Accounting for toponyms with multiple classifiers in Colonial Valley Zapotec

by Jasper Nash

This senior thesis was submitted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Linguistics

Bryn Mawr College

Fall 2022

Abstract

Toponyms are place names which have certain common linguistic properties across languages. Toponyms are typically formed by the combination of descriptive morphemes and noun classifiers (Ursini 2017). Southern Zapotec (SZ) languages are primarily spoken in the Southern Sierra Madre region of Oaxaca, Mexico. Southern Zapotec toponyms appear to have similar structures compared to those in the four languages analyzed by Ursini (2017), with spatial classifiers and descriptive morphemes merging to create a phrase which carries spatial features. Beam de Azcona (2012) categorizes Southern Zapotec toponyms according to whether there are no classifiers, one classifier, or multiple classifiers, as well as whether they contain relational nouns.

This paper assesses a corpus of Colonial Valley Zapotec (CVZ) toponyms to determine whether the analysis in Beam de Azcona (2012) of toponyms with multiple classifiers in Southern Zapotec languages can be extended to describe the CVZ data as well. The corpus was created through the keyword searching of contemporary Spanish translations of Zapotec manuscripts on the digital text explorer Ticha. Then, the corpus was analyzed by standardizing spelling and sorting alphabetically to identify any pairs of toponyms that differ only by a singular morpheme.

Through the identification of these toponym pairs, evidence was found to support both types of toponyms with multiple classifiers as described in Beam de Azcona (2012). Many examples demonstrate optionality with the initial classifier in a toponym with multiple classifiers, while a couple of examples suggest possible optionality with an internal classifier. The results are consistent with the categories provided by Beam de Azcona (2012), although a few data points raise questions for further inquiry.

Acknowledgements

This thesis was made possible by the support of many folks. First of all, I would like to express my deepest appreciation to my thesis advisor, Professor Brook Lillehaugen, whose classes and work inspired me to pursue this topic and who helped me navigate the entire process with constant encouragement. Many thanks as well to my second faculty reader, Professor Jane Chandlee, who provided thoughtful comments and guidance.

I would like to extend my sincere thanks to my writing partner, Nicha, who has been there to keep me on track and to help me get everything out of my head and into the world. I am also grateful for my senior seminar group members, who have provided both feedback and encouragement throughout the semester.

Lastly, I'd like to recognize my roommates, my partner, and my family, who have all been there for me at every deadline, keeping me motivated and offering support whenever I needed it. Thank you all.

Table of Contents

1. Introduction

Toponyms are place names, whose distributional properties and semantic interpretations are guided by their similar morphological structures (Ursini 2017). Toponyms in Southern Zapotec (SZ) languages, as described by Beam de Azcona (2012), tend to comprise several morphemes. Beam de Azcona (2012) provides a classification of SZ toponyms according to their structure: classified noun phrases (with either a single or multiple classifiers), toponyms with relational nouns, and toponyms without a relational noun or classifier.. She further breaks down toponyms with multiple classifiers into two categories based on the optionality of the classifiers.

The Ticha online database of Colonial Valley Zapotec (CVZ) manuscripts provides a rich source of CVZ toponyms that can be easily accessed through searching by keyword. Using the transcriptions of contemporary Spanish translations of CVZ texts as a point of entry, I created a corpus of CVZ toponyms from Ticha documents. The majority of the toponyms in this corpus include at least one classifier, and many of them contain two or more classifiers. In this paper, I analyze how CVZ toponyms with multiple classifiers fit the paradigm of Southern Zapotec toponyms established by Beam de Azcona (2012).

