

**The Documentation and Development of a Spelling System for San Bartolomé Quialana  
Zapotec**

By: Graham Mauro

December 2019

Advised by Dr. Amanda Payne and Dr. Brook Lillehaugen

Haverford College

### **Abstract**

This thesis documents the spelling system created by Aurora Sánchez Gómez for Zapotec (Otomanguan) as spoken in San Bartolomé Quialana, Oaxaca. Beyond documentation, areas of over and underrepresentation of the underlying phonology are identified, and possible solutions are recommended. To my knowledge, there is currently no phonological analysis of the phonology of Quialana Zapotec. Thus, the phonological analysis of the neighboring town of San Lucas Quiaviní is used (Munro 2004). Data is drawn from the Quialana (Lillehaugen 2019b) and Quiaviní Talking Dictionaries (Lillehaugen et al. 2019a). A background and description of the Talking Dictionary project is provided.

## Table of Contents

1. Introduction
2. Talking Dictionaries
  - 2.1. Features of a Typical Entry
3. Zapotec Languages
  - 3.1. Zapotec Writing Systems
4. Tlacolula Valley Zapotec Talking Dictionaries
5. San Bartolomé Quialana Talking Dictionary
  - 5.1. NSF REU Project
    - 5.1.1. My positionality
  - 5.2. Aurora Sánchez Gómez
    - 5.2.1. Spelling System
6. Data
7. Quialana Phonology and Previous Research
8. Note on Quiaviní Orthography
9. Quiaviní Phonology
10. Consonant Representation in Aurora's Spelling System
  - 10.1. Orthographic Development
    - 10.1.1. Perspectives
  - 10.2. Areas for Development and Suggested Changes
    - 10.2.1. [k] as <k>, <c>, and <qu>
    - 10.2.2. Nasals and Laterals
    - 10.2.3. [f]
    - 10.2.4. Retroflex
    - 10.2.5. [x] and [g] as <g> and <gu>
    - 10.2.6. [tʃ] as <tx> and <ch>
    - 10.2.7. [j] as <ll> and <y>
  - 10.3. Summary of Orthographic Developments
11. Further Study
12. Conclusion

## 1. Introduction

Half of all world languages are estimated to fall silent within the next century (Language Conservancy)<sup>1</sup>. Like the loss of biological diversity, the mass disappearance of languages brings with it an endangerment of identity, culture, and ways of life (Harrison 2007).

As our world grows steadily more digital, the ability to write in one's language has become increasingly important. Lillehaugen describes in detail the importance of working collaboratively with communities that want to develop writing systems and how the implementation of said systems on social media platforms can work to combat the grossly false yet routine comment that indigenous languages are ancient and not meant for the modern world (Lillehaugen 2016, 2019).

It is with a belief in the importance of this work that this paper is written. Drawing on my field work alongside Lillehaugen, K. David Harrison, and fellow undergraduate students in the summer of 2019, this paper documents the spelling system created by Aurora Sánchez Gómez which is currently in use by her for Zapotec (Otomanguean) as spoken in her pueblo of San Bartolomé Quialana, Oaxaca. I have also identified areas in which the current spelling system is potentially over or under representative of the language's phonology and provided possible solutions.

1. I would like to thank everyone who made the writing of this thesis possible. In particular, my advisors Dr. Amanda Payne and Dr. Brook Lillehaugen for their encouragement, support, and advice throughout this process. This work would not have been possible without the support of Aurora Sánchez Gómez whose work this thesis is based around and who has inspired and supported me since we first began working together. Thank you also to Dr. Felipe Lopez, K. David Harrison and everyone involved in the Talking Dictionaries Project for creating an opportunity for this work which was supported and funded through NSF REU grant #1461056.

The data throughout this paper is drawn from the San Bartolomé Quialana Zapotec Talking dictionary (Lillehaugen et al. 2019b). Given the centrality of this dictionary to the contents of the paper, the following sections are dedicated to explaining the project.

## **2. Talking Dictionaries**

The Talking Dictionaries are a collection of 120 online, free to access dictionaries which together hold more than 150,00 entries (Harrison et al. 2019). The Dictionaries cover a variety of different languages and are built to be flexible such that they can meet the varying motivations of the different communities they are built for (Harrison et al. 2019).

The first Talking Dictionary went online in 2006 – the Tuvan Talking Dictionary (Harrison & Anderson 2006). This, in turn, was the result of digitizing field recordings and was based on the print dictionary published by the same authors three years earlier (Anderson & Harrison 2003). The Talking Dictionary format is an attempt to remedy many of the shortcomings of print dictionaries. One that Harrison et al. (2019) mention in particular is the common criticism of missing words from any print dictionary. As will be detailed in later sections, the online medium allows content to be mutable and contextualized in a way that a print dictionary cannot be (Harrison et al. 2019). Of particular importance for endangered languages, the Talking Dictionaries are online. This allows easier access for the speech community as well as the diaspora (Harrison et al. 2019). This does rely on the availability and use of internet within these communities and the surrounding areas. Within the context of the Zapotec communities in Oaxaca for which we have Talking Dictionaries, internet is available and widely used – primarily through smartphones. More than just

increasing accessibility, having online resources in indigenous languages can help to combat the damaging ideology that indigenous languages are relics of the past instead of contemporary, living languages (Lillehaugen 2016, 2019).

### 2.1. Features of a Typical Entry (Harrison et al. 2019)

The Talking dictionary format is flexible and each dictionary is unique. That said, there are elements that can be found in every entry. Figure 1 below shows a typical entry in the Talking Dictionaries with callouts for each of these elements. Specifically, this entry comes from the San Bartolomé Quialana Zapotec Talking Dictionary (Lillehaugen et al. 2019b, entry 2392).

**Figure 1. A Typical Entry in the Zapotec Talking Dictionaries**

1 gi toh

2  listen

4 (Spanish) (flor del cerro amarilla usada en la festividad de semana santa y navidad)

5 

3 Speaker: Floriana Hernandez Martinez

6 bookmark

The headword in the language (Quialana Zapotec) is found at 1. Some Dictionaries feature multiple orthographies. All are built to allow for some variation in spelling (Harrison et al. 2019). The callout at 2 highlights the namesake of the Talking Dictionaries. By clicking the ear symbol, a user can listen to a native speaker say the word. That speaker is cited for their contribution at 3. The Zapotec Talking Dictionaries are

trilingual, so the space for glosses (4) features translations into both Spanish and English. To the right of the gloss(es), 5 indicates where media can be found relating to the entry. This media is meant to add context and can be a picture, video, or Tweet. Finally, 6 shows how a user can get a unique URL for that entry.

### 3. Zapotec Languages

Zapotec is a language family within Otomanguean consisting of around 450,000 speakers – primarily in Southern Mexico (Lillehaugen 2019). The Zapotec family consists of a large variety of languages. Of particular relevance to this paper are the Tlacolula Valley (seen in Figure 3) Zapotec languages found in the state of Oaxaca (Figure 2, Google Maps).

**Figure 2. Location of the State of Oaxaca**



Tlacolula Valley Zapotec is within the Central Zapotec branch of the family. There is a fair amount of disagreement about the number of Zapotec languages and how to distinguish between language and dialect. Ethnologue lists over 50 different languages within Zapotec while still classifying all five of the languages shown in Figure 3 as Western

Tlacolula Zapotec (Ethnologue). The decision to classify as either a language or a dialect is not an arbitrary distinction. The choice to use one term or the other may be motivated by the general concept of mutual intelligibility between dialects but not languages. This breaks down in areas such as the Tlacolula Valley where the degree of mutual intelligibility can vary significantly. Beyond this, there is also the question of how the terminology has (and continues to) play a role in the perception and treatment of the speakers and their communities. The use of 'dialect' has routinely been used to downgrade indigenous languages (Zapotec languages included) to a level below majority languages like Spanish (Lillehaugen 2019). In light of this, I will refer to each pueblo's Zapotec as a separate language variety with the acknowledgment that there are varying levels of intelligibility between each pueblo.

