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



American Council of Learned Societies

Syntax in a larger context

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 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

Trees and roots

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

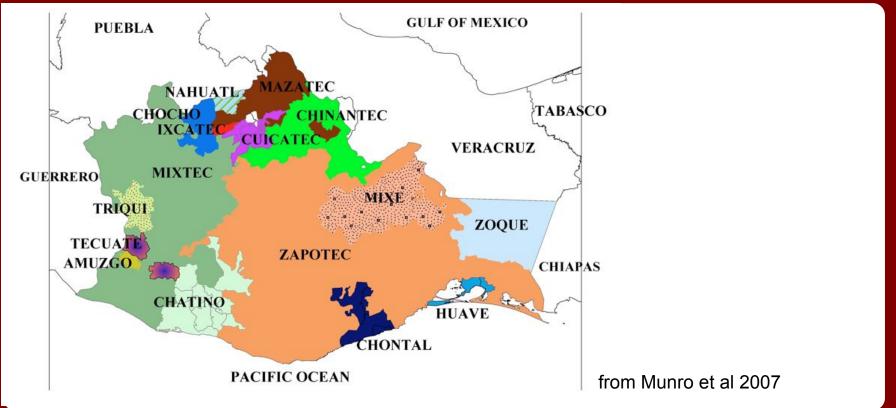
- Locate source material
 - Printed material (often in rare book sections of libraries)
 - Handwritten materials (often in archives)
 - Make digital copies of these when not available
- Transcription of material for study
- Interlinear glossing; lexical analysis

Zapotec

Oaxaca



Oaxaca



Zapotec languages

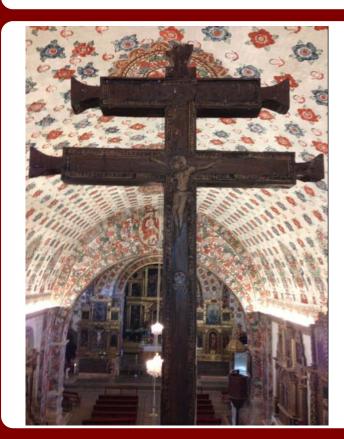
- 400,000 450,000 speakers of all Zapotec languages
- Communities are shifting to Spanish in more and more contexts
- All Zapotec varieties are endangered
- Zapotec has no de facto official status
- Most speakers of Zapotec do not write their language
- Discrimination against speakers of Zapotec (and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

