Zapotec Language Activism and Talking Dictionaries

K. David Harrison[1], Brook Danielle Lillehaugen[2],

Jeremy Fahringer[3], Felipe H. Lopez[4]

[1] Swarthmore College, 500 College Ave., Swarthmore PA 19081, and New York Botanical Garden, 2900 Southern Blvd, The Bronx, NY 10458

[2] Haverford College, 370 Lancaster Ave., Haverford PA 19041

[3] Swarthmore College, 500 College Ave., Swarthmore PA 19081

[4] University of California, San Diego, 9500 Gilman Dr., Literature Dept 0410, La Jolla CA 92093

E-mail: harrison@swarthmore.edu, blilleha@haverford.edu, jfahrin1@swarthmore.edu, lieb@ucla.edu

Abstract

Online dictionaries have become a key tool for some indigenous communities to promote and preserve their languages, often in collaboration with linguists. They can provide a pathway for crossing the digital divide and for establishing a first-ever presence on the internet. Many questions around digital lexicography have been explored, although primarily in relation to large and well-resourced languages. Lexical projects on small and under-resourced languages can provide an opportunity to examine these questions from a different perspective and to raise new questions (Mosel, 2011). In this paper, linguists, technical experts, and Zapotec language activists, who have worked together in Mexico and the United States to create a multimedia platform to showcase and preserve lexical, cultural, and environmental knowledge, share their experience and insight in creating trilingual online Talking Dictionaries in several Zapotec languages. These dictionaries sit opposite from big data mining and illustrate the value of dictionary projects based on small corpora, including having the flexibility to make design decisions to maximize community impact and elevate the status of marginalized languages.

Keywords: lexicography; collaboration; endangered languages; Zapotec

1. Introduction

Dictionaries, whether print or digital, are much more than just an organized collection of words with definitions or translations. At their best, they are living repositories of the collective knowledge base compiled by and for the community that owns it. When shared with outsiders, they can also provide a window into a culture, its traditions, beliefs, and values. Dictionaries can serve as accurate records of the historical and contemporary state of a language, to the extent that they are inclusive of variation and not highly standardized. Finally, dictionaries can shape the future development of a language. They can influence the vitality of a language, expanding its domains of use. And, intentionally or not, they contribute to processes of standardization.

The Zapotec lexicography project draws inspiration from Native American Language online dictionaries such as the Passamaquoddy-Maliseet Language Portal (https://pmportal.org) and the Lenape Talking Dictionary (http://www.talk-lenape.org). Both of these projects represent an intense effort by these respective communities to reclaim, record, share, and generationally transmit their severely endangered languages. They serve as proof that indigenous languages can thrive thanks to the digital activism of community members and linguists. Methodology—not just final product—is of central importance in the creation of these dictionaries: the collaborative practices as well as the resulting resources can be interventions in contexts where discrimination and detrimental linguistic ideologies conspire to silence languages, such as those described by Sicoli (2011) based on ethnographic research in Southern Sierra Zapotec communities.

In this paper, we describe the ideation, design, building, and sharing of a suite of five Zapotec Talking Dictionaries. We also discuss unintended uses and effects, and what we hope are positive impacts on community practices and ideas about language maintenance.

2. Talking Dictionary project history

In 2003, one of the current study's co-authors, Harrison, published a print Tuvan Dictionary (Anderson & Harrison, 2003), and upon distributing it to the community received questions and feedback suggesting that some words were missing. Such experiences are ubiquitous—the very same sentiment is noted on page 1 of *Making Dictionaries* (Frawley et al., 2002: 1) in a long list of challenges when making dictionaries. This led to the realization that print was not the best medium for dictionaries, and so we began to put some of our many lexical field recordings into a searchable online format, which became the "Tuvan Talking Dictionary" launched in 2006 (Harrison & Anderson, 2006). We have considerably expanded the platform since then in terms of technological capabilities, design, and community participation. As of 2019, we have 120 Talking Dictionaries in varying stages of development. The total number of lexical entries is 150,000+, with the largest collections being the Gutob Talking Dictionary (Anderson & Harrison, 2016) with 13,338 entries, and the Siletz Dee-ni Talking Dictionary (Anderson & Harrison, 2007) with 10,552 entries. The lexical entries come from a variety of sources: (a) field recordings, (b) recordings made during digital lexicography workshops we have hosted, and (c) recordings made on an ongoing basis by online co-authors working in places such as India, Mexico, and Vanuatu.

The Valley Zapotec Talking Dictionaries began in 2012 with the creation of the Tlacolula Valley Zapotec Talking Dictionary. Another of the current work's co-authors, Lillehaugen, created this using already existing audio recordings as a mock-up to show members of the San Lucas Quiaviní and Tlacolula de Matamoros communities as a way to gauge interest in developing the dictionary further. During a field trip to Oaxaca during summer 2013, Lillehaugen met with the authorities and community members in

San Lucas and Tlacolula. The feedback was clear: members of both communities were interested in developing the dictionaries further, and both wanted dictionaries that represented only the language variety as spoken in their community. Thus, while the Zapotec of San Lucas Quiaviní and the Zapotec of Tlacolula de Matamoros may be considered dialects of the same language on linguistic grounds—both are classified with the ISO 639-3 code [zab] (Eberhard et al., 2019)—that was not the relevant criterion for interested community members. In response to this, the mock-up dictionary was split into two dictionaries in 2013—the first two Valley Zapotec Talking Dictionaries: Tlacolula de Matamoros (Lillehaugen et al., 2013) and San Lucas Quiaviní Zapotec (Lillehaugen et al., 2109a), the latter co-authored and locally directed by a native Zapotec speaker and co-author of this paper, Felipe H. Lopez. Soon thereafter, two additional communities joined in: San Jerónimo Tlacochahuaya (Lillehaugen & García Guzmán et al., 2019) and Teotitlán del Valle (Lillehaugen and Chávez Santiago et al., 2019). Most recently, the Talking Dictionary for San Bartolomé Quialana was started in summer 2019 (Lillehaugen et al., 2019b).