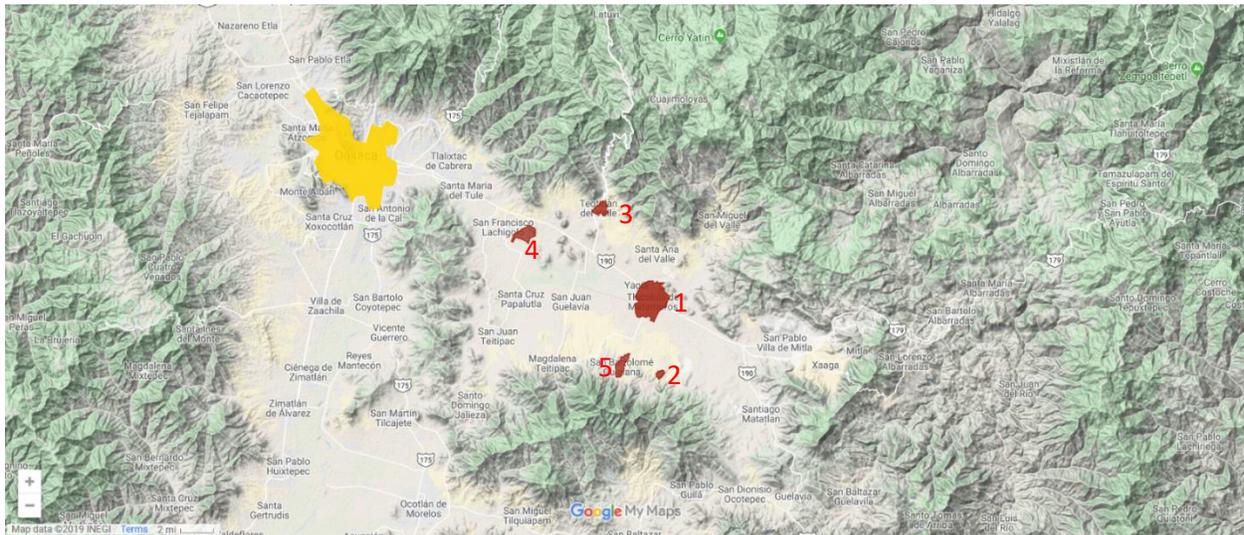
### **3.1. Zapotec Writing Systems (Munro et al. 2007)**

The first Zapotec writing systems were not alphabets, they were pictographic. However, by the time of the conquest by the Spanish and the enforced dominance of the Spanish language, these systems had already fallen into disuse (Urcid 2011). The history of alphabetic representations of Valley Zapotec began in the 1500s with the earliest written document dating back to 1565 – much of the writing in that period was done under religious authorities such as Juan de Córdova, his contemporaries, and the many Zapotec people who worked whether credited or not (Oudijk 2011). While Zapotec is represented in a variety of different ways throughout Oaxaca today, there is no official or unified writing system for Tlacolula Valley Zapotec. Existing systems range from ad hoc spelling systems to full orthographies developed with the assistance of linguists (Munro et al. 2007).

#### 4. Tlacolula Valley Zapotec Talking Dictionaries

Figure 3 (Google Maps) shows a map of the Tlacolula Valley with highlights on the 5 pueblos with Talking Dictionaries (highlighted in red) as well as the capital of the state – Oaxaca de Juárez (highlighted in yellow).

**Figure 3. Tlacolula Valley and Location of Zapotec Talking Dictionaries.**



The Zapotec Talking Dictionaries began in 2013 with Tlacolula de Matamoros (1) (Lillehaugen et al., 2013) and San Lucas Quiavini (2) (Lillehaugen & Lopez et al., 2019, Harrison et al. 2019). Dictionaries in Teotitlan Del Valle (3) (Lillehaugen & Chávez Santiago et al., 2019) and San Jerónimo Tlacoahuaya (4) (Lillehaugen & García Guzmán et al., 2019) brought the number of Zapotec Talking Dictionaries to 4 between 2013 and 2019 (Harrison et al. 2019). The Talking Dictionary of focus in this paper, San Bartolomé Quialana (5) (Lillehaugen et al., 2019b), was started in the summer of 2019.

## 5. San Bartolomé Quialana Zapotec Talking Dictionary

As of writing this, the Quialana Talking Dictionary has over 570 entries including more than 520 audio files and nearly 200 images. The local language activists leading the Quialana Dictionary are Aurora Sánchez Gómez and Floriana Hernández Martínez.

### 5.1. NSF REU Project

My involvement in the Talking Dictionary Project began with being accepted to take part in the NSF Grant #1461056 REU Site (PI K. David Harrison): “Building Digital Tools to Support Endangered Languages and Preserve Environmental Knowledge in Mexico, Micronesia, and Navajo Nation.” As part of this grant, the undergraduate students were taught the basics of ethnographic linguistics and how to work on and add to the Talking Dictionaries. In the summer of 2019, we traveled with Dr. Lillehaugen and Dr. Felipe Lopez (a language activist, poet, professor, and co-author and leader of the San Lucas Quiaviní Talking Dictionary) to Oaxaca where we stayed for a month expanding the Zapotec Talking Dictionaries along with our co-authors.

While all of us worked throughout the four different pueblos, each of us was assigned 1-2 languages to spend the majority of our time working on. I focused on San Lucas Quiaviní and San Bartolomé Quialana Zapotec. Between the two, most of my time was spent with Quialana Zapotec working with Aurora. While I was there to work on whatever aspect of the Talking Dictionary Aurora wished to focus on, it is undoubtedly true that I also brought my own ideas and perspectives into that work as well the work done throughout this paper.

### **5.1.1. My Positionality**

It is important to note that, in this work, everyone has a positionality and their own motivations. Particularly when your work is inseparable to some degree from matters of identity and culture, those motivations should be explicitly discussed.

My primary motivation behind writing this paper is twofold. First, I want to produce something that is of use to Aurora and can help her continue her revitalization work. Second, I seek produce a thesis which makes concrete contributions based on original research in accordance with linguistic theory. To balance these two, my orthographic recommendations generally include a discussion of multiple possible solutions before coming down on the side of a single answer.

The use of ‘recommendation’ is intentional. This is not my writing system nor is it my language or culture. As such, I can make recommendations, but the decisions are ultimately for the community to make. As a co-author of the Quialana Talking Dictionary, a member of that community, and the person I have primarily worked with, my work will be presented directly to Aurora.

### **5.2. Aurora Sánchez Gómez**

Aurora has a degree in tourism management and is an activist pursuing the documentation and revitalization of her mother tongue – San Bartolomé Quialana Zapotec. She is the director of the Casa Cultura in Quialana and one of two local co-authors and dictionary leaders for the Quialana Talking Dictionary (Personal Communication Nov. 2019).

She and I have been in communication throughout the writing of this paper and I am immensely thankful for her support and assistance.

Aurora often talks about a desire to teach more children to speak Zapotec and sees the lack of interest and command of the language amongst children as a major danger to her language. Many of the children that can speak Zapotec in Quialana do so by heavily relying on Spanish borrowings – another point of contention for Aurora (Personal Communication 2019).

Once I became more familiar with Quialana Zapotec, a typical workflow for Aurora and me would consist of her writing a word in her spelling system and then asking me and other students to say the word back to her. She said that if we managed to pronounce it correctly, she knew her spelling was right. If we failed to say it correctly, she would try different spelling variations until we pronounced it as she wanted. This reflects one of her major goals. That is, to have an orthography that someone could use to correctly pronounce the word without further instruction. She wants this both for teaching the language today, as well as preparing for the possibility that there may one day be no native speakers of Quialana Zapotec (Personal Communication 2019).

### **5.2.1. Spelling System Development**

When we first arrived in Quialana in the summer of 2019, we asked Aurora whether people in her pueblo wrote in Zapotec. She told us they did not. The next day, we returned to record words for the dictionary. To guide our work, Aurora made a word list. In Spanish and Zapotec. It was then that we realized that while in general people in her pueblo do not

write in Zapotec, Aurora does. It is important to note that we later learned that a number of other community members also write in Zapotec. Many of these speakers would ask Aurora for advice in writing when creating word lists for elicitation.