Zapotec writing



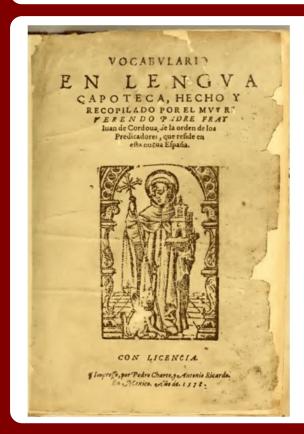


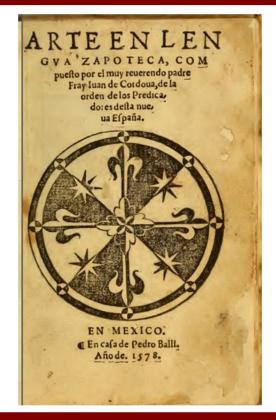
Zapotec alphabetic writing

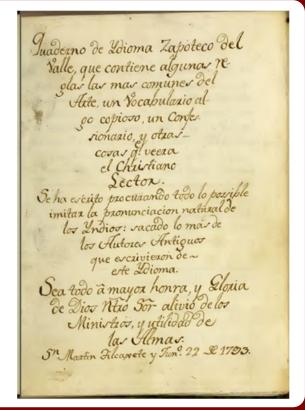




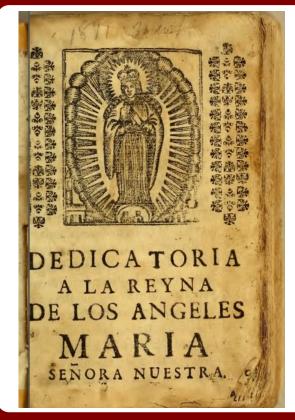
Meta-linguistic texts

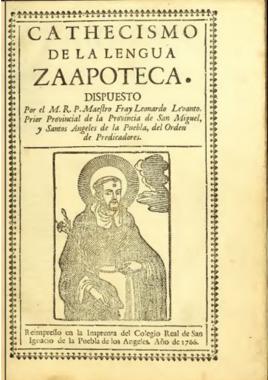


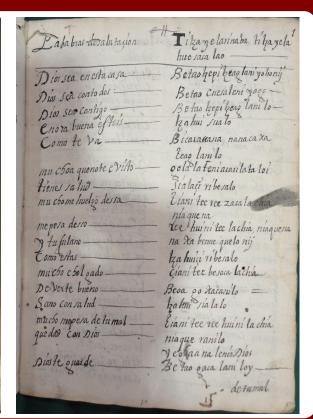




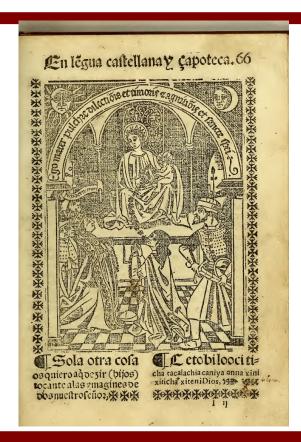
Religious texts







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.

Trees and roots -- Zapotec

- 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.

Ticha

http://ticha.haverford.edu





Welcome to Ticha

 $\label{lem:condition} A \ collection \ of \ Colonial \ Zapotec \ resources \ including \ transcribed \ texts, \ translations, \ and \ linguistic \ analyses.$

Arte en Lengua Zapoteca

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

Explore

Handwritten Manuscripts

Examine manuscripts in Colonial Valley Zapotec.

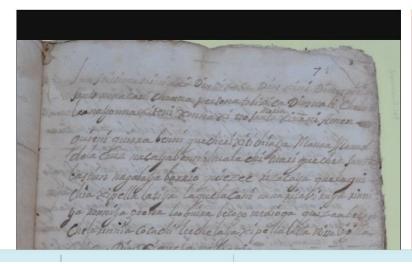
Explore

Bibliography

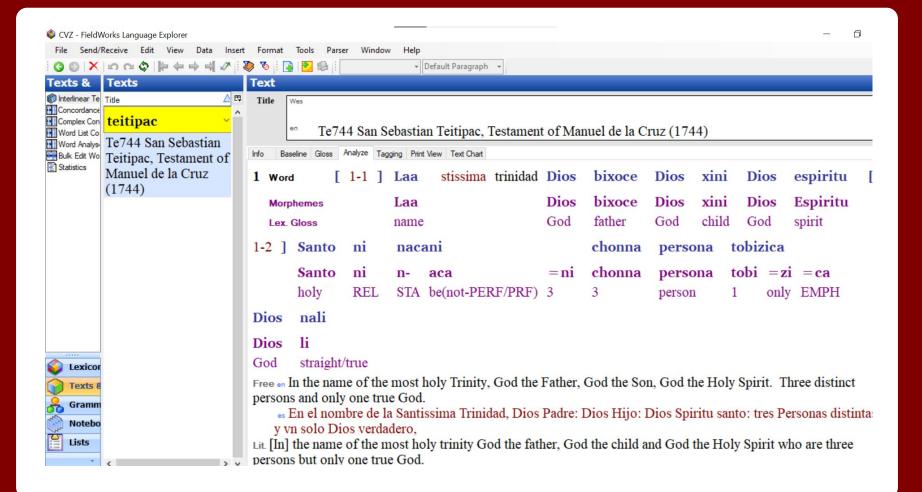
Look for related works on Colonial Valley Zapotec and its cultural context.

Explore















Transcription English translation Analysis

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
quieni quirra benni quecheri xiti chiaya sicanaa Manuel
dela Cruz nacaya benn[i] huala chi tuari guechee Sans
bastian nagabaya barrio quiezee rizacaya quelagui
chia xipella latiya laquelacani anna ricaviruya rinni
ya ronniya goona Loo quira becogo medioga quizaa becogo
Chela rinnia cocachi leeche laya xipella latia missa vigilla
Chi Cue Dios xiquela na baania --Chela rinniya yooholichiya huane solar xitenia quea
Nane xini chapaya Manuela Agustina dela Cruz
Chela rinnia napaya tobi cue layoo teche taani quiebichi na
Chaga vizaa Miguel ernandes nezaa sosilla sochi nachaga