3. Design and features

The intention in designing the Talking Dictionaries was to create a multimedia resource (audio, video, photo, text, maps) for small languages that went beyond traditional dictionary design and content. The user experience would be paramount, while the back-end design would be secondary. It would be a living, constantly expanding resource that was community-authored, community-owned, and fully attributed by name to all contributors. The interface would be easy to access (online, on smartphones, or even as a paper printout), and would use simple iconography (for example, an ear icon for sound files, as seen in Figure 1). It would be rich in content from the very first encounter: the user would never be confronted with only a blank "search" box on the front page or a null search result. Regardless of what a user searched for, some content would always appear. The back-end design would not be based on specialist or proprietary software, but would use widely available and well-supported database software: we chose MySQL, an open-source relational database management system. (Note that the examples illustrated throughout this text show the English language interface of the Talking Dictionary, since this article is in English. The Zapotec Talking Dictionaries also have a Spanish-language interface.)

Figure 1: Ear icon for sound files

We learned from our experience with other digital lexicography efforts, such as LEXUS (Kemps-Snijders & Wittenburg, 2006), a tool designed at MPI-Nijmegen and launched in 2001. LEXUS prioritized information architecture over user experience. As a result, it proved to be of limited appeal beyond those individual researchers who were uploading their lexical data, and even such specialized users needed to engage in significant learning and adaptation to the unique environment (Wojtylak, 2012). As one reviewer observed: "LEXUS is obviously a tool for the creation and maintenance of a lexicon rather than for the visually appealing presentation of lexical data" (Kochetova, 2009: 244). It remained in circulation until 2017, but is no longer supported. Likewise, Bergenholtz and Bothma (2011: 55) warn of "information death" resulting from an overwhelming presentation of content, especially in a digital context. We hoped to avoid these scenarios of limited user appeal and technical support leading to obsolescence.

Design decisions for any particular Talking Dictionary are made together with local co-authors. For example, the dictionaries were originally designed with a computer web-based interface in mind, but it quickly became clear that most users were accessing these Talking Dictionaries from their smartphones. Future development was thus optimized with smartphone use in mind. As all of the dictionaries are supported from the same back-end structure, solutions for needs that originate in one community can end up providing design enhancements for other Talking Dictionaries.

3.1 Acknowledgement and crediting of expertise

A key feature of the dictionaries is that expertise is identified and credited with every entry (see Figure 2). This overt recognition of knowledge holding is a form of decolonizing lexicography, and a response to anthropological and linguistic practices which erased the names of local experts, as well as the larger issue of the exclusion of indigenous authority, as argued in Anderson and Christen, 2019. In these dictionaries, you hear and see the names of the experts behind each word. Copyright is explicitly mentioned in each dictionary as belonging to the community, and authorship of the dictionaries includes the Zapotec co-authors.
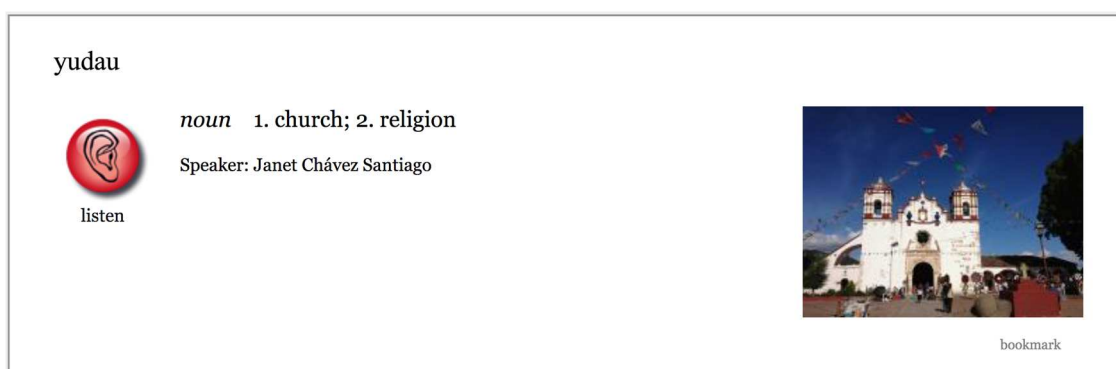

Figure 2: Speaker credit at the entry level
(Lillehaugen & Chávez Santiago et al., 2019: entry 248)

3.2 Semantic domains

Dictionaries can be organized and browsed based on a set of semantic categories (e.g., kinship terms, food, botany) that are dynamic and customizable at the dictionary level. Any new semantic domain can be created and immediately available. This freedom allows for flexibility and creative experimentation with ways of interacting with the words in a dictionary, and for highlighting domains of (specialized) knowledge that are important at the local level. For example, Teotitlán del Valle is a weaving town, so the category "weaving" is crucial and contains scores of entries, as seen in the list on the left in Figure 3. "Weaving" does not (currently) occur as a semantic category in the dictionary for San Jerónimo Tlacochahuaya, an agricultural town, illustrated in the list of semantic domains on the left in Figure 3.
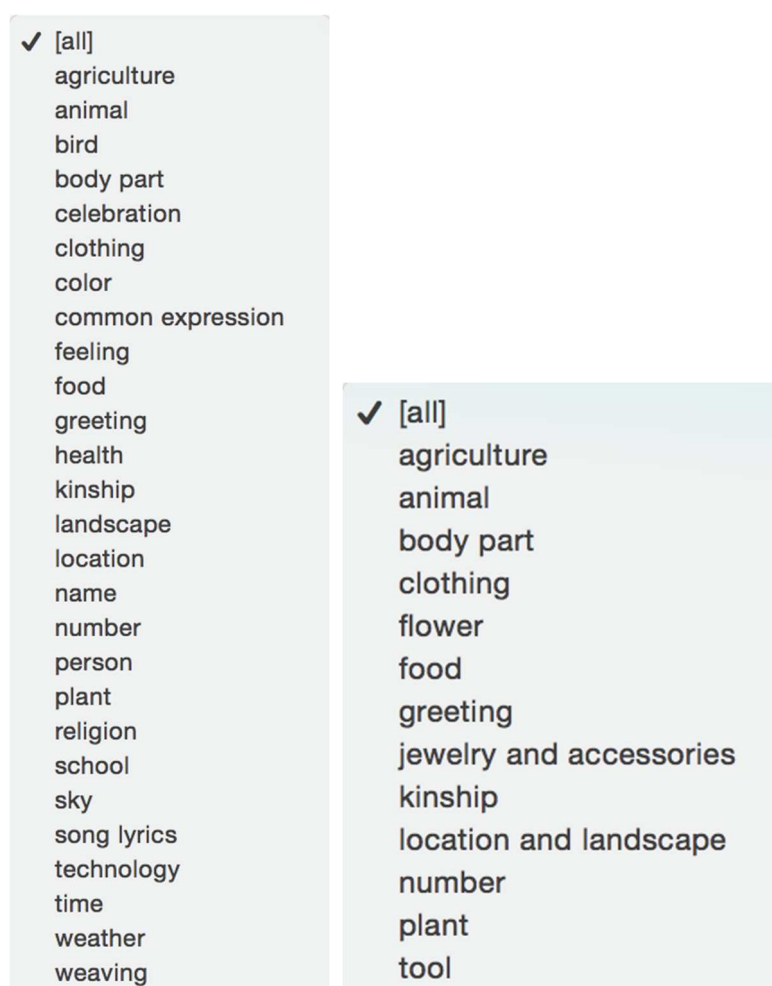
Figure 3: Semantic domains in two Valley Zapotec Talking Dictionaries
(left Teotitlán; right Tlacochahuaya)