Aurora remembers facing the severe discrimination that many indigenous people faced in the school systems. She was told that she would never be able to learn to read and write in Spanish unless she stopped speaking Zapotec. Aurora wanted to prove them wrong. With the support of her parents and grandparents, she began to write in Zapotec (Personal Communication Nov. 2019).

Her spelling system is heavily influenced by Spanish orthography (e.g. complimentary distribution between <c> and <qu> graphemes according to Spanish orthographic rules) and nearby Zapotec writing systems that are being developed (e.g. the use of *i* to represent the mid high vowel). The system is highly complex and features marking of different glottalizations. The recommendations of this paper will focus on places within the spelling system where there is inconsistency in representation due to over or underrepresentation of the phoneme inventory.

## 6. Data

Unless otherwise specified, all Zapotec language data used throughout this paper will be drawn from the San Bartolomé Quialana (Lillehaugen et al. 2019b, <https://talkingdictionary.swarthmore.edu/quialana/>) and San Lucas Quiavini (Lillehaugen et al. 2019a, <http://talkingdictionary.swarthmore.edu/sanlucasquiavini/>) Talking Dictionaries. Full citations for these dictionaries as well as the three other Zapotec Talking Dictionaries

(Teotitlán - Lillehaugen and Chávez Santiago et al. 2019, Tlacoahuaya – Lillehaugen and García Guzmán et al. 2019 and Tlacolula – Lillehaugen et al. 2013) can be found in the references section. The reader is encouraged to follow the provided links to both better understand and further explore the language.

Throughout this paper *italics* will be used for words in Zapotec, while single quotes ('word') will be used for the English or Spanish glosses. Graphemes will appear as <c> while their associated phoneme will appear as /k/.

## 7. Quialana Phonology and Previous Research

As of the writing of this paper, a phonological analysis of Quialana Zapotec has not been published nor has any linguistic work on Quialana Zapotec. While the general lack of research does provide an exciting opportunity to make contributions to both the community and to the field, the lack of a full phonological analysis is a challenge to the present work. Thankfully, there is a relatively large body of research on San Lucas Quiaviní Zapotec – including a full documentation of the phonology (Munro and Lopez, et al. 1999). It will be assumed that the Quiaviní phonology closely reflects Quialana phonology and any analysis done using phonology will consider the Quialana and Quiaviní phoneme inventories the same.

This is a reasonable assumption for a number of reasons. Initially, Ethnologue classifies the two as a single language (iso code Zap) (Ethnologue). The two pueblos border one another (See Figure 3) and are less than four miles apart. There is frequent interaction between the two pueblos contemporarily and historically both due to proximity and

through shared interactions in markets (Lopez personal communication). Further, this assumption is consistent with the work we have done in Quialana. While it is only anecdotal without a formal study, we did not notice any missing phonemes and I've noted only one additional phoneme (detailed in section 9). Rather, the phonologies seemed to differ primarily through varying frequencies of usage for certain phonemes and phonations.

### **8. A Note on Quiaviní Orthography**

There are currently two orthographies used in the Quiaviní Talking Dictionary. The first was introduced in the original dictionary of Quiaviní Zapotec (Munro and Lopez et al. 1999). This orthography is highly detailed and includes graphemes for every consonant, vowel, and phonation type. Both breathiness and creakiness can also be expressed in two levels of intensity by reduplicating the marking (a grave accent over a vowel for creaky and a <h> following a vowel for breathiness). The second orthography was developed from the 1999 orthography in an attempt to make writing and reading more accessible (Munro et al. 2007). It is much quicker to learn but underrepresents the phonological contrasts. Importantly, it does not distinguish tone, breathiness, or creakiness. When featuring Quiaviní Zapotec, this paper uses the 1999 orthography whenever it is available. This is done for two reasons. First, the 1999 Quiaviní Zapotec orthography more closely resembles Aurora's spelling system. Second, when Quiaviní examples are used, it will often be to show the differences between the two languages which are made more evident by the more detailed orthography. A full explanation of the 1999 Quiaviní orthography is not given in this paper. All examples are explained such that it is not necessary. If interested, refer to the

original dictionary (Munro and Lopez et al. 1999) for the 1999 orthography and (Munro et al. 2007) for the modified orthography.

### 9. Quiaviní Phonology (Munro 2004)

The Quiaviní phonology that is used throughout this paper is from the work of Munro and her colleagues, specifically drawing from her summary of the language in 2004. The original print dictionary (Munro and Lopez et al. 1999) will also be consulted when additional data is needed. This paper does not attempt to fully document the phonology of Quialana Zapotec. Thus, any phonological claims presented are based in qualitative observations formed through my work in Quialana and documented in the Talking Dictionaries when possible. Table 1 below shows the consonant inventory presented by Munro.

**Table 1. Quiaviní Consonant Inventory** (adapted from Munro 2004)

	Labial	Dental-Alveolar	Alveolar-Palatal	Retroflex	Velar
Fortis Stop	<p>	<t>			<k>
Lenis Stop	<b>	<d>			<g>
Fortis Affricate		<ts>	<ch>		
Fortis Fricative	<f>	<s>	<x>	<x:>	<j>
Lenis Fricative		<z>	<zh>	<zh:>	
Fortis Nasal	<mm>	<nn>			<nng>
Lenis Nasal	<m>	<n>			<ng>
Fortis Lateral		<ll>			
Lenis Lateral		<l>			
Flap		<r>			
Trill		<rr>			
Glide	<w>		<y>		

Quiaviní Zapotec does not have a voicing distinction. Instead, it has a fortis/lenis distinction for sonorants and obstruents (Munro 2004). In addition to the above consonants, Quialana Zapotec features post-alveolar affricates [ɟʒ]. Table 2 provides various examples of the use of [ɟʒ] (represented in Aurora’s spelling system as <dx>) in the Quialana Talking Dictionary.

**Table 2. Uses of [ɟʒ] in the Quialana Talking Dictionary.**

Quialana Zapotec Headword	English Gloss	Talking Dictionary Link
<i>dxíidx</i>	‘pineapple’	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4820">https://talkingdictionary.swarthmore.edu/quialana/?entry=4820</a>
<i>a udxá</i>	‘full’	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5024">https://talkingdictionary.swarthmore.edu/quialana/?entry=5024</a>
<i>ndxáab</i>	‘bad’	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5048">https://talkingdictionary.swarthmore.edu/quialana/?entry=5048</a>

While this is absent in Quiaviní Zapotec, Lillehaugen has documented the use of [ɟʒ] in Tlacolula Zapotec (Lillehaugen 2006 p.16-17).

Quiaviní Zapotec has six vowel which combine into ten diphthongs. Additionally, vowels can hold one of four phonation types (creaky voice, breathy voice, checked, or modal) (Munro 2004). Tale 3 shows each of the six vowels and diphthongs.

**Table 3. Quiaviní Vowel Inventory** (Adapted from Munro 2004)

Vowel	Diphthongs beginning with this vowel
[a]	[ai] [au]
[e]	[ei] [eu]
[i]	[ia] [ie] [iu]
[o]	
[u]	[ua] [ue]
[ɨ]	[ɨi]

Quiaviní Zapotec vowels can also take four tones (high, low, rising, and falling) (Munro 2004). Munro describes these tones as being predictable based on the phonation pattern and specifies 27 different phonation patterns and the tone associated with each pattern. Vowels can be more or less creaky/breathy in comparison to other vowels of the same phonation. For example, in Quiaviní Zapotec the word for air *bihih* (Lillehaugen et al. 2019, entry 378) is immensely breathy (and an example of sound symbolism) in comparison to the word for good/fine *gweenahg* (Lillehaugen et al. 2019, entry 2745).

The lenis palatal-alveolar affricate [dʒ] is used where Quiaviní Zapotec would use the lenis palatal-alveolar fricative [ʒ]. Table 4 shows two examples of near minimal pairs between the two languages showing the usage of the affricate in Quialana. While this correspondence is common throughout the dictionary, it is not universal. Table 4 also shows two examples of the use of the fricative in Quialana Zapotec. A more specific analysis would have to be done to

determine whether or not it occurs in specific environments – the below pairs suggest that [dʒ] may be used post vocally while [ʒ] remains in other environments.