Transcription English translation Analysis

- 1-1. Laa stissima trinidad Dios bixoce Dios xini Dios espiritu¹
 laa s[an]tissima trinidad Dios bixoce Dios xini Dios es
 name most.holy trinity God father God child God sp
 [In] the name of the most holy trinity God the father, God the child
- 1-2. Santo ni nacani chonna persona tobizica Dios nali Chela santo ni n-aca=ni chonna persona tobi=zi=ca Dios nali chonna persona tobi=zi=
- 1-3. Laa nayonna xiteni xonna xi too Santa Maria⁴ xinani Amen laa na-yonna xiteni xonnaxi too⁵ Santa Maria xi-na=ni name STA-sacred of Lady great Santa Maria POSS-mot [in] the sacred name of the Great Lady Saint Mary, his mother, Am

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

 2010-2021 ACLS Digital Extension Grant to create an
 - 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



Caseidyneën Saën - 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

READ BOOK



Chapters

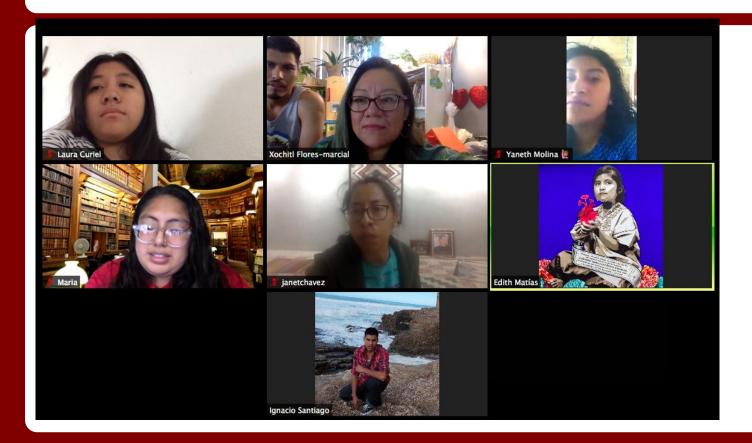
- 1. Ticha: an Introduction
- 2. Colonial Documents and Archives
- 3. Reading an Interlinear Analysis
- 4. Numbers
- 5. Language Shift
- 6. Twitter and Zapotec Language Activism
- 7. The Written and Intellectual Legacy of the Zapotecs
- 8. Reclaiming our Languages

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 <u>feedback and additional contributions</u>. 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



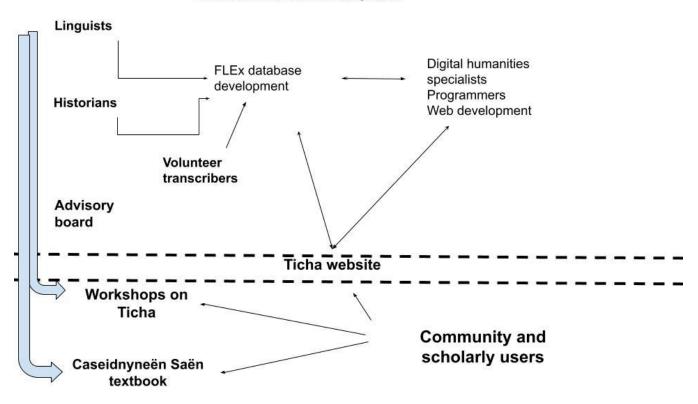
Going beyond linguist expertise

- 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.

Behind the scenes development



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.

Some linguistic topics of investigation

- TAM morphology
- Pronominal systems
- Resumptive pronouns
- Locational verbs
- Valence alternation

- Negation
- Toponyms
- Word order and topicalization
- Conjunctions

Some historical and anthropological issues

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

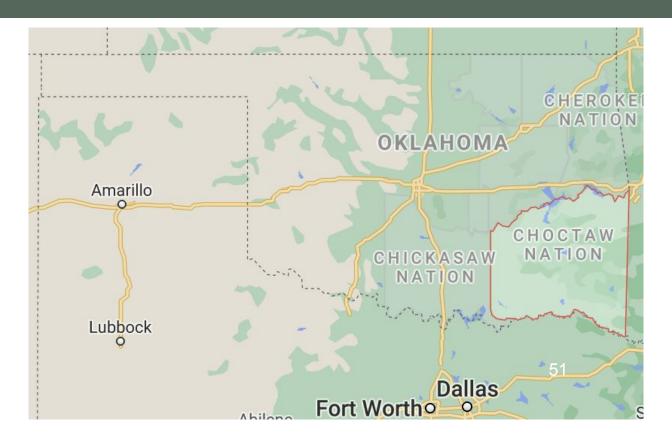
- Gender and property
- Kinship and marriage