2. Background

2.1. Toponyms

Toponyms are place names, including the names of cities, towns, bodies of water, and more: e.g., *Paris*, *Philadelphia*, and *Mississippi River*. Blair & Tent (2021) provide a linguistic typology of toponyms with reference to Australian English place names; they focus on expressions of the naming intention, i.e., the ways that the motivations and intentions behind naming can be carried out, and they separate toponyms into different categories based on these

expressions: descriptive, associative, evaluative, occurrent, copied, eponymous, and innovative. For example, *Cape Manifold* is considered a descriptive expression because it was selected due to the number of high hills above it (Blair & Tent 2021).

Ursini (2017) examines the cross-linguistic traits of toponyms through an analysis of four languages of increasing morphological complexity (English, Mandarin, Italian, and Finnish). Ursini (2017) found that in all four of the languages studied, toponyms have restricted syntactic distributions. In English, toponyms cannot distribute with articles or quantifiers in complement position. Examples (1)-(2) show this with the quantifier *every*[1]. (The symbol * means that the inclusion of a given lexical item makes a sentence ungrammatical.)

1. *The boys are sitting in front of *∅/the/every car* (Ursini 2017, 2)
2. *The boys live north of ∅/*the/*every Northampton* (Ursini 2017, 2)

The same pattern holds for the three other languages examined (Ursini 2017). Ursini's examination of the morphological structure of toponyms in these four languages indicates that both productive and non-productive compounding and affixation rules are involved–i.e., that the rules which describe the compounding of morphemes describe the structure and still apply regardless of whether those morphemes are seen elsewhere in the language (2017). These rules all use the same building blocks: descriptive morphemes (nouns and adjectives) and spatial heads (including classifiers, case markers, and suffixes).

Using the framework of Type Logical Syntax, Ursini (2017) demonstrates a simple process of combining morphemes to create toponyms. In Type Logical Syntax, different types are defined that can be related to each other through rules of multiplication ("$\cdot$") and division

[1] Ursini (2017) does not consider appelative constructions, such as *every city called "Paris",* which can refer to locations sharing a toponym.

("/"). Descriptive morphemes are designated as type *p*, spatial heads are described as *p'/p*, and toponyms, which Ursini defines as NPs carrying spatial features, are given the type *p'*. Through the multiplication rule, spatial heads and descriptive morphemes can be merged to create a toponym. For example, the descriptive morpheme *North* (type *p*) merges with the spatial classifier *hampton* 'settlement' (type *p'/p*) to form a phrase of type *p'*, *Northampton*. This formula would yield the same results if the order of the descriptive morpheme and the spatial classifier were reversed (Ursini 2017).

2.2 Zapotec languages

Zapotec refers to a group of languages in the Otomanguean family which are indigenous to southern Mexico and are primarily spoken in the state of Oaxaca (Lillehaugen et al. 2016). With respect to time depth and diversity, Zapotec languages are comparable to Romance languages and the Otomanguean family to the Indo-European family (Lillehaugen et al. 2016). Colonial Valley Zapotec refers to a variety of Zapotec from the Central branch of the family, which is attested in a written corpus from the colonial period of Mexico, 1521-1821 (Lillehaugen et al. 2016).

Southern Zapotec (SZ) languages are spoken in the Southern Sierra Madre region of Oaxaca (Beam de Azcona 2012). Beam de Azcona (2012) examines toponyms in SZ languages and describes the different types of toponyms by separating them into categories. She found that toponyms in SZ languages are typically made up of both descriptive morphemes and spatial classifiers, as well as occasionally relational nouns. Her divisions are as follows: classified noun phrases, with either a single independent classifier or multiple classifiers; toponyms with relational nouns; and toponyms without a relational noun or classifier (Beam de Azcona 2012).

Classified noun phrases appear in more than just toponyms in Southern Zapotec (Beam de Azcona 2012). In these phrases, a noun with a generic meaning serves as the head of the classified noun phrase. For example, in the Coatec phrases *má lwê* 'winged animal' and *má yìch* 'furry animal', the classifier *má* indicates that the phrase is a type of animal, and the second morpheme defines the type of animal (Beam de Azcona 2012). In a toponym with a single classifier, the same pattern occurs. The morpheme *la'tz* 'llano, valle; plain, valley' in Miahuatec appears in examples such as *La'tz Naróo* 'Llano Grande', where a single classifier merges with a descriptive morpheme to form the place name (Beam de Azcona 2012).