3.3 Dictionaries without standardized orthographies

Our design anticipates the possibility of making a dictionary for a language with an emerging orthography or no orthography. Some communities that want to develop dictionaries lack an orthography and may not have plans to develop one, or may still be working towards a consensus about the best orthography. We can accommodate multiple orthographies (showing various alternate spellings under a single entry), or none at all. In the case where no orthography exists, we do not attempt to devise one, as we believe this work is best done by the community itself, not by outsiders. Instead we can use IPA transcription (which we fully acknowledge is not appealing to indigenous speech communities), while leaving blank for future use the field where an orthographic spelling would go. Another option is to use the spelling preferences of the speaker, whether or not those preferences are part of a systematic set of decisions on how to spell the sounds of the language.

The San Lucas Quiaviní Zapotec dictionary uses the orthography defined in Cali Chiu (Munro et al., 2008). All other Valley Zapotec dictionaries exist in the absence of a standardized orthography. For these dictionaries, we use the spelling preferences of the speaker. This means that words and sounds may be spelled inconsistently throughout the dictionary, and single words may be spelled more than one way. While this may make some lexicographers and linguists uncomfortable, it reflects the current practice and linguistic reality of these speech communities, where individuals are used to deciphering personal spelling decisions, like those used, for example on store signage or social media. Trying to read Zapotec on Facebook and Twitter is a frequent experience for co-author Felipe H. Lopez, and two examples of Twitter exchanges across orthographic differences in Valley Zapotec can be seen in Figures 4 and 5 in Lillehaugen (2019). Moreover, "[c]ommunity discussions about standardized orthographies can sometimes become unproductive, and these debates can even impede other advances in increasing the use of the language. In some cases, these disagreements can turn into 'orthography wars' (Hinton, 2014), draining the precious time and energy of the activists involved" (Lillehaugen, 2016: 367).

The Valley Zapotec practice around spelling represents the full continuum between idiosyncratic, ad hoc spelling choices on one end, and fully developed, orthographic systems based on a phonological analysis of the language on the other. Our Talking Dictionaries support writing systems at any point in this continuum, and are flexible to change as decisions about writing choices in a community emerge. The design of the dictionary allows us to include multiple spellings in a single entry and multiple entries for a single word, each potentially with a different spelling. For example, 'flower' in the San Jerónimo Tlacochahuaya dictionary is currently spelled both *gie* and *gie'*, the latter marking the word-final glottal stop. In the San Bartolomé Quialana dictionary we find even more diversity in the spelling of 'flower', which currently includes: *gi*, *gui*, *glli*, and *gyi'i*.

This feature has had the added benefit of representing dialect variation within a pueblo, without being forced to choose one of the forms as the headword. Multiple parallel entries and pronunciations, where no one is primary or authoritative, can exist side by side. While we currently have between one and four pronunciations for each word, as Garrett points out, hearing multiple voices is a benefit to those who might be using a dictionary in language learning efforts.

> In the online Yurok dictionary, we have tried to include audio examples of as many words and short phrases as possible, spoken by as many fluent speakers as possible. Users report greatly appreciating a chance to hear the range of variation that would have been present in the speech community when Yurok was still used as a first language in many households... Users can hear recordings as spoken by six fluent speakers recorded in the 2000s. (Garrett, 2019: 201)

3.4 Integration of multimedia

The heart of each entry is an audio recording of the word by a native speaker, and this is one reason why spelling and orthographic choices need not be critical constraints. In addition to translations in multiple languages, other types of multimedia can be included in any particular entry. Entries may include images or photos. These might be the photographs or artwork of participants in the dictionaries, such as in Figure 4 or Figure 5, or other images available for use under creative commons licenses, such as those from iNaturalist.org in Figure 6, and the line drawings from the ILV / SIL Artwork for Literacy in Mexico (ILV, 2004) in Figure 7. The experience of browsing the Talking Dictionaries is enriched by these visual complements, which is also akin to real-life experiences that speakers have with images when forming cognitive representations of lexical items, especially in the domain of specialized knowledge (Faber, 2012: 225-226).
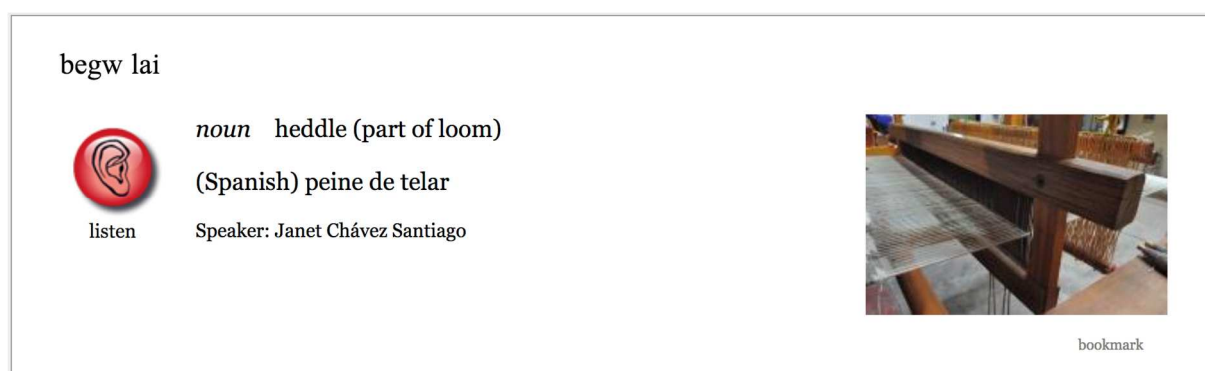