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



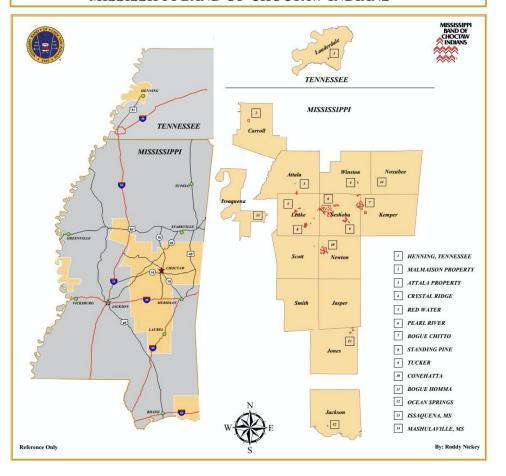


Choctaw Nation of Oklahoma today: 223,279 citizens

Indian Territory Boundaries, 1889

(Pre-Organic Act) Peoria Quapaw Modoc Wyandotte Shawnee Kaw-No Man's Land Cherokee Osage Attached to the Twin Territories Seneca Cherokee Outlet (Nez Perce) via Organic Act of 1890 Opened by Land Run **Nation** Otoe-Missouria-September 16, 1893 Unassigned Lands Cheyenne and Opened by the Creek First Land Run Iowa Arapaho April 22, 1889 and **Nation** Opened by Land Run After Allotment April 19, 1892 Texas Caddo Lands Opened by Lottery After Allotment July 9, 1901-August 6, 190 Potawatomi and Shawnee Kiowa, Comanche, and Apache Choctaw **Greer County** Lands Opened by Lottery Claimed by Texas Chickasaw After Allotment until 1896 July 9, 1901-August 6, 1901 **Nation** Nation **Grazing Lands** Known later as "Big Pasture" Lands opened by Sealed Bid December 1906

MISSISSIPPI BAND OF CHOCTAW INDIANS



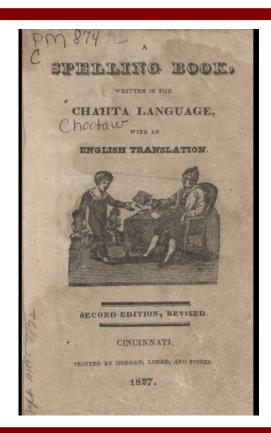
Choctaw history

- 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



ULLA I KATIKISMA: CHILD'S CATECHISM IN CHOCTAW: BEING A TRANSLATION OF DR. WATTS' SECOND CATECHISM FOR CHILDREN. Revised. BOSTON: PRINTED FOR THE AMERICAN BOARD OF COMMISSIONERS FOR FOREIGN MISSIONS, BY CROCKER & BREWSTER.

Choctaw Intelligencer 1851



WEDNESDAY, OCT. 15, 1851

The Choctaw Intelligencer .- Nan huoli holisso ilupput wensta aiyuka kak o toba hoke.

tvli-holisso tuklo hoke; tikbanli atobbi

Holisso-atoba iluppak o nholissa chi mikmut iluppak o kuna hosh han im akaniohmi hokmut, Alsobrook micha Dwight itatuklo kak o i holisso chashke; mikma holisso aya ka pishno akke; mikma noussu ya ng p indi hoshil atobba hinla hoke.

Nanta holi o harak moma ket a chulona ili periodici di

Hatuk maa kalent Johna loh kia akestininchi i ponna, mikmet isht aisht aya, keyukmut nana ho kanichmi puta kuno ngksika maya kut nan im ithina achukmut, isht anumpulit, i hopoyuksa i ponnu: itunowa chatuk: yohmi kia hatak moma kut ilap nan im akaniohmi ik aiono, iluppak a nahukmalie akassininahi ili ishaw

A standing driny of Bibles.

