philadelphia Innovation Award – Honorable Mention
Choctaw language and history team:

George Aaron Broadwell
University of Florida

Christina Snyder
Penn State University

Julie Reed
Penn State University

Edward Green
Penn State University

Matthew Tyler
Cambridge University
Choctaw language and history team:

Michael Stoop
University of Florida

Seth Katenkamp
University of Florida

Frankie Hiloha Bauer
University of North Carolina

Jamie Henson
Penn State University
Topics of investigation for Choctaw

● Tense
● Evidentiality and quotation
● Status of freedmen in Indian Territory
● Gender and inheritance
● ’Whooping' as an expression of indigeneity and its suppression in the 19th century
● Adoption and legitimization of children
Xtyoztën yuad - Yakoke -Thank you - Gracias
Publications and talks from Ticha


Publications from Ticha, continued