begw lai

noun   heddle (part of loom)

(Spanish) peine de telar

listen   Speaker: Janet Chávez Santiago

bookmark

Figure 4: Original photography (Lillehaugen & Chávez Santiago et al., 2019, entry 921)

Figure 5: Original artwork by María Mercedes Mendez Morales
(Lillehaugen & García Guzmán et al., 2019, entry 2858)

Figure 6: Creative commons photos (iNaturalist.org) in lexical entry
(Lillehaugen et al. 2019a, entry 3764)
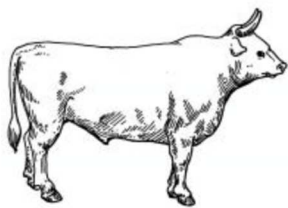
Figure 7: Creative commons artwork (ILV, 2004) in lexical entries
(Lillehaugen et al., 2019a, entry 458)

The entry in Figure 7 illustrates a further feature of the platform, which allows a tweet to be linked from, and embedded in, an entry. These languages have very little in terms of a native speaker written corpus, and most of that corpus is "born digital" on Twitter. In fact, many of the same Zapotec co-authors of the dictionaries are also individuals who write on Twitter in their language (see Lillehaugen, 2016 & 2019). Twitter is not the only domain of digital language activism, as Zapotec language can also be found on YouTube in a variety of contexts including language lessons created by native speaker teachers, such as those created by Talking Dictionary co-author Moisés García Guzmán (https://www.youtube.com/user/BnZunni), Zapotec language materials created for the purpose of language documentation, including expressly for the purpose of illustrating a lexical entry, and other Zapotec language materials that aren't expressly for teaching or documenting Zapotec language. The video embedded in the entry in Figure 8 is a 3-minute long episode of a documentary web series (Dizhsa Nabani; García Guzmán et al., 2018). This episode illustrates the preparation of beans with narration in Zapotec, and adds significant cultural context to the entry.
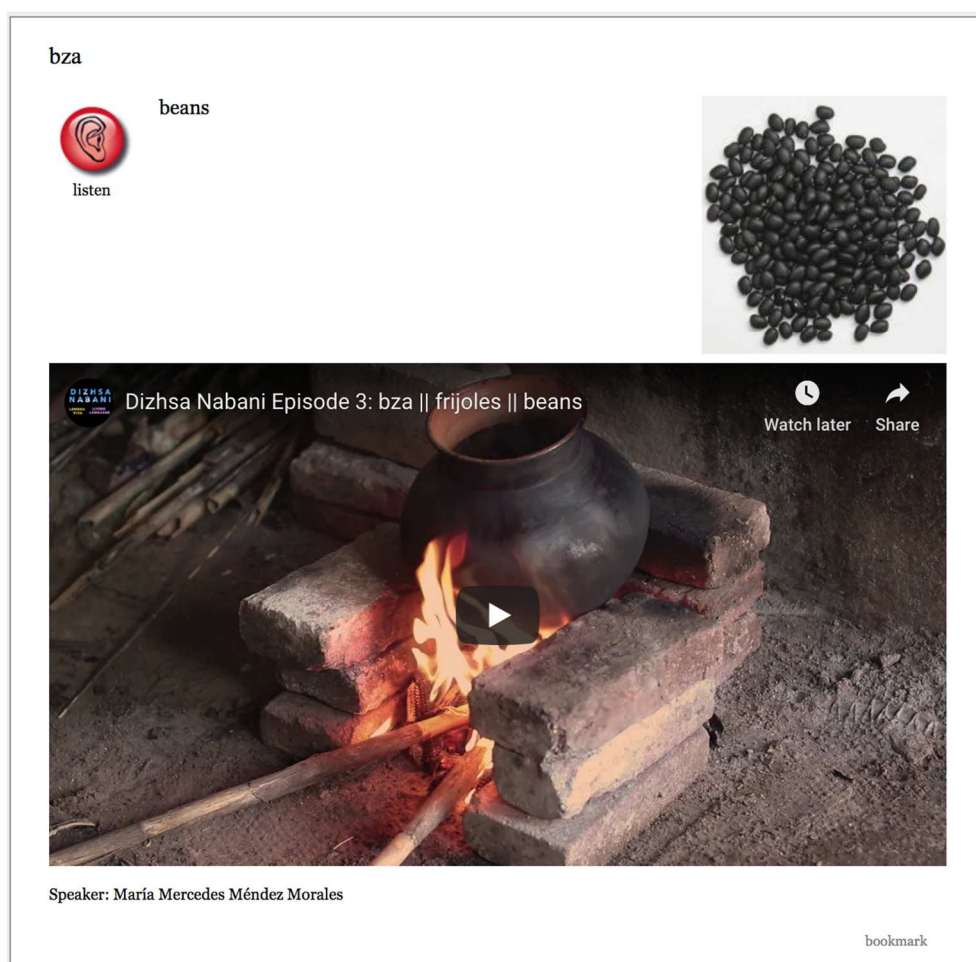

Figure 8: Embedding and linking of YouTube videos in lexical entries
(Lillehaugen & García Guzmán et al., 2019, entry 3471)

Consistent with Biesaga's (2017: 232) observation regarding the use of illustrations in her sample study, multimedia is used extensively for entries in the semantic domain of plants in the Zapotec Talking Dictionaries. When videos are associated with entries for names of plants, they often include a (monolingual) Zapotec scientific explanation, illustrating the more encyclopaedic nature of some of the multimedia components (cf. Biesaga, 2017: 1). In a Zapotec Talking Dictionary being developed by another team for a Northern Sierra Zapotec language, Macuiltianguis (Foreman & Martinez et al., 2019), the embedded videos are used to show verb paradigm information in an accessible format, revealing some of the range of possibilities teams are exploring in the utilization of embedded multimedia.