The role of each classifier is more complex in SZ toponyms with multiple classifiers. Beam de Azcona (2012) defines two categories to explain the two modes of behavior exhibited in these toponyms. In the first type, the additional classifier provides more information about the place. In the second type, a classifier which is normally used with a common noun but would typically be dropped in a toponym, is retained. In type 1, the classifier that is farther to the left would be considered less obligatory than one closer to the end (Beam de Azcona 2012). For example, Beam de Azcona (2012) shows three different ways that the place 'Miahuatlán' is referred to in different SZ languages. In San Agustín Mixtepec, a variety of Miahuatec, it is called *Yìs Dòʔo* [town holy(.thing)]. In the Coatec varieties San Baltazar Loxicha and Santa María Coatlán, the same morphemes appear: *Yêzh Doʔ* [town holy(.thing)]. However, in the Coatec variety Campo Nuevo, there is an additional classifier at the front, *làt* 'plain', which does not appear in the other two: *Làt Yêzh Doʔ* [plain town holy(.thing)] (Beam de Azcona 2012). The absence of this classifier from the other examples shows a level of optionality with the left-most classifier.

In contrast, the left-most classifier in the second type would be more obligatory than one that is further to the right (Beam de Azcona 2012). Classified noun phrases that are embedded in SZ toponyms typically drop the classifier from the common noun phrase (Beam de Azcona 2012). In the Miahuatec variety San Baltazar Loxicha, when the noun phrase *yáa dón* 'guarumbo' becomes part of a toponym with the classifier *Yó* 'river', the whole toponym becomes *Yó'Dón* 'Arroyo Guarumbo', and the classifier *yáa* 'tree' is left out (Beam de Azcona 2012). Toponyms where a classifier in this position is not omitted define Beam de Azcona's second type. For example, in San Baltazar Loxicha, there is a place called *Là'tz Yáa Guín* 'Llano Palmiche' [plain tree palm] (Beam de Azcona 2012). The retention of the classifier *yáa* 'tree' in the phrase which means 'palm tree', in contrast to the example of *Yó'Dón*, shows optionality with the classifier in the right-most position.

In this paper, I examine Colonial Valley Zapotec toponyms with multiple classifiers through a corpus of 90 data points which I have compiled from an online corpus of digitized manuscripts. Using Ursini 2017 as a reference point for typology and morphosyntactic analysis of toponyms in other languages, I explore how Beam de Azcona's framework for understanding toponyms with multiple classifiers in Southern Zapotec accounts for the Colonial Valley Zapotec data.

3. Methods

Ticha (https://ticha.haverford.edu; Lillehaugen et al. 2016) is a digital explorer for a corpus of Colonial Valley Zapotec texts (Broadwell et al. 2020). The documents displayed on Ticha include Zapotec language manuscripts and also contemporary translations of Zapotec documents into Early Modern Spanish. Using the search capabilities of Ticha, I built a corpus of

Colonial Valley Zapotec toponyms. The Spanish documents on Ticha have all been transcribed, which facilitates automatic searching by keyword. To identify toponyms in the manuscripts, I have utilized a common practice of the translators to Early Modern Spanish which was that they would transcribe Zapotec place names rather than translate them. These copied Zapotec toponyms are frequently introduced by some form of the Spanish word *nonbrado/nombrado* 'named', which makes them easily searchable regardless of the usage of specific classifiers or the type of toponym.