**Table 4. Examples of /ʒ/ and /dʒ/ usage in Quialana Zapotec**

Quiavini Zapotec	Quialana Zapotec
<i>a wzhya</i> <sup>1</sup> 'to be full'	<i>a udxá</i> <sup>2</sup> 'full'
<i>dii'zh</i> <sup>3</sup>  'language, speech, word'	<i>diidx</i> <sup>4</sup> 'word'
<i>zh:u'ub</i> <sup>5</sup> 'corn'	<i>zhuub</i> <sup>6</sup> 'corn'
<i>rzhiéz</i> <sup>7</sup> 'laugh, smile'	<i>rzhiu</i> <sup>8</sup> 'useful'

1. <http://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=3746> , 2. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5024>

3. <http://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=102> , 4. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5142> ,

5. <http://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=95> , 6. <https://talkingdictionary.swarthmore.edu/quialana/?entry=4771> ,

7. <http://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=525> , 8. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5035>

The vowel systems show substantially more difference than the consonant inventories.

Based on my own observations working with Aurora, Quialana Zapotec tends to use less breathy and creaky voice, occasionally using checked vowels instead. The example used previously to show extreme breathiness in Quiavini (*bihih*) can be said with a modal vowel in Quialana Zapotec (Aurora Personal Communication 2019). It is worth noting that this is not true across all speakers and some Quialana Zapotec speakers may still use breathy voice.

Nonetheless, this difference led Aurora to often joke that Zapotec speakers from Quiavini, “speak with so much air” (Personal Communication 2019). Such a comment also provides further proof that the two pueblos have frequent interaction which has led to reflection and comparison of the two languages

The vowel differences are substantial enough (and the system itself complex enough) that it is beyond the scope of this paper to suggest improvements in the spelling system’s representation of vowels. Thus, this paper concerns itself only with recommendations for the representation of consonants.

**10. Consonant Representation in Aurora’s Spelling System**

As aforementioned, due to the lack of a phonologic analysis for Quialana Zapotec, it will be assumed that the consonant inventories of Quiavini (Munro 2004) and Quialana are equivalent. Each consonant will be lined up with the grapheme(s) in Aurora’s spelling system which represent(s) it as well as a word example and the link to that entry in the Quialana Talking Dictionary. Table 5 shows the current spelling system in two different formats.

**Table 5. Consonant Inventory and Associated Graphemes**

	Labial	Dental-Alveolar	Alveolar-Palatal	Retroflex	Velar
Fortis Stop	<p>	<t>			<c>,<qu>,<k>
Lenis Stop	<b>	<d>			<g>,<gu>
Fortis Affricate		<ts>	<tx>,<ch>		
Lenis Affricate			<dx>		
Fortis Fricative	*	<s>	<x>	<x>	<g>,<gu>
Lenis Fricative		<z>	<zh>	<zh>	
Fortis Nasal	<m>	<n>			<n>
Lenis Nasal	<m>	<n>			<n>
Fortis Lateral		<l>			
Lenis Lateral		<l>			
Flap		<r>			
Trill		<rr>			
Glide	<w>		<ll>,<y>		

Phone	ASG Grapheme	Example	Talking Dictionary Entry Link	Notes
p	p	<i>puraad</i> 'fast'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5023">https://talkingdictionary.swarthmore.edu/quialana/?entry=5023</a>	
b	b	<i>bich</i> 'cat'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4956">https://talkingdictionary.swarthmore.edu/quialana/?entry=4956</a>	
t	t	<i>tap</i> 'four'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4908">https://talkingdictionary.swarthmore.edu/quialana/?entry=4908</a>	
d	d	<i>dam</i> 'owl'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4937">https://talkingdictionary.swarthmore.edu/quialana/?entry=4937</a>	
k	k, c, qu	<i>kiit</i> 'to play'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5155">https://talkingdictionary.swarthmore.edu/quialana/?entry=5155</a>	
g	g, gu	<i>ga</i> 'nine'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4925">https://talkingdictionary.swarthmore.edu/quialana/?entry=4925</a>	
m (fortis and lenis)	m	<i>maay</i> 'animal'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=3907">https://talkingdictionary.swarthmore.edu/quialana/?entry=3907</a>	
n, ŋ, (fortis and lenis)	n	<i>náax</i> 'chocolate'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=2581">https://talkingdictionary.swarthmore.edu/quialana/?entry=2581</a>	
r	r	<i>rúub</i> 'basket'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4999">https://talkingdictionary.swarthmore.edu/quialana/?entry=4999</a>	
r	rr	<i>burre</i> 'donkey'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4853">https://talkingdictionary.swarthmore.edu/quialana/?entry=4853</a>	Spanish borrowing
f	*			See note below
s	s	<i>sáa</i> 'to walk'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5099">https://talkingdictionary.swarthmore.edu/quialana/?entry=5099</a>	
z	z	<i>zit</i> 'distant'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5021">https://talkingdictionary.swarthmore.edu/quialana/?entry=5021</a>	
ʃ, ʂ	x	<i>xa</i> 'what'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5167">https://talkingdictionary.swarthmore.edu/quialana/?entry=5167</a>	
ʒ, ʝ	zh	<i>zhít</i> 'cat'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4955">https://talkingdictionary.swarthmore.edu/quialana/?entry=4955</a>	
x	g	<i>gity</i> 'squash'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4751">https://talkingdictionary.swarthmore.edu/quialana/?entry=4751</a>	
l (fortis and lenis)	l	<i>laas</i> 'laas'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5010">https://talkingdictionary.swarthmore.edu/quialana/?entry=5010</a>	
ʈ	ʈs	<i>ʈsi</i> 'ten'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4909">https://talkingdictionary.swarthmore.edu/quialana/?entry=4909</a>	
tʃ	tx, ch	<i>txa</i> 'a, an, one'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=3276">https://talkingdictionary.swarthmore.edu/quialana/?entry=3276</a>	
ɖʒ	dx	<i>dxambeu</i> 'mood god'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4918">https://talkingdictionary.swarthmore.edu/quialana/?entry=4918</a>	Not in Munro 2004
j	ll, y	<i>lláx</i> 'avocado'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5001">https://talkingdictionary.swarthmore.edu/quialana/?entry=5001</a>	
w	w	<i>wí</i> 'a type of plant'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4823">https://talkingdictionary.swarthmore.edu/quialana/?entry=4823</a>	When not word final

\* As of the writing of this paper, the Quialana Talking Dictionary does not have any Spanish borrowings featuring an [f] nor have I definitively encountered one while working with Aurora.

## **10.1. Orthographic Development**

The following sections address the areas of the consonant orthography which I believe could be developed further to make the orthography more effectively accomplish Aurora's goals. That statement should cause some level of discomfort. There is no universal standard for a perfect orthography nor is there even a single concept of what it means to improve an orthography. Thus, before individual areas of the Aurora's spelling system can be examined, the broader discussion of what it means to develop orthographies must be had.

### **10.1.1. Perspectives**

Work such as that done by Hull (2017) displays how questions of orthographies cannot be separated from their context and how competing orthographic ideas are often reflective of competing social and/or academic interests (Hull 2017).

Cahill (2011), Smalley (1964), and Casquite and Young (2017) all attempt to remedy this by providing higher level guiding metrics for orthographic development. Smalley (1964) posits maxims of: Acceptability (the community accepts and wants to use the orthography), Ease of Learning (given the linguistic background of the community, it is easy for members to learn), Representation (the orthography accurately represents the phonology of the language), Ease of Reproduction (current writing technology can be easily used to write in the orthography), and Ease of Transfer (knowledge of the other languages speakers are likely to know aids in the use of the orthography). Essentially, Smalley (1964) argues that an orthography should reflect the linguistic features of a language while being easy to use and teach given the broader linguistic background of the speech community (Smalley 1964).