4. Methodology

The Zapotec Talking Dictionaries have grown since 2013, through the work of linguists, undergraduate students, and Zapotec speakers. Each Zapotec dictionary has local Zapotec co-authors, and the work is highly collaborative, with many community members participating in the creation of the dictionaries. Intense periods of work, usually in the summer, also serve as training and research experience for undergraduate students and Zapotec co-authors. While the original intention was to set up work during the summer that the Zapotec partners would continue during the year, we have come to accept that the natural rhythm of the project varies over the course of the year, with lots of additions of words over the summer, and slower work during the academic year that may focus on technical advances, corrections, and linguistic analysis.

Three of the co-authors of the San Lucas Quiaviní Zapotec Talking Dictionary (Munro, Lopez, and Lillehaugen) were involved in the creation of two print dictionaries on Tlacolula Valley Zapotec: the tri-lingual San Lucas Quiaviní Zapotec dictionary (Munro & Lopez et al., 1999) and the fourth volume of the textbook *Cali Chiu?* (Munro et al., 2008), which uses a revised, simplified orthography based on Munro and Lopez et al. (1999). These print dictionaries were our starting points for the San Lucas Quiaviní Talking Dictionary.

For both this and the other Zapotec Talking Dictionaries we use a variety of elicitation techniques based on: (i) legacy published sources that are out of copyright or, if copyrighted, up to 10% can be used in fair use; (ii) prepared word lists, such as SIL's Rapid Word Collection (http://www.rapidwords.net/), or photos from iNaturalist.org; (iii) community generated lists (e.g. photo elicitation); (iv) born-digital corpora (like Zapotec Twitter); and (v) pedagogical materials that are published or informally circulated. Finally, (vi) the more established Zapotec Talking Dictionaries can now serve as starting wordlists for the newer ones.

New lexical data can be gathered by photo elicitation techniques or thematic conversations (around kinship, foodways, etc.). Part of our workflow involves having

large public workshops in the pueblos, where interested community members are invited to record words or otherwise contribute to the project. These workshops facilitate a type of crowdsourcing, which, as Čibej et al. (2015: 71) put forward, "could have lasting consequences on the nature of lexicographic work... as well as the perception, use, and life-cycle of the lexicographic product". Such community workshops may be held outside in public spaces, as seen in Figure 9, or in collaborators' homes, as shown in Figure 10, both from Teotitlán de Valle. On more than one occasion, local experts have arrived with their own, often extensive, word lists and dictionaries that they have compiled out of a love for the language, such as the one that can be seen in Figure 10, created by Froilán Carreño Gutiérrez, which contained an impressive number of names of local avifauna. The hand drawn images in Figure 5 came from another personal dictionary created by the teacher María Mercedes Mendez Morales.

Figure 9: Talking Dictionary workshop in Teotitlán del Valle. Photo credit Brook Lillehaugen.

As Zapotec speakers across generations are involved in the creation of the Talking Dictionaries, the methods employed have also facilitated intergenerational language learning. For example, while documenting the semantic domain of weaving in Teotitlán del Valle, Janet Chávez Santiago, the local director of the Talking Dictionary, learned specialized terminology previously unknown to her. Likewise, Felipe H. Lopez learned

many names for medicinal plants as part of dictionary work in San Lucas Quiaviní. In the creation of the dictionaries, even fluent speakers can learn and share knowledge. We have also observed dictionary work fostering meta-linguistic conversations between speakers. While language revitalization work is complex and there is no simple panacea, we value every positive step forward, each small shift in ideology, and all these small moments of language learning.

5. User experience and creative uses

We view the Zapotec Talking Dictionaries as living projects: as both resources and sites for collaboration. As such, user experience with the dictionaries and novel, unexpected uses are particularly inspiring. We cannot always know ahead of time exactly how individuals will want to interact with the dictionary. As attested in the following quote from Teotitlán del Valle Zapotec Talking Dictionary co-author Janet Chávez Santiago, sometimes this may change for a user over time:

> When I started teaching Zapotec as a second language I did not have any kind of pedagogical material printed or online. I basically was creating my own material as my classes were developed. In 2012, when I met Professor Lillehaugen for the first time, I told her about the lack of material in my Zapotec variant. She mentioned to me about the Talking Dictionary platform and, without hesitating, offered me her support to create one for my Zapotec variant. At the very beginning I saw the Talking Dictionary as a resource to help my students to practice and learn the language outside the classroom, but soon I understood that it is more than what you can listen to and see on your screen as the final "product". For me, the Talking Dictionary is a tool that helps to document the language and, most importantly, thanks to the technical and academic support, it is a tool that reunites and involves a community of speakers that reflect on the language before recording any word, and on how the speakers want to show and teach the language to the outside through the dictionary.
>
> (Janet Chávez Santiago, personal communication June 6, 2019)

Figure 10: Janet Chávez Santiago recording words with Froilán Carreño Gutiérrez, seen holding his personal dictionary of Teotitlán del Valle Zapotec. Photo credit Brook Lillehaugen.