This method of data collection introduces multiple potential sources of deviation from the original Zapotec writing. It relies on the judgment of the translator to Early Modern Spanish who copied it originally in determining what constituted the toponym and in accurately reproducing it, and it depends on the decisions of whoever transcribed the documents so that they could be searched. Given these factors, I cross-referenced each toponym in my corpus (which were all initially found via the Spanish translations) with the corresponding original Zapotec document, examining the transcription for accuracy and looking at the context to determine the boundaries of the toponym. In Colonial Valley Zapotec writing, spelling was highly variable. The same word might be written in multiple ways even within one text. Common variations include that "vowels may be written singly or doubly, with or without accent marks, and the vowel <o> frequently alternates with <u>, and <i> (also written as <y> or <ij>) with <e>; among consonants, there is frequent variation in stops, yielding <p>/<b>, <t>/<d>, and <c, qu>/<g(u)> alternations; <r> may alternate with <t>; and the fricatives [s] and [z] may be written as <s>, <z>, <ç>, or <c>" (Foreman & Lillehaugen 2017). Thus, variations such as *quinaa* and *guina* will be considered to represent the same morpheme in this paper.

The complete list of toponyms generated for this study, which can be found in full in Appendix A, does not constitute every toponym that appears in the manuscripts on Ticha, and should not be considered exhaustive. Additionally, any reference to a toponym which I have judged to be a duplicate of a toponym already present in my corpus has been omitted from this paper, as my analysis is concerned with types not tokens. Because these toponyms in this corpus were found in bills of sale and last wills and testaments, they are primarily fields/pieces of land along with a few landmarks used to describe the location of pieces of land.

4. Data

For this paper, classifiers were identified based primarily on a table of common classifiers that occur in SZ toponyms in Beam de Azcona (2012). Some morphemes were included as semantic analogs to those classifiers. Other morphemes were included as classifiers due to behavior that was consistent with the classifiers from the chart in Beam de Azcona (2012), i.e., regular occurrence in the left-most position and apparent optionality. The question of whether such classifier-headed strings are noun phrases or noun-noun compounds is addressed in Beam de Azcona (2016), but is out of the scope of this paper.

4.1 Optionality with Initial Classifier

Several pairs of data points seem to show optionality with the first classifier, which is consistent with Beam de Azcona's first category (2012). In example (1), the toponym consists of three morphemes, of which I consider the first two to both be classifiers.

(1) guinaa-guiya-yana (Al642[2]–40r;17)

[2] For the citation of the location of a toponym in a Ticha document, I will use this format: (TextCode–PageNumber;LineNumber). I use the text code indicated by Ticha and the page/line number at which the toponym starts.

sown.field[3]-high.place[4]-corn[5]

Example (2) shows a toponym with two morphemes, which are the same as the last two morphemes of example (1).

(2) quiya-yana (Al642-40r;15)

high.place-corn

Example (2) could refer to the same location as example (1), due to their appearance in the same document, Al642, and the overlap in morphemes. Example (2) does not include the classifier *guinaa* 'sown field' which occurs at the beginning of example (1). Because it is likely that these are both referring to the same location, this provides evidence for the omission of initial classifiers in certain CVZ toponyms.

Examples (3) and (4) present another pair of toponyms in which the only difference is the presence of *quinaa* 'sown field' at the beginning of example (3) and its absence from example (4). Both examples are found in the same document and are presumed from context to refer to the same location.

(3) quinaa-queeya-huee (Al711b–46r;8)

sown.field-high.place-fold[6]

(4) queeya-huee (Al711b–46r;15)

high.place-fold

In examples (5) and (6), the same apparent optionality of *quinaa* 'sown field' appears. These variations also appear in the same document.

(5) quinaa-quela-xaloo (Al700–48v;3)

---

[3] Córdova 401v: "tierra para sembrar…quinaa"
[4] Córdova 024r: "alto lugar…quiaa"
[5] Córdova 252r: "maçorca de mayz…yana"
[6] Córdova 144r: "dobladura…huee"

sown.field-cornfield[7]-?

(6) quela-xaloo (Al700–48v;8)

cornfield-?