Nitak tikba yatuk pilla ka pit a nukfillit, nana akaniohmi yatuk puta ka ephinsa ka okla il aiyuka gehut mihinti kut hatak tunup, tushka chipota nan isht il otvininchi im asha bicka tuk ahoba chatuk; okla kunimi kut holisso holitopa yimmi, ubanumpuli okla ha i tvnvpi kak osh hatak tunup ikbit alasha beka tok okla kanima ho hatak tunup, yakohmi kut laua, miko kia yakolimi kak bano ho anukchieto hosh ahanta hokholitopa Baibil ut iksho, keynkmut, otvni keyn kanimpo pulla beka tok; Afommi achofakma vlhtoba kvt vba Chitokaka yvt okla kvnima ho ai ititakla ka ilap im apehlichika ik ai oteninchekma, nan ai vihpeso im antia okla nan im annkfilla bieka beka tok oke. Amba yakni kunima ka vba Chitokaka yut ilap im apehlichika alieliehikms, Baibil, et yemhosh ningha kut na m half nukshopa keyn, hatak puta kut holitobli, yakoh felma chatuk keyn hoke. Himonasi yak kin yakohmi ka il ithvijashke; nan ik vihpesot abanta hosh nana ka nuktahla pulla hoke.

chatuk oke; volimicha hatak tekeh boli et hatak bika ya it im anumpli keyu, nashoka ya pisa keyu hosh it im anumpuli chatuk oke; si akostininchi tuk a ahni hokut yohmicha na na moma immaiya Chihowa, i nan vllipisa yash o kobuffi tuk okut, micha nitak ishtaiopik muno, il anohonin he mak okut.

Hatak kanimi kut tekchi ya i noko wakma, mallikma, nukshopa hokut, iakaiyet teli hochito akia isht bohli kut asha meya chatuk: hatuk katiohmi hokak osh, yohmi kut asha chishbn. Himak pilla hono hechik yohmo kashke: amba hvchit tekchi ya ma, okia yumma itintakla ya holisso hush holitoblashke. Yohmikma ohoyo hychia ma, hatak ut hychi hullokma ilap ma hush i hullashka.

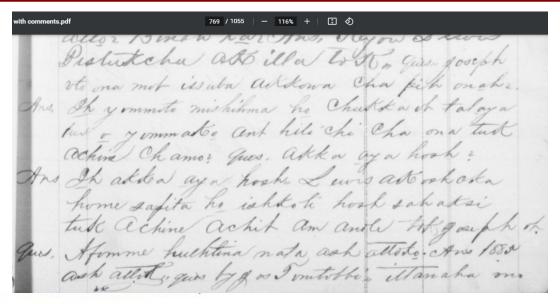
Ohoyo hush kanimi kut hatak huchi masha kia, hatak um asha huchik. ahno hosh, hatak inla ya impa kia bench mayakmut ai impa nuta ai itti hulbit hush maya, hush issashke,

Yohmikma batak hesh kanimi ket ohoyo um asha huchik ahno hosh, yo okla nana ho [yemma ai ititakla ohoyo hani itonowa ka hliohlit hesh itunohonwa na, hvehi pisa li chatuk hysh issashke.

Hatak tekehi bali ut yakulimi chama ai oteni, aholiopa, hatak met tuk, shavali tak a hesh beliokraet, Chihowa Jimatia Muha, ayakolimi hush kanimi kat ugtehi tok finchia thut, ilap ak och i blet ble pramma, hush scha chi kia ushu, yormi kia a upulit glynichi I ponna mulhi mi hatuk osh, bacak tunun I lana bi luchi, ohoyo ik som iksho hoka. Tohto Okchaki.

> Solmon miko nana akaniohmi tok a haklashke

> Himouna ma ohoyo tuklot aboha a chufa ashwa chatuk osh, itatuklot ella



Ans | Keyow Lewis Pistukcha 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 hush sahaksi tuk achine achit am anoli tok Joseph ut-

Ques Afomme hulhtina nata ash attoko.



sr:ds

ona

house

tuk

nom

achine

stand

past

chamo

chamo

past

yummak -o ant hili -chi -cha ona -tuk achine

one

chi

beyond

cha

that

yes

hili

Lex. Gloss

ant

yummako

that acc come and stand:pl causative sr:ss arrive past to seem so Free Yes, There is a house at some distance from the Church. He left his horse there.

The Choctaw National Records

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, focussing 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.

Choctaw trees and roots

- 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.