5.1 Comparing languages

Just as Chávez Santiago noted that her view of the dictionary changed over time, so has our perception of particular features of the Talking Dictionaries. One such example is the "compare languages" feature, illustrated in Figure 11 for the word 'guava'. This feature was originally designed with the linguist user in mind, as a way to give insight into the complex dialect continuum of the Tlacolula Valley. When speakers could not remember a particular word, we showed this view and asked if they would like to hear it in other varieties. After listening to other ways of saying the word in nearby pueblos, speakers were often reminded of the word in their language, saying things like, *ah, that sounds like what my grandmother used to say.* Moreover, participants were interested in hearing the words in other varieties even when they knew the word in their own language. After the fact, we realized this should not have surprised us, given the intense level of both active multilingualism / multi-dialectalism and passive understanding of other varieties that exists in the Tlacolula Valley. The dictionary became another space where speakers could hear multiple varieties of Zapotec, echoing events like the weekly

market in Tlacolula, where many Zapotec languages and varieties can be heard side by side.
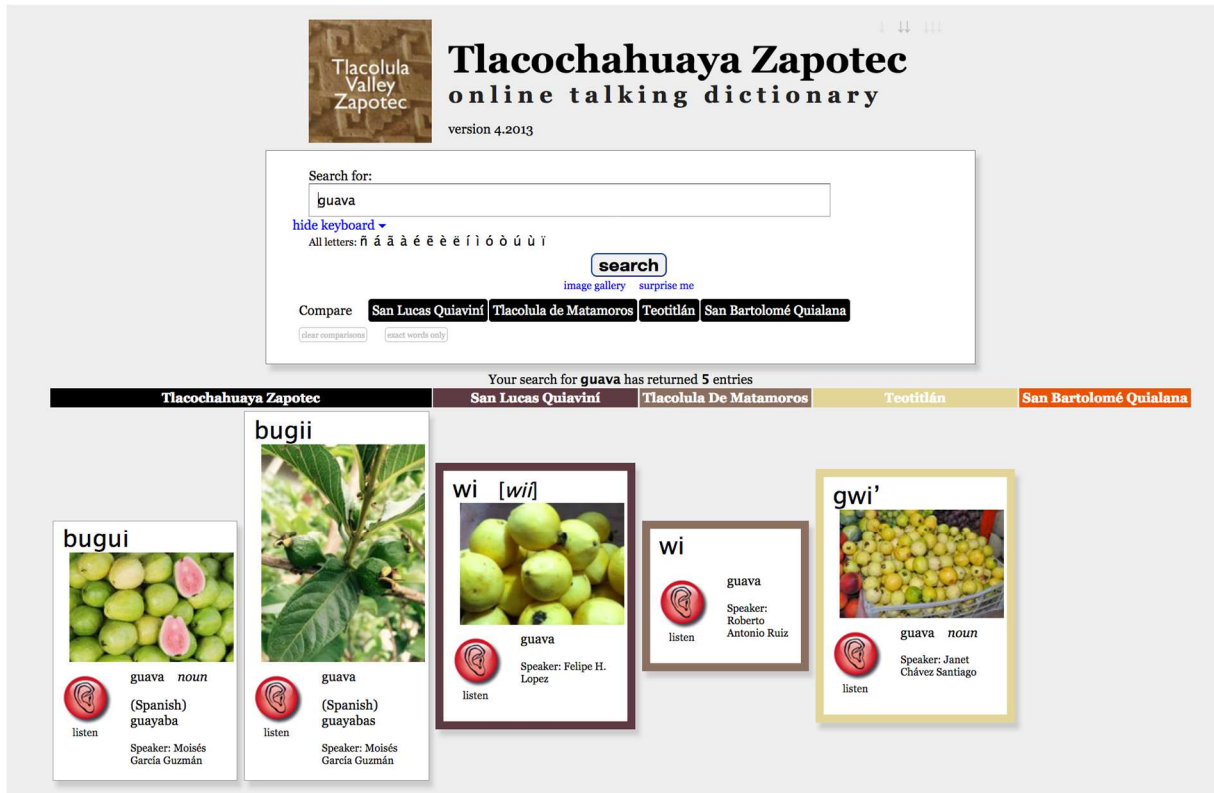

Figure 11: The compare dictionaries feature for "guava".

5.2 Comparing languages

Another area of unexpected development was growing synergies with other digital platforms. As described above, tweets can be embedded in entries. Likewise, each entry has a unique URL, and these entries can be shared through social media. This type of two-way relationship can also be seen in relation to the documentation of the natural world and biodiversity. We utilize photographs from crowd-sourced naturalist sites like iNaturalist.org and Fishbase.org. This, too, creates a two-way relationship with a global network of scientists, conservationists, and amateur or expert naturalists. Fishbase, in particular, has declared an interest in including "vernacular" names for species, and is open to contributions in Zapotec or any other language.

Two of the co-authors of this paper are also involved in an online text explorer for a corpus of Zapotec language texts written in the Mexican Colonial period: Ticha (http://ticha.haverford.edu; Lillehaugen et al., 2016). As this digital scholarship project focuses on a historical corpus of Zapotec texts, the team wanted to be careful to "not reinforce [the] harmful false ideology that Zapotec language and people are only of the past, frozen in time" (Broadwell et al., to appear). The team describes how the Zapotec Talking Dictionaries were used as one intervention:

One way we addressed this was by bringing Zapotec voices to the site. Figure 3 [Figure 12 in this text] shows one of the resources available on Ticha: a vocabulary of the most common words found in the corpus... Wherever possible, we connect these lexical entries for historical forms of words with their modern counterparts, by linking entries in Ticha's Vocabulary with entries in online Talking Dictionaries for several Valley Zapotec language varieties... The design came out of one of the in-person workshops in Oaxaca. As the room full of Zapotec speakers from different communities in the Valley of Oaxaca worked through understanding one of the Colonial era texts together, a pattern of practice emerged. For each word, speakers would go around the table, saying the modern cognate in their variety of Zapotec. The text was read, performed—even echoed—in a multitude of modern Zapotec languages. Ticha's Vocabulary is our attempt to realize this in a digital format. (Broadwell et al., to appear)