In a short overview created for the Symposium on Developing Orthographies for Unwritten Languages, Cahill (2011) further expands these ideas by defining an effective orthography as satisfying the phonology, meeting the interests of all invested groups, and being realistically implementable. Cahill focuses on explaining how satisfying this trifecta is often a balancing act and how, in particular, social and cultural factors can compete to pull the language in different directions (Cahill 2011). Relevant to the current discussion, he points out the common use of <qu> or <c> to represent [k] in many Latin American orthographies to more closely resemble the orthography of Spanish. He provides a contrasting example with Kiche. In this case, the community more recently wanted to use <k> specifically because it contrasted with the Spanish Orthography. This reflects a central tension between the desire to establish a unique identity free from the influence of colonizing languages and the desire to create orthographies that can be easily learned by people who are often native speakers of the colonizing language (Cahill 2011). This leads to an important point. Particularly in the case of endangered languages, it is important for outside linguists to acknowledge their positionality and ensure that their work is done in collaboration with members of the speech community. We can make suggestions and recommendations, but it is ultimately the community's decision and linguists should be willing to work within that framework even if it means using an orthography that is not ideal for linguistic research.

Casquite and Young (2017) build on this idea by detailing a number of ways for speech communities to work collaboratively to develop their own orthography in conjunction with advice from linguists. They conclude that local co-author involvement and agency in

orthographic decisions is an important means of empowerment and revitalization (Casquite and Young 2017).

With these works in mind, the following sub-sections identify areas of Aurora's spelling system where there is overrepresentation or underrepresentation of the phonology. These are areas that should be considered for revision moving forward. Solutions will be suggested based on Aurora's goal to build an orthography that a non-Zapotec speaker could use to accurately produce Quialana Zapotec words (Personal Communication Nov. 2019). In its purest form, this goal is extremely difficult (if not impossible) to accomplish. The orthography would need to accurately and fully distinguish between every allophone of the language in what would more or less become a phonetic transcription. This would be unwieldy and difficult to learn. However, if the orthography does not represent allophones but instead focuses only on representing contrastive differences, it may be concise enough to easily learn while still accomplishing the heart of Aurora's goal. Non-speakers would not be able to perfectly reproduce the Zapotec word without further instruction, but for the purposes of this paper it will be considered the best compromise between representation and learnability.

## **10.2. Areas for Development and Suggested Changes**

### **10.2.1. [k] as <k>,<c>,<qu>**

This is perhaps the most complicated area in the consonant orthography as it not only involves three separate graphemes all used to represent the same phoneme, but also involves the question of distance from the Spanish orthography. The use of <k> is one of Aurora's most recent additions to the orthography and is therefore somewhat rare in the Talking Dictionary –

occurring only 12 times (in comparison to almost 100 entries featuring <c>). <qu> is also found only 12 times, but this may be due to the restricted environment. <c> and <qu> are used complementarily as one would expect from Spanish. That is, <qu> is used preceding <i>, <i> and <e>. <c> is used in all other environments. <k> is used across both environments. Table 6 below shows a number of usages of these graphemes in comparable environments.

**Table 6. Representations of /k/**

Use of <c>	Use of <qu>	Use of <k>
	nquit <sup>1</sup> ‘white’	nkits <sup>2</sup> ‘white’
	quitxiracti’i <sup>3</sup> ‘healthy’	xkit <sup>4</sup> ‘robin’
Cali chiu <sup>5</sup> ‘where are you going?’		kalainée sáraab <sup>6</sup> ‘they are in love’
Zac laizhi <sup>7</sup> ‘good day (midday)’		rcuenk <sup>8</sup> ‘contractions’

1. <https://talkingdictionary.swarthmore.edu/quialana/?entry=4857>, 2. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5242>, 3. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5027>, 4. <https://talkingdictionary.swarthmore.edu/quialana/?entry=4930>, 5. <https://talkingdictionary.swarthmore.edu/quialana/?entry=2388>, 6. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5223>, 7. <https://talkingdictionary.swarthmore.edu/quialana/?entry=9>, 8. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5227>

My recommendation is to move to only using <k>. Simplifying the representation of [k] will make the orthography simpler and thus easier to learn. Using only <qu> would be an odd choice as it would expand upon a grapheme used only in fairly limited contexts and could be confusing to native Spanish speakers. Choosing <c> could similarly cause issues as the Spanish grapheme alternation is motivated by <c> representing [s] when preceding <i> and <e>. The disadvantages of using <k> are twofold deriving from the movement away from the Spanish orthography. First, it is a grapheme not found in Spanish and thus may have less familiarity to L1 Spanish speakers (which many children in Quialana are). While this could make acquisition

slightly more difficult, it is a simple rule that could be quickly taught and learned regardless of linguistic background. However, we do not want to oversaturate the orthography with unfamiliar graphemes. Second, it creates more distance from Spanish. This is not necessarily a problem and could even be an advantage depending on the perspective of the community (Cahill 2011). Given that Aurora introduced the use of <k> herself, I will assume that this is not a disqualifying factor. As with all of these recommendations, this is a judgement left for Aurora and the rest of her community.

Finally, this does leave the question of how to handle Spanish borrowings which may use <c> or <qu>. If Aurora wishes to create maximum distance from the Spanish orthography, the orthography could insist on re-writing these graphemes as <k> when borrowed. Otherwise, the borrowings could be left as they are and the use of <c> and <qu> would become a signal that the word has been borrowed. This allowance is preferable particularly given the case of Spanish words of the form [k]V such as *que* ('what') which would likely be very uncomfortable for native Spanish speakers to write as *ke*. Allowing the use of <c> and <qu> in borrowings avoids this discomfort.

### **10.2.2. Nasals and Laterals**

There are currently two graphemes representing nasal phonemes in the orthography: <n> and <m>. They represent a predicted six phonemes [n, m, ŋ] and their fortis counterparts. In the Quiaviní 1999 orthography there is a distinction made between fortis and lenis nasals as they are contrastive. There are currently no minimal pairs/triplets in the Quialana Talking Dictionary that would clearly indicate the nasals as contrastive. However, returning to the

assumption that the two phonologies are equivalent, it can be assumed that nasals in Quialana Zapotec are contrastive and thus must have some form of orthographic differentiation.

This same analysis can be applied to <l> representing lenis [l] and fortis [l̪]. I would recommend following the example of the 1999 Quiaviní Orthography and using duplication of the grapheme. This doesn't require the learning of a distinct grapheme and could easily be taught (if it is not already intuitive). In this sense, the fortis/lenis distinction could be made orthographically even if they are allophonic without too much concern for falling intelligibility or acquisition.

### 10.2.3. [f]

The only attested uses of [f] in the Quiaviní Talking Dictionary are from Spanish borrowings (Munro & Lopez et al. 1999)– some of which are listed below.

**Table 7. Selected uses of /f/ in the Quiaviní Talking Dictionary**

Quiaviní Zapotec	Spanish	English	Talking Dictionary Entry Link
<i>fa'alld</i>	'falda'	'skirt'	<a href="https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=546">https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=546</a>
<i>fá'sihlly</i>	'fácil'	'easy'	<a href="https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=1992">https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=1992</a>
<i>fot</i>	'foto'	'photo'	<a href="https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=457">https://talkingdictionary.swarthmore.edu/sanlucasquiavini/?entry=457</a>

As of now, the Quialana Talking Dictionary does not contain any Spanish borrowings with [f] and thus it currently does not attest to any grapheme. Given that this is a phoneme only featured in borrowings, it is reasonable that its representation should be borrowed as well. While there may be a desire of the community to not include clear borrowings at all, if they choose to, I recommend the use of <f>.

The inclusion of dominant language borrowings is often a point of discussion and returns to the question of distance from Spanish and its orthography. If the community wishes to not include any borrowings in the Quialana Talking Dictionary, there will be no words using [f] and thus it would seem unnecessary to assign a grapheme to it. However, outside of the context of the Talking Dictionary, there would still be pressure to represent [f]. Minimally, this would manifest in the need to write Spanish names (such as *Flori*). If speakers are willing to represent it, I would expect that they are willing to use <f> as there is no Zapotec equivalent.

#### **10.2.4. Retroflex**

Both alveolar-palatal fricatives in the consonant inventory ([ʃ] and [ʒ]) are accompanied by the corresponding retroflexes ([ʂ] and [ʐ] respectively). The retroflexes are both represented in the 1999 Quiaviní orthography as distinct from their non-retroflexed partners. Assuming they are distinct phonemes in Quialana and must have a distinct representation, they present a difficult problem.