Although the potential classifiers in examples (7) and (8) have not been glossed, they seem to follow in the pattern established with *quinaa* ‘sown field’ and have therefore been included for the representation of possible optionality with *cha*. These variations also appear within one document, increasing the likelihood that they both refer to the same location.

(7) cha-cho-xi.loguaa (Al711c–56v;3)

?-behind[8]-POSS[9].forehead[10]

(8) choo-xi.loogua (Al711c–56r;28)

behind-POSS.forehead

4.2 Optionality with Internal Classifier

The next examples show evidence of optionality with the right-most classifier. In examples (9) and (10), which are both from one document, the only difference in morphemes is the presence of *lachi* ‘flat land’ in the middle of (9) but not in (10). This may indicate that *lachi raqueya* constitutes its own classified noun phrase, as in Beam de Azcona’s second category (2012).

(9) quinaa-lachi-raqueya (Al711c–56v;1)

sown.field-flat.land[11]-?

(10) quinaa-raqueeya (Al711c–56v;9)

[7] Córdova 213r: “haça de trigo mayz…quela”
[8] Córdova 408v: “tras…choo”
[9] Córdova 268v: “mio cosa mia…xi”
[10] Córdova 200v: “frente…loo cuaa”
[11] Córdova 069v: “campo o tierra llana…lache”

sown.field-?

The distinction between examples (11) and (12) is the morpheme *loo* ‘face’, which is the third morpheme in (12) (after two classifiers) but which does not appear in (11). Although *loo* is a relational noun and not a classifier like *quinaa* ‘sown field’, the apparent internal optionality may still reflect the same type described by Beam de Azcona. These variations do not appear in the same document, so there is a significant chance that these two toponyms do not refer to the same location; however, the three shared morphemes provide reason to consider the possibility that they might.

(11) guina-guego-yaga (Al686b–69v;3)

sown.field-river[12]-tree[13]

(12) quinaa-queco-loo-yaga (Al697–62v;9)

sown.field-river-face[14]-tree

4.3 Additional Toponyms with Multiple Classifiers

Further pairs of toponyms raise questions about the structure of CVZ toponyms. Examples (13) and (14) are two toponyms from one document where one starts with the classifier *quina* ‘sown field’ and the other starts with the classifier *xolare* ‘house plot’, a borrowing from Spanish, followed by the same two morphemes in each. While it is possible that these are not referring to the same location, they are both present in the same document and the context makes it plausible that they are. If this is the case, the ability to substitute a Spanish borrowing for a Zapotec classifier is interesting and merits further examination.

(13) quina-quego-eza (Al711c–56r;24)

[12] Córdova 360v: “rio generalmente…quego”
[13] Córdova 036r: “arbol generalmente…yaga”
[14] Córdova 205r: “gestro o rostro…loo”

sown.field-river-?

(14) xolare-quego-hueza (Al711c–56r;19)

house.plot[15]-river-?

The final example is not a pair, but a single toponym notable for its length. Example (15) seems to embed one toponym into another using *chacho* 'be located separate'.[16] The entire sequence of morphemes presented here was interpreted as one toponym in the contemporary Spanish translation, which suggests that the embedded toponym may have become part of the overall unit rather than a separate description.

(15) quinaa-quego-ezaa-chacho-quego-cho-xi.loguaa (Al711c–56v;9)

sown.field-river-?-be.located.separate-river-behind-POSS.forehead

For a list of toponyms analyzed for this paper, see table A1 in Appendix A. Additionally, the toponyms which were identified for this paper are listed in full in table A2 of the appendix.