George Aaron Broadwell

University of Florida

Janet Chávez Santiago

Teotitlán del Valle

Laura Curiel

California State University, Northridge, San Lucas Quiaviní & Oaxacalifornia

Xóchitl M. Flores-Marcial

California State University, Tlacolula de Matamoros & Oaxacalifornia

Moisés García Guzmán

San Jerónimo Tlacochahuava

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 Quiavin

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

Michael Zarafonetis

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

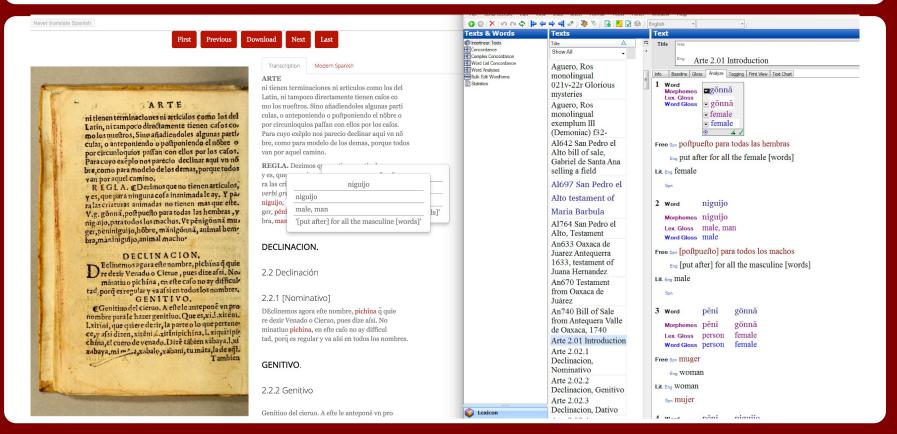
Jamie Henson

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

Berklein, F. (2014). Fortis/Lenis Affricates in Colonial and Modern Valley Zapotec [Undergraduate thesis, Swarthmore College].

Broadwell, G. A., & Lillehaugen, B. (2013). Considerations in the creation of an electronic database for Colonial Valley Zapotec. International Journal of LASSO, 32(2), 77-110.

Broadwell, G. A. (2015). The Historical Development of the Progressive Aspect in Central Zapotec. International Journal of American Linguistics, 81(2), 151-185. https://doi.org/10.1086/679223

Broadwell, G. A., García Guzmán, M., Lillehaugen, B. D., Lopez, F. H., Plumb, M. H., & Zarafonetis, M. (2020). Ticha: Collaboration with Indigenous communities to build digital resources on Zapotec language and history. Digital Humanities Quarterly, 14(4).

Broadwell, G. A., López, F. H., Dubcovsky, A., Flores-Marcial, X., García Guzmán, M., Lillehaugen, B. D., & Zarafonetis, M. (2021). Caseidyneën Saën – Learning Together: Colonial Valley Zapotec Teaching Materials [Online book]. Retrieved from http://ds-wordpress.haverford.edu/ticha-resources/modules/

Publications from Ticha, continued

Broadwell, G. A., López, F. H., Lillehaugen, B. D., Dubcovsky, A., Flores-Marcial, X., García Guzmán, M., Plumb, M. H., & Zarafonetis, M. (2021). Ticha: Collaboration with Indigenous communities to build digital resources on Zapotec language and history. In K. M. Eide & M. Fryd (Eds.), The perfect volume: Papers on the perfect (Studies in Language Companion Series, 217) (pp. 164-177). Amsterdam: John Benjamins. https://doi.org/10.1075/slcs.217.07bro

Foreman, J., & Lillehaugen, B. D. (2017). Positional verbs in Colonial Valley Zapotec. International Journal of American Linguistics, 83(2), 263-305.

Lillehaugen, B. (2016). The syntax of preverbal subjects in Colonial Valley Zapotec. In VII Syntax of the World's Languages International Conference and Workshop.

Lillehaugen, B. D., & Flores-Marcial, X. (2023, May 15). Ticha: Exploring Antiracist Markup and Community Engaged Digital Scholarship. Paper presented at the Women Writers Project Antiracist Markup Practices Symposium, Northeastern University.

Metzger, J. (2019). Colonial Valley Zapotec Effects on Bilingual Spanish [Undergraduate thesis, Swarthmore College].

Nash, J. E. (2022). Accounting for toponyms with multiple classifiers in Colonial Valley Zapotec [Undergraduate thesis,