Given that it is the first instinct when many people begin to transcribe the retroflexes, the first recommendation is to represent them with the grapheme for the non-retroflex fricative following an <r>. This could generally be considered a naïve transcription as the

retroflex fricatives are often mis-transcribed as a [r] followed by a non-retroflex fricative. Using this representation bears the distinct advantage of being intuitive for those learning the language and coming from a linguistic background which lacks retroflexed consonants. By representing the retroflex fricatives as digraphs, the orthography would be perpetuating the idea that it is two sounds (a retroflex followed by a fricative) instead of one. In terms of the necessary usability of the orthography, this is not too severe of a problem and may be of little concern to the community - particularly given that digraphs are commonly used in orthographies (Aurora's current system included). This option is preferable to the digraph below that has already been rejected by speakers in Quiaviní as seen through the differences between the 1999 and 2007 Quiaviní Orthographies (Munro and Lopez et al. 1999, 2007).

That option is to follow the 1999 Quiaviní Orthography and use a <:> following the fricative. This is the mirror of the previous option. No other part of the orthography uses <:> so it cannot introduce ambiguity. Once a reader has learned the rule, there would be no doubt as to whether or not the fricative is retroflex. However, that clarity exists because it is highly unlikely that the reader would have seen <:> used as anything other than a punctuation marker. Thus, it is clear but likely to cause discomfort to speakers. If the use of a colon were particularly undesired, another diacritic could be used to the same effect with the same advantages and disadvantages.

Using another familiar grapheme would be confusing to a user (who likely expects to use <r> to represent a retroflexed sound if they have any intuition given the previously mentioned mis-transcriptions) and introduce the same ambiguity.

However, the use of <rz> and <rx> does introduce a possible ambiguity as to whether it is a retroflexed fricative or a retroflexed tap followed by a non-retroflexed fricative. An example of this is *rzhiu* ‘useful’ (Lillehaugen 2019b, Entry 5035) which is a tap followed by a non-retroflexed fricative. This becomes more troubling with the possibility of a retroflex fricative following a trill and leading to a cluster such as: <rrrz>. Seeking to avoid this, my final suggestion for representing retroflex fricatives is the use of a superscript <r̥> to create <x̥> and <zh̥>. This prevents any ambiguity. Unfortunately, it does so in a similar way as the use of a colon – by making use of a grapheme that is otherwise unused in the orthography. As a balance, the superscript <r̥> would likely be more familiar and intuitive to speakers than a colon. The most pressing disadvantage of this option is the difficulty of using it digitally. The <r̥> is not natively available on Spanish or English keyboards. While it would not be difficult to create a digital keyboard that gave this option, it would nonetheless create a barrier to usage.

This is not a clear-cut question, but I would recommend the use of <r̥> following the grapheme for the fricative. The use of a colon was rejected by speakers and led to the simplified orthography of Quiaviní Zapotec not marking retroflex (Munro et al. 2007). In keeping with Aurora’s desire for a highly representative orthography, some level of representation is needed. While the use of <r> before the fricative does allow more immediate digital accessibility, the two or three <r> clusters it allows create a more permanent drawback.

#### **10.2.5. [x] and [g] as <g> and <gu>**

This is the only example in which a single grapheme represents more than one phone. The Quiaviní orthography does not distinguish between the two phones as they are not

contrastive. Rather, both are allophones of /g/. Therefore, it is possible to leave these as allophones represented by the same graphemes.

If this is the preferred course of action, I would recommend keeping with <g> and <gu> while using them in accordance with Spanish orthographic rules. It is also possible to use only <g> as I already recommended the removal of <qu>. The important difference here is that <g> would still have implications to a L1 Spanish speaker and may feel misleading unlike the previously recommended <k>.

Distinguishing between the sounds is a more difficult task. The only way would be to introduce a second grapheme. There is sense that using <g> helps orient the speaker to know that it is some type of a velar sound (as most users would be Spanish speakers familiar with the use of <g> to represent a velar voice plosive). Because of this, I would be hesitant to abandon <g> and would instead opt for adding a second grapheme. In order to try to keep a sense of consistency in the orthography, I would recommend using <gh> to represent [x]. This draws inspiration from the use of <zh> to represent [ʒ]. Within <zh> we see the use of <h> following an already present grapheme to indicate a different phoneme. By choosing [x] over [g] to bear this extra <h>, there is also some level of consistency in allowing fricatives to be represented by two graphemes while all plosives are represented by a single one. To avoid unnecessary complexity, I would recommend leaving them both to be represented by <g> and <gu> unless it is shown that they are contrastive in Quialana Zapotec.

### 10.2.6. [tʃ] as <tx> and <ch>

This is a fairly straightforward case of two graphemes <tx> and <ch> being used interchangeably to represent a single affricate [tʃ]. The central tension in choosing between the two representations lies within the question of distance from the Spanish orthography. Table 8 shows two examples differing in grapheme usage.

**Table 8. Representations of [tʃ]**

Uses of <ch>	Uses of <tx>
xchíí <sup>1</sup> ‘cena’	kali txigibid <sup>2</sup> ‘where are you all going?’
Cháan <sup>3</sup> ‘(greeting to an elderly person)’	txa <sup>4</sup> ‘a, an, one’

1. <https://talkingdictionary.swarthmore.edu/quialana/?entry=47942>, <https://talkingdictionary.swarthmore.edu/quialana/?entry=2387>, 3. <https://talkingdictionary.swarthmore.edu/quialana/?entry=2396>, 4. <https://talkingdictionary.swarthmore.edu/quialana/?entry=3276>

Using <ch> would likely make acquisition easier for Spanish and English speakers as it would be a familiar grapheme lining up with the expected phoneme. For better or worse, this would make the orthography more closely resemble Spanish and provide more consistency cross-linguistically. This would also appear familiar to speakers of Teotitlán and Tlacolula Zapotec – both of which use <ch> to represent [tʃ] (Lillehaugen 2006 p. 16-17, Lillehaugen and Chávez Santiago et al. 2019).

In contrast, <tx> would provide more consistency intra-lingually. The voicing pair to [tʃ] ([dʒ]) is represented with <dx> – the voicing pair of [t] followed by an <x>. This is logical when focusing within the orthography. It provides a consistent rule in which the addition of an <x> after a plosive is used to represent the nearest affricate of the same voicing. Further, the combination of a plosive (<t> representing [t]) followed by a fricative (<x> representing [ʃ])

logically leads to an affricate which contains elements of both manners of articulation. That said, it's unlikely that any non-linguists would find this helpful. Finally, this would make all affricates represented by a <t> (or a <d>) followed by a fricative (<s> or <x>). This would further provide internal consistency to the orthography. The only disadvantage is that this representation would be unfamiliar to a Spanish speaker and would distance the orthography from the Spanish orthography. The familiarity/acquisition issue I believe is more than balanced out by the orthographic consistency provided. It thus is a question of allowable distance from Spanish. If this is not a concern, I would recommend the exclusive use of <tx>.

**10.2.7. [j] as <ll> and <y>**

Finally, there is the question of representing [j]. Currently both <ll> and <y> are used interchangeably (or at very least they are not consistently used in distinct environments). Table 9 shows their uses in two selected examples.

**Table 9. Representations of [j]**

Use of <ll>	Use of <y>
llab <sup>1</sup> 'to fall'	yabdu'u <sup>2</sup> 'a type of plant'
llú'u <sup>3</sup> 'house'	yú'u <sup>4</sup> 'to enter'

1. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5146>, 2. <https://talkingdictionary.swarthmore.edu/quialana/?entry=4791>, 3. <https://talkingdictionary.swarthmore.edu/quialana/?entry=4987>, 4. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5104>

Both grapheme representations can be found in Spanish although they are distributed complementarily in the Spanish orthography. This means that the familiarity/acquisition argument could go either way as each grapheme would seem familiar in some contexts and

uncomfortable but pronounceable in others. Thus, the primary focus of this decision can be focused upon how each representation acts within the orthography.