5. Analysis

Beam de Azcona (2012) separates SZ toponyms with multiple classifiers into two categories. The first category involves toponyms where the first classifier is more optional and is adding information about the place. This category is represented in four pairs of toponyms from the CVZ corpus, which each have one variation that uses the additional classifier at the beginning and one variation that excludes it. These examples provide evidence that within Colonial Valley Zapotec, some degree of optionality was present with respect to initial classifiers that provided extra information, which is consistent with the cross-linguistic data from the Southern Zapotec languages. There are many further examples in the full corpus presented in

[15] Ticha (Lillehaugen et al. 2016): "solar…plot of land for a home"
[16] Córdova 035v: "aquende luego alli…chachoogaa"

table A2 (see Appendix A), which are toponyms that contain multiple classifiers and do not have a complement that omits a classifier. The majority of these likely fall in this first category, given the prevalence of examples of this type and the classifiers involved. However, without the complementary examples, they cannot be conclusively assigned to this type.

Beam de Azcona's second category captures the opposite situation of the first: the initial classifier is more obligatory, and the internal/right-most classifier is more likely to be omitted (2012). In these situations, Beam de Azcona (2012) asserts that the internal classifier is primarily associated with an embedded noun phrase, and results from the infrequent circumstance where that classifier is not dropped. In the corpus developed for this paper, this category was less common than the first one. There are two pairs of toponyms which may possibly fit into this category, although one of these pairs comes from separate documents and is thus less likely to be referencing a common location. Nevertheless, examples (9) and (10) show internal classifier optionality which suggests that Beam de Azcona's second category of toponyms with multiple classifiers is also applicable to CVZ toponyms.

One observation of note is that the toponyms with an omitted classifier tended to appear later in each document than the associated toponym which included the classifier, though this was not always the case. It is possible that the mention of a toponym in full creates conditions that render the omission of a classifier more acceptable in that context.

Additional examples that do not directly correspond with Beam de Azcona's two categories raise questions for future research. Firstly, the example of the Spanish borrowing *xolare* 'house plot' substituting for *quina* 'sown field' in examples (13) and (14) prompts questions about type one optionality–is there a connection between this ability for a classifier to be dropped and the ability for a classifier to be replaced? Can borrowings fill the same roles as

Zapotec words? Further, example (15) complicates the categorization of toponyms with multiple classifiers into types, which are primarily distinguished under a model that is limited to two classifiers, by possibly embedding an entire toponym with its own classifier(s) into a toponym which already contained multiple classifiers. How can complex toponyms with more than two classifiers be categorized?

6. Conclusion

A corpus of Colonial Valley Zapotec toponyms was created from digitized manuscripts on the Ticha digital text explorer to investigate whether the categories of Southern Zapotec toponyms as outlined in Beam de Azcona (2012) were attested in CVZ data. This paper specifically examined the two types of toponyms with multiple classifiers that Beam de Azcona (2012) defined. The first type involves optionality with the first/left-most classifier, while the second involves optionality with any other internal classifier. Pairs of toponyms were identified that seem to support the existence of both types in Colonial Valley Zapotec, although greater evidence was found for the first type. Some data points raise further questions about the nature of CVZ toponyms, how borrowings from Spanish fit into the Zapotec structures, and how toponyms with three or more classifiers relate to this scheme of categorization. Overall, the categories created to describe Southern Zapotec toponyms appear applicable to Colonial Valley Zapotec data as well.

Works Cited

Beam de Azcona, Rosemary G. 2012. Southern Zapotec Toponyms. In Brook Danielle Lillehaugen & Aaron Huey Sonnenschein (eds.), *Expressing location in Zapotec*. Muenchen, Germany: LINCOM Europa.

Beam de Azcona, Rosemary G. 2016. Zapotecan Languages. In *Oxford Research Encyclopedias, Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.73 Google Scholar

Blair, David & Jan Tent. 2021. A revised typology of place-naming. *Names* 69(4). 30-47. https://doi.org/10.5195/names.2021.2260

Broadwell, George Aaron, Moisés Guzmán, Brook Danielle Lillehaugen, Felipe H. Lopez, May Helena Plumb, and Mike Zarafonetis. 2020. Ticha: Collaboration with Indigenous communities to build digital resources on Zapotec language and history. *Digital Humanities Quarterly* 14(4). http://www.digitalhumanities.org/dhq/vol/14/4/000529/000529.html

de Córdova, Fr. Juan. 1987 [1578]. *Vocabulario en lengua çapoteca*. México: Ediciones Toledo (INAH).