Figure 12: Connections to historical corpus on Ticha

5.3 Connecting communities in Oaxaca and the diaspora

As is the case for many indigenous communities in southern Mexico, there is a large diaspora community of Zapotecs in the United States. In the case of Valley Zapotec communities, California (or "Oaxacalifornia"), and especially the greater Los Angeles

area, has become home to hundreds of thousands of Zapotec people (Lopez & Runsten, 2004). The Zapotec Talking Dictionaries, thus, serve—and are created by— transnational communities. Felipe H. Lopez notes that he appreciates being able to point members of his community to the Talking Dictionary when they ask him for resources on how to write their language—and he starts by pointing out that this community exists "on both sides of the border":

> The Talking Dictionary is a very useful and important resource for my community both in San Lucas Quiaviní, Oaxaca, and for many of those who live in California in order to preserve the language. Some members of the pueblo of Quiaviní, on both sides of the border, have requested help to either learn how to write the language or for written material to teach Zapotec to their children or to learn it themselves. Most people who speak the language don't write. For instance, a young woman who leads a local group of young Zapotecs in San Lucas asked me to help them write poems and local stories in Zapotec. In another case, an artisan couple wanted to incorporate Zapotec writing in their promotions for their local textiles and promote their work in various public social spaces in the City of Oaxaca. They reached out to me to help them with written materials in Zapotec. More recently, a Quiaviní woman asked me to help her to translate a Spanish song into Zapotec. Also, within the diaspora community there has been an interest in learning more about the Zapotec language. For example, a Zapotec college student reached out to help her get a Zapotec dictionary or other written materials that would help her to learn her parent's language. Additionally, a couple of parents with young children living in Los Angeles have requested material that they can give their children to learn Zapotec.

As the co-author of the print dictionary of his language (Munro & Lopez et al., 1999), Lopez often receives such requests. He further noted, "before the online dictionary, there was little I could do [in response to these requests] because the San Lucas Quiaviní dictionary is very hard to obtain since it is no longer in print and is expensive. Now with this Talking Dictionary, I can refer people to it, not only to see written Zapotec but also to listen to it. Most of all, it is free."

5.4 Reaffirming kinship relations and Zapotec identity

Given the diasporic nature of Valley Zapotec communities, it may not be surprising that Zapotec users pay close attention to the identity of the individual speaking for each entry. Many users want to make sure they understand who this person is and how they are related to them, displaying knowledge of kinship relations across the diaspora. This tracking and re-affirming of kinship ties is also a reaffirmation of ethnic identity and mirrors other modes of affirmation of belonging to a pueblo.

Even within the pueblo, the identity of speakers is of central importance in the user experience with the Talking Dictionaries, as these words and this knowledge are never separate from those who came before, as expressed clearly by Moisés García Guzmán, Secretary of Culture of the pueblo of San Jerónimo Tlacochauaya, and co-author of the Talking Dictionary for his language (Lillehaugen & García Guzman et al., 2019):

> The Talking Dictionary personally means that we are able to document all this knowledge in a digital platform and share it with others, and that ultimately leads to better preservation efforts for the language and the community. It's been great to share words that at the same time remind me of stories that my grandmother taught me and be able to link them.

> Moisés García Guzmán, personal communication, June 6, 2019

6. Looking forward

We plan to continue developing and expanding the Talking Dictionary platform as a living cultural repository that is community-owned, inclusively co-authored, and fully attributed. The history of lexicography may have largely belonged to the empowered gatekeepers of knowledge enforcing static linguistic norms (Mugglestone, 2011). But the future of lexicography—as envisioned in our work, and in the other papers in this volume—looks very different. It is a future of words technologized, yet remaining under collective ownership and individual authority. Digital lexicography has the potential for constant expansion and can reflect dynamic language change and variation. Finally, by supporting local agency over linguistic resources, the Talking Dictionaries can play a positive role in community-based language revitalization and maintenance. As Moisés García Guzmán said to us: "This is the best tool that we could ever have to save our languages."

7. Acknowledgements

and Floriana Hernández Martínez for allowing us to work with them and learn alongside them. *Xtyozën yuad!*

8. References

Anderson, J. & Christen, K. (2019). Decolonizing Attribution: Traditions of Exclusion. *Journal of Radical Librarianship* 5, pp. 113–52.

Anderson, G. D.S. & Harrison, K. D. (2003). *Tuvan Dictionary.* München: Lincom Europa.

Anderson, G. D.S. & Harrison, K. D. (2007). *Siletz Talking Dictionary.* Living Tongues Institute for Endangered Languages. http://siletz.talkingdictionary.org

Anderson, G. D.S. & Harrison, K. D. (2016). *Gutob Talking Dictionary.* Living Tongues Institute for Endangered Languages. http://www.talkingdictionary.org/gutob.

Bergenholtz, H., & Bothma, T. J. D. (2011). Needs-adapted Data Presentation in e-Information Tools. *Lexikos / AFRILEX-reeks* 21: 53-77.

Biesaga, M. (2017). Pictorial Illustrations in Encyclopaedias and in Dictionaries – a Comparison. In I. Kosem, C. Tiberius, M. Jakubíček, J. Kallas, S. Krek, & V. Baisa (eds.), *Electronic lexicography in the 21st century. Proceedings of eLex 2017 conference*, pp. 221-236.

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

Eberhard, D. M., G. F. Simons, & Fennig, C. D. (eds.). (2019). *Ethnologue: Languages of the World.* Twenty-second edition. Dallas, Texas: SIL International. Online: http://www.ethnologue.com.

Frawley, W., Hill, K. C. & Munro, P. (2002). *Making Dictionaries: Preserving Languages of the Americas.* Berkeley: University of California Press.

García Guzmán, M., Brashear, H., Deutch, L., Evans, S. K., Funari, V., Goldberg, K., Jauregui-Volpe, M., Lillehaugen, B. D., Ogborn, E., Palmarini, L. & Rodgers, C. (producers). (2018). *Dizhsa Nabani – Lengua Viva – Living Language.* United States: Haverford College. Online: https://doculabs.haverford.edu/dizhsanabani/.