Using only <ll> has clear disadvantages. First, for simplicities sake, it is preferable to avoid using digraphs when a single grapheme could suffice. This helps avoid confusion around whether they are separate phonemes. In this case, there could be some confusion about whether it is a distinct sound or a re-articulated or lengthened <l>. There is also the possibility of needing to represent a fortis [l]. This issue was raised in Section 5.2.2. The suggested grapheme for this task was <ll>. Therefore, using <ll> to also represent [j] could introduce another area of ambiguity. Unless it is remarkably more comfortable to use <ll> out of context, it would seem that the only advantage <ll> has over <y> is that it does not have the disadvantages of <y>.

Those disadvantages can be boiled down to one essential point, <y> is already used in the orthography to represent a voiceless word final vowel. While this is a potential concern, the restricted environment and the fact that they would exist in complementary distribution make this concern relatively minor compared to those associated with the use of <ll>. It is because of this that I would recommend the use of <y> to represent [j].

### **10.3. Summary of Orthographic Suggestions**

The table below shows how the orthography would represent each consonant if all of the suggestions made throughout section 5.2 were to be adopted.

**Table 10. Consonant Inventory and Modified Associated Graphemes.**

	Labial	Dental- Alveolar	Alveolar- Palatal	Retroflex	Velar
Fortis Stop	<p>	<t>			<k>
Lenis Stop	<b>	<d>			<g>,<gu>
Fortis Affricate		<ts>	<tx>		
Lenis Affricate			<dx>		
Fortis Fricative	<f>	<s>	<x>	<xʳ>	<g>,<gu>
Lenis Fricative		<z>	<zh>	<zhʳ>	
Fortis Nasal	<mm>	<nn>			<nng>
Lenis Nasal	<m>	<n>			<ng>
Fortis Lateral		<ll>			
Lenis Lateral		<l>			
Flap		<r>			
Trill		<rr>			
Glide	<w>		<y>		

Phone	ASG Grapheme	Example	Talking Dictionary Entry Link	Notes
p	<p>	<i>puraad</i> 'fast'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5023">https://talkingdictionary.swarthmore.edu/quialana/?entry=5023</a>	
b	<b>	<i>bitx</i> 'cat'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4956">https://talkingdictionary.swarthmore.edu/quialana/?entry=4956</a>	
t	<t>	<i>tap</i> 'four'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4908">https://talkingdictionary.swarthmore.edu/quialana/?entry=4908</a>	
d	<d>	<i>dam</i> 'owl'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4937">https://talkingdictionary.swarthmore.edu/quialana/?entry=4937</a>	
k	<k>	<i>kiit</i> 'to play'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5155">https://talkingdictionary.swarthmore.edu/quialana/?entry=5155</a>	
g	<g>, <gu>	<i>ga</i> 'nine'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4925">https://talkingdictionary.swarthmore.edu/quialana/?entry=4925</a>	
m (lenis)	<m>	<i>maay</i> 'animal'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=3907">https://talkingdictionary.swarthmore.edu/quialana/?entry=3907</a>	
m (fortis)	<mm>			
n, ŋ (lenis)	<n>	<i>náax</i> 'chocolate'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=2581">https://talkingdictionary.swarthmore.edu/quialana/?entry=2581</a>	
n, ŋ (fortis)	<nn>			
r	<r>	<i>rúub</i> 'basket'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4999">https://talkingdictionary.swarthmore.edu/quialana/?entry=4999</a>	
r	<rr>	<i>burre</i> 'donkey'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4853">https://talkingdictionary.swarthmore.edu/quialana/?entry=4853</a>	Spanish borrowing
f	<f>			
s	<s>	<i>sáa</i> 'to walk'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5099">https://talkingdictionary.swarthmore.edu/quialana/?entry=5099</a>	
z	<z>	<i>zit</i> 'distant'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5021">https://talkingdictionary.swarthmore.edu/quialana/?entry=5021</a>	
ʃ	<x>	<i>xa</i> 'what'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5167">https://talkingdictionary.swarthmore.edu/quialana/?entry=5167</a>	
ʂ	<xʳ>			
ʒ	<zh>	<i>zhít</i> 'cat'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4955">https://talkingdictionary.swarthmore.edu/quialana/?entry=4955</a>	
ʒ	<zhʳ>			
x	<g>, <gu>	<i>gity</i> 'squash'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4751">https://talkingdictionary.swarthmore.edu/quialana/?entry=4751</a>	
l	<l>	<i>laas</i> 'thin'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5010">https://talkingdictionary.swarthmore.edu/quialana/?entry=5010</a>	
l (fortis)	<ll>			
ʈ	<ts>	<i>tsí</i> 'ten'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4909">https://talkingdictionary.swarthmore.edu/quialana/?entry=4909</a>	
tʃ	<tx>	<i>txa</i> 'a, an, one'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=3276">https://talkingdictionary.swarthmore.edu/quialana/?entry=3276</a>	
ɖʒ	<dx>	<i>dxambeu</i> 'moon god'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4918">https://talkingdictionary.swarthmore.edu/quialana/?entry=4918</a>	Not in Chavez-peon
j	<y>	<i>yáx</i> 'avocado'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=5001">https://talkingdictionary.swarthmore.edu/quialana/?entry=5001</a>	
w	<w>	<i>wí</i> 'a type of plant'	<a href="https://talkingdictionary.swarthmore.edu/quialana/?entry=4823">https://talkingdictionary.swarthmore.edu/quialana/?entry=4823</a>	When not word final

Finally, Table 11 below provides a number of words in Quialana Zapotec written in Aurora's current spelling system with a side by side comparison of how they would be written if all the above suggestions were to be implemented.

**Table 11. Comparison of Current and Suggested Systems.**

Current Spelling System	Implementation of Suggestions
<i>cadxiíichy</i> <sup>1</sup> 'angry'	<i>kadxiítxy</i>
<i>Ila lláaly</i> <sup>2</sup> 'copal tree'	<i>ya yáaly</i>

1. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5061>, 2. <https://talkingdictionary.swarthmore.edu/quialana/?entry=5190>

These examples were chosen specifically to show variation but it is worth noting that a significant number of spellings would be unaffected by the suggestions presented throughout this paper. The orthography is still recognizable and primarily distinguishes between previously equivalently represented phonemes (*Ila lláaly* vs. *ya yáaly*) or simplifies the system by using a single grapheme where multiple are currently used (*cadxiíichy* vs. *kadxiítxy*).

It is also apparent that the suggested orthography looks far less like the Spanish orthography. As previously discussed, this is a common area of tension. Ultimately this will likely be a deciding factor in whether Aurora and her community adopt these suggestions – whether they want to keep the orthography close to Spanish to ease learning or move away from Spanish orthography to establish a clearly different identity.

## 11. Further study

The current work will be partially translated into Spanish and sent to Aurora for her to review and use as she sees fit. Exactly which portions are translated is dependent on how Aurora wishes to use this work and what she would find helpful in her own work. The translation will begin in early 2020.

The most apparent area for further study is the creation of a full phonological analysis of Quialana Zapotec. Of particular importance to further papers of this type is the documentation of the vowel system so as to allow for continued orthographic development. Further understanding the phonology of Quialana Zapotec will also provide the basis for deeper comparisons of the similarities and differences from Quiaviní Zapotec as well as the multitude of Tlacolula Valley Zapotecs.

The Talking Dictionaries also present a rich resource for computational analyses and studies. Creating Talking Dictionaries for a language simultaneously creates a corpus of great use to the community and provides hundreds to thousands of words in the language – most with audio and glosses in at least one high resource language – which are already digitally organized and formatted. Importantly, the Talking Dictionaries also work off of generally the same framework regardless of language. Thus, if a process could be established using the Talking Dictionary format, it could be quickly applied to a variety of endangered languages. Possible future directions include semi-automated vowel classification (drawing on the audio files) and basic grapheme to phoneme converters or speech synthesis/recognition systems.

## 12. Conclusion

By providing a description of the Talking Dictionaries, their features, and their importance to endangered languages, this paper began by establishing a contextual background for San Bartolomé Quialana Zapotec. My positionality in this work was detailed and Aurora Sánchez Gómez was introduced along with the original inspiration for her spelling system. As Quialana Zapotec lacks a phonological analysis, the phonology of neighboring Quiaviní Zapotec was examined drawing primarily from the extensive work led by Munro. Aurora's spelling system was documented and areas of over and under representation of the hypothesized phonology were identified. Each area was discussed, and a suggested solution was given. Finally, a summary of these developments was provided, and possible directions of future study were listed.