Foreman, John & Brook Danielle Lillehaugen. 2017. Positional Verbs in Colonial Valley Zapotec. *International Journal of American Linguistics* 83(2). https://doi.org/10.1086/689846

Lillehaugen, Brook Danielle, George Aaron Broadwell, Michel R. Oudijk, Laurie Allen, May Plumb, and Mike Zarafonetis. 2016. Ticha: a digital text explorer for Colonial Zapotec, first edition. Online: http://ticha.haverford.edu/

Ursini, Francesco-Alessio. 2017. On the structure of toponyms. In Lívia Kortvélyessy, Pavol Štekauer & Salvador Valera (eds), *Word formation universals*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Appendix A

This appendix consists of two corpuses of CVZ toponyms, the first of which is a smaller subset of the second. These corpuses should not be considered exhaustive. Additionally, duplicates have been omitted for the purposes of this paper. The headers for the two tables are the same and are repeated at the beginning of each new page. The "ID" field is a unique identifier created for this paper, and the category refers to the sort of place a toponym is, based primarily on the context of the Spanish translation, the Zapotec document, and the initial classifier as relevant. The "toponym" field indicates a place name as it appears in the original Zapotec text, based on my own transcription using the Spanish translation as a reference. "Morphological analysis" includes an approximate breakdown of the toponym into separate morphemes, with as many glosses as possible. The "text code" field is for a combination of letters (indicating location) and numbers (indicating year) which originated on Ticha. The "page" field indicates the page of a physical document on which the toponym appears. This is taken from Ticha's labels and copies the designation of "r" for *recto* and "v" for *verso* after a page number to indicate the particular side of a leaf of paper. The "line" header refers to the number of the line in which the toponym starts, counting down from the top.

Table A1. Selected Colonial Valley Zapotec toponyms.

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
1	Land	chacho xilo guaa	cha-cho-xi.loguaa ?-behind-POSS.forehead	Al711c	56v	3
2	Field	choo xilo cua	choo-xi.locua behind-POSS.forehead	Al642	40r	17
3	Field	guina guego yaga	guina-guego-yaga sown.field-river-tree	Al686b	69v	3
4	Field	guinaa guiya yana	guinaa-guiya-yana sown.field-high.place-corn	Al642	40r	17
5	Field	queeyahuee	queeya-huee high.place-fold	Al711b	46r	15
6	Field	quela xaloo	quela-xaloo cornfield-?	Al700	48v	8
7	Field	quina quego eza	quina-quego-eza sown.field-river-?	Al711c	56r	24
8	Field	quinaa lachi raqueya	quinaa-lachi-raqueya sown.field-flat.land-?	Al711c	56v	1
9	Field	quinaa q[ue]coloo yaga	quinaa-queco-loo-yaga sown.field-river-face-tree	Al697	62v	9

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
10	Field	quinaa que la xa loo	quinaa-quela-xaloo sown.field-cornfield-?	Al700	48v	3
11	Field	quinaa queeyahuee	quinaa-queeya-huee sown.field-high.place-fold	Al711b	46r	8
12	Field	quinaa quego ezaachacho quego cho xiloguaa	quinaa-quego-ezaa-chacho-quego-cho-xi.loguaa sown.field-river-?-be.located.separate-river-behind-POSS.forehead	Al711c	56v	9
13	Field	quinaa raqueeya	quinaa-raqueeya sown.field-?	Al711c	56v	9
14	Place	quiya yana	quiya-yana high.place-corn	Al642	40r	15
15	Field	xolare quego hueza	xolare-quego-hueza house.plot-river-?	Al711c	56r	19

Table A2. Full list of toponyms in Colonial Valley Zapotec collected for this paper.