Garrett, A. (2019). Online Dictionaries for Language Revitalization. In L. Hinton, L. Huss, & G. Roche (eds.) *The Routledge handbook of language revitalization.* Abington-on-Thames: Routledge, pp. 197-206.

Harrison, K. D. & Anderson, G. D. S. (2006) *Tuvan Talking Dictionary.* http://tuvan.talkingdictionary.org

Hinton, L. 2014. Orthography wars. In M. Cahill & K. Rice (eds.), Developing orthographies for unwritten languages. Dallas: SIL International Publications, , pp. 139–168.

Instituto Lingüistico de Verano, A. C. (2004). *Arte para la Alfabetización en México (Art for Literacy in Mexico).* Online: http://www.mexico.sil.org/es/publicaciones/arte4lit.

Kemps-Snijders, M. & Wittenburg, P. (2006). LEXUS- a web-based tool for manipulating lexical resources. International conference on Language Resources and Evaluation, Genoa, Italy. Accessed at: https://pure.mpg.de/rest/items/item_61151_3/component/file_61152/content. (15 July 2019)

Kochetova, K. (2009). LEXUS: from Max Planck Institute for Psycholinguistics Nijmegen. *Language Documentation and Conservation.* Vol. 3, No. 2 (December 2009), pp. 241-246.

Čibej, J., Fišer, D. & Kosem, I. (2015). The role of crowdsourcing in lexicography. In I. Kosem, M. Jakubiček, J. Kallas, & S. Krek (eds.), *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference.* Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd., pp. 70-83.

Faber, P. (2012). A Cognitive Linguistics View of Terminology and Specialized Language. Berlin/Boston: De Gruyter Mouton.

Foreman, J., Martinez, M. with D. Arellano, R. Cabrera, L. Closner, K. Grimaldo, & F. Pérez Ruiz. 2019. *Macuiltianguis Zapotec Talking Dicitionary,* version 0.1. Living Tongues Institute for Endangered Languages. https://talkingdictionary.swarthmore.edu/macuiltianguis/

Lillehaugen, B. D. (2016). Why write in a language that (almost) no one can read? Twitter and the development of written literature. *Language Documentation and Conservation* 10: 356-392. Online: http://hdl.handle.net/10125/24702.

Lillehaugen, B. D. (2019). Tweeting in Zapotec: social media as a tool for language activists. In J. C. Gómez Menjívar & G. E. Chacón (eds.) *Indigenous Interfaces: Spaces, Technology, and Social Networks in Mexico and Central America,* 202—226. Tucson: University of Arizona Press.

Lillehaugen, B. D., Antonio Ruiz, R. & Antonio Ruiz, J. with C. Batten, H. M. Felker, A. Mannix, K, D, McCormick, & R. E. Weissler. (2013). *Tlacolula de Matamoros Zapotec Talking Dictionary*, pilot version. Living Tongues Institute for Endangered Languages. Online: http://www.talkingdictionary.org/tlacolula.

Lillehaugen, B. D., Broadwell, G.A., Oudijk, M.R., Allen, L., Plumb, M.H., & Zarafonetis, M. (2016). Ticha: a digital text explorer for Colonial Zapotec, first edition. Online: http://ticha.haverford.edu/.

Lillehaugen, B. D. & Chávez Santiago, J. with A. Freemond, N. Kelso, J. Metzger, K. Riestenberg, & K. D. Harrison. (2019). *Teotitlán del Valle Zapotec Talking Dictionary*, version 2.0. Living Tongues Institute for Endangered Languages. Online: http://www.talkingdictionary.org/teotitlan.

Lillehaugen, B. D. & García Guzmán, M. with K. Goldberg, M. M. Méndez Morales, B. Paul, M. H. Plumb, C. Reyes, C. G. Williamson, & K. D. Harrison. (2019). *Tlacochahuaya Zapotec Talking Dictionary*, version 2.1. Living Tongues

Institute for Endangered Languages. Online: http://www.talkingdictionary.org/tlacochahuaya.

Lillehaugen, B. D., Lopez, F. H. & Munro, P. with S. M. Deo, G. Mauro, & S. Ontiveros. (2019a). *San Lucas Quiaviní Zapotec Talking Dictionary*, version 2.0. Living Tongues Institute for Endangered Languages. Online: http://www.talkingdictionary.org/sanlucasquiavini.

Lillehaugen, B. D., Sánchez Gómez, A. & Hernández Martínez, F. with S. M. Deo, G. Mauro, S. Ontiveros, & K. D. Harrison. (2019b). *San Bartolomé Quialana Zapotec Talking Dictionary*, version 1.0. Living Tongues Institute for Endangered Languages. Online: http://www.talkingdictionary.org/quialana.

Mosel, U. (2011). Lexicography in endangered language communities. In P. Austin & J. Sallabank (eds.) *The Cambridge Handbook of Endangered Languages (Cambridge Handbooks in Language and Linguistics)*. Cambridge: Cambridge University Press, pp. 337–353.

Mugglestone, L. (2011). *Dictionaries: A Very Short Introduction*. Oxford: Oxford Univ. Press.

Munro, P., Lillehaugen, B. D. & Lopez, F. H. (2008). *Cali chiu? A course in Valley Zapotec*, vol. 1–4. Lulu publishing, www.lulu.com.

Munro, P. & Lopez, F. H. with O. V. Mendez, R. Garcia, & M. R. Galant. (1999). *Di'csyonaary x:tee'n Dii'zh Sah Sann Lu'uc (San Lucas Quiaviní Zapotec dictionary / Diccionario zapoteco de San Lucas Quiaviní)*. Los Angeles: UCLA Chicano Studies Research Center Publications.

Sicoli, M. (2011). Agency and ideology in language shift and language maintenance. In T. Granadillo and H. A. Orcutt-Gachiri (eds.) *Ethnographic contributions to the study of endangered languages*. Tucson: University of Arizona Press, pp. 161-176.

Wojtylak, K. (2012). LEXUS for creating lexica Version 3.0. Accessed at https://www.mpi.nl/corpus/manuals/manual-lexus.pdf (15 July 2019)