This paper is meant not to be the end of this conversation, but rather the beginning. Whether these suggestions are implemented is Aurora's decision (and more than likely another dialogue). On a more general note, I hope that this paper will begin more work on Quialana Zapotec and encourage a wider range of uses for the Talking Dictionaries as a whole.

## References

- Anderson, G. D.S. & Harrison, K. D. (2003) *Tuvan Dictionary*. München: Lincom Europa.
- Cahill, Michael. 2011. "Non-Linguistic Factors in Orthographies," Linguistic Society of America Annual Meeting, Symposium for Developing Orthographies for Unwritten Languages.
- Casquite, Mansueto, and Catherine Young. 2017. "Hearing Local Voices, Creating Local Content: Participatory Approaches in Orthography Development for Non-Dominant Language Communities." In *Creating Orthographies for Endangered Languages*, edited by Mari C. Jones and Damien Mooney, 54–68. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/9781316562949.003>.
- Chávez-Peón, Mario E. n.d. "Tone-Bearing Units in Quiaviní Zapotec: The Split between Fortis and Lenis Sonorants," .
- Chávez-Peón, Mario E. n.d. "Loanword Phonology in San Lucas Quiaviní Zapotec (SLQZ)," . "Zapotec." n.d. Ethnologue. Accessed December 9, 2019.  
<https://www.ethnologue.com/language/zap>.
- Esposito, Christina M., and Sameer ud Dowla Khan. 2012. "Contrastive Breathiness across Consonants and Vowels: A Comparative Study of Gujarati and White Hmong." International Phonetic Association. *Journal of the International Phonetic Association*; Cambridge 42 (2): 123–43. <http://dx.doi.org/10.1017/S0025100312000047>.
- Harrison, K David, Brook Danielle Lillehaugen, Jeremy Fahringer, and Felipe H Lopez. 2019. "Zapotec Language Activism and Talking Dictionaries," Proceeding of eLex 2019.
- Harrison, David K., Brook Danielle Lillehaugen, Jeremy Fahringer, Felipe H. Lopez, Saul Ontiveros, and Chantal Reyes. 2019. "Zapotec Talking Dictionaries: Collaboration, Documentation, & Language Activism." Symposium Presentation presented at the Indigenous Languages Today, UPenn, October 26 2019.
- Harrison, K. David. 2007. *When Languages Die: The Extinction of the World's Languages and the Erosion of Human Knowledge*. Oxford University Press.  
<https://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195181920.001.0001/acprof-9780195181920>.
- Hull, Kerry. 2017. "When Letters Represent More Than Sounds: Ideology versus Practicality in the Development of a Standard Orthography for Ch'orti' Mayan." In *Creating Orthographies for Endangered Languages*, edited by Mari C. Jones and Damien Mooney, 142–54. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/9781316562949.008>.
- "What Is Language Loss?" n.d. The Language Conservancy (blog). Accessed November 21, 2019.  
<https://www.languageconservancy.org/understanding-the-issue/what-is-language-loss/>.
- Lillehaugen, Brook Danielle. 2016. "Why Write in a Language That (Almost) No One Can Read? Twitter and the Development of Written Literature," September, 356–93.
- Lillehaugen, Brook Danielle. 2019. "Otomanguean Languages". In Siddiqui, Daniel et al. *Routledge Handbook on North American Languages*. Pp. 331-364. New York: Taylor & Francis/Routledge Press.

- Lillehaugen, Brook Danielle. "Tweeting in Zapotec: social media as a tool for language activists." *Indigenous Interfaces: Spaces, Technology, and Social Networks in Mexico and Central America* (2019): 202-226.
- Lillehaugen, Brook Danielle. "Expressing location in Tlacolula valley Zapotec." PhD diss., University of California, Los Angeles, 2006.
- Munro, Pamela. 1996. "Making a Zapotec Dictionary." *Dictionaries: Journal of the Dictionary Society of North America* 17 (1): 131–55. <https://doi.org/10.1353/dic.1996.0020>.
- Munro, Pamela, and Felipe H. Lopez. 1999. *Di'csyonaary x:Tèe'n Di'zh Sah Sann Lu'uc = San Lucas Quiaviní Zapotec Dictionary*. Los Angeles: UCLA Chicano Studies Research Center Publications.
- Munro, Pamela. 2004. "Zapotec Grammar without Tears (Except Perhaps for the Grammarian)." <https://escholarship.org/uc/item/57j8t96p>.
- Munro, Pamela, Brook Lillehaugen Danielle, and Felipe H. Lopez. 2007. *Cali Chiu? A Course in Valley Zapotec*. Vol. 1.
- Oudijk, Michel R. 2011. "El texto más antiguo en zapoteco." *Tlalocan* 15 (December). <https://doi.org/10.19130/iifl.tlalocan.2008.191>.
- Roberts, David. 2008. "The Two Kabiye Orthographies: A Sociolinguistic and Linguistic Comparison." *Written Language & Literacy* 11 (1): 49–72. <https://doi.org/10.1075/wll.11.1.05rob>.
- Smalley, W. 1964. How shall I write this language? In W. A. Smalley (ed.), *Orthography studies: Articles on new writing systems. (Helps for translators 6)*. London: United Bible Societies, pp. 31–52.
- Uchihara, Hiroto, and Gabriela Pérez Báez. 2016. "Fortis/Lenis, Glides and Vowels in Quiaviní Zapotec." *Glossa; London* 1 (1): 1–24. <http://dx.doi.org/10.5334/gigl.13>.
- Uchihara, Hiroto. 2016. "Tone and Registrogenesis in Quiaviní Zapotec." *Diachronica* 33 (2): 220–54. <https://doi.org/10.1075/dia.33.2.03uch>.
- Urcid, Javier. 2011. "The Written System as a Cultural Code". *Their Way of Writing. Scripts, Signs, and Pictographs in Pre-Colombian America*. Edited by Boone, Elizabeth Hill and Urton, Gary. 111-143.

#### Zapotec Talking Dictionaries

- Lillehaugen, Brook Danielle, Roberto Antonio Ruiz, and Josefina Antonio Ruiz with Caroline Batten, Helen Felker, Alex Mannix, Katherine Donovan McCormick, and Rachel Elizabeth Weissler. 2013. *Tlacolula de Matamoros Zapotec Talking Dictionary*, pilot version. Living Tongues Institute for Endangered Languages. <http://www.talkingdictionary.org/tlacolula>
- Lillehaugen, Brook Danielle, Felipe H. Lopez, Pamela Munro, with Savita M. Deo, Graham Mauro, and Saul Ontiveros. 2019a. *San Lucas Quiaviní Zapotec Talking Dictionary*, version 2.0. Living Tongues Institute for Endangered Languages. <http://www.talkingdictionary.org/sanlucasquiavini>
- Lillehaugen, Brook Danielle & Janet Chávez Santiago, with Allison Freemond, Neal Kelso, Jaime Metzger, Kate Riestenberg, & K. David Harrison. 2019. *Teotitlán del Valle Zapotec Talking Dictionary*, version 2.0. Living Tongues Institute for Endangered Languages. <http://www.talkingdictionary.org/teotitlan>
- Lillehaugen, Brook Danielle and Moisés García Guzmán, with Kathryn Goldberg, María Mercedes Méndez Morales, Benjamin Paul, May Helena Plumb, Chantal Reyes, Cecilia G. Williamson, & K.

David Harrison. 2019. *Tlacoahuaya Zapotec Talking Dictionary*, version 2.1. Living Tongues Institute for Endangered Languages. <http://www.talkingdictionary.org/tlacoahuaya>

Lillehaugen, Brook Danielle, Aurora Sánchez Gómez, and Floriana Hernández Martínez, with Savita M. Deo, Graham Mauro, Saul Ontiveros, & K. David Harrison. 2019b. *San Bartolomé Quialana Zapotec Talking Dictionary*, version 1.0. Living Tongues Institute for Endangered Languages. <http://www.talkingdictionary.org/quialana>