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
1	Ditch	piacee yala	piacee-yala	Al714	72r	28
2	Field	choo xilo cua	choo-xi.locua behind-POSS.forehead	Al642	40r	17
3	Field	dellee	dellee	Al642	40r	5
4	Field	la cha laci	lacha-laci flat.land-?	Al714	72r	29
5	Field	lachi xooaya	lachi-xooaya flat.land-?	Al714	72r	7
6	Field	quela xaloo	quela-xaloo corn.field-?	Al700	48v	8
7	Field	queeyahuee	queeya-huee high.place-fold	Al711b	46r	15
8	Field	quinaa ceeni too lacha hua	quinaa-ceeni-too-lacha-hua sown.field-?-?-flat.land-?	Al711c	56r	22
9	Field	quinaa chahua	quinaa-chahua sown.field-?	Al711c	56r	25
10	Field	guinaa ciallaa	guinaa-ciallaa sown.field-?	Al649	36r	21

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
11	Field	guina g[ue]guitoo	guina-gueguitoo sown.field-?	Al686b	69v	2
12	Field	quinaa ichilooa	quinaa-ichi-looa sown.field-?-?	Al686b	69r	22
13	Field	quinaa lachi huaa	quinaa-lachi-huaa sown.field-flat.land-?	Al711c	56v	6
14	Field	quinaa lachi raqueya	quinaa-lachi-raqueya sown.field-flat.land-?	Al711c	56v	1
15	Field	quinaalachiyal ana	quinaa-lachi-yalana sown.field-flat.land-?	Al697	62v	10
16	Field	quinaa lachi zaque	quinaa-lachi-zaque sown.field-flat.land-?	Al697	62v	10
17	Field	guina larieza	guina-lari-eza sown.field-?-?	Al686b	69r	24
18	Field	guinaa lichi hui chiy	guinaa-lichi-hui-chi sown.field-?-?-?	Al649	36r	22
19	Field	guina llooyacaa	guina-lloo-yaca sown.field-land[17]-wood	Al649	36r	21

[17] Córdova 401v: “tierra para sembrar…yoo”

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
20	Field	quinaa loo hui	quinaa-loohui sown.field-?	Al697	62v	7
21	Field	guina ni chii	guina-nichii sown.field-?	Al686b	69r	23
22	Field	guina guego etaa	guina-guego-etaa sown.field-river-?	Al686b	69v	1
23	Field	quina quego eza	quina-quego-eza sown.field-river-?	Al711c	56r	24
24	Field	quinaa quego ezaachacho quego cho(?) xiloguaa	quinaa-quego-ezaa-chacho-quego-cho-xi.loguaa sown.field-river-?-be.located.separate-behind-POSS.forehead	Al711c	56v	9
25	Field	guina guego huichii	guina-guego-huichii sown.field-river-?	Al686b	69v	3
26	Field	quinaa queco hui lee	quinaa-queco-hui-lee sown.field-river-?-?	Al697	62v	6
27	Field	guinaa gueco laya	guinaa-gueco-laya sown.field-river-?	Al649	36r	20
28	Field	quinaa queco lloo [hua?]	quinaa-queco-lloo-hua sown.field-river-land-?	Al697	62v	9

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
29	Field	quinaa quecoloo	quinaa-queco-loo sown.field-river-?	Al697	62v	6
30	Field	guina gueco lochoo	guina-gueco-lochoo sown.field-river-?	Al686b	69v	2
31	Field	quinaa q[ue]coloo yaga	quinaa-queco-loo-yaga sown.field-river-?-wood	Al697	62v	9
32	Field	quinaa que conii	quinaa-queco-nii sown.field-river-?	Al697	62v	7
33	Field	quinaa queco ru cui	quinaa-queco-ru.cui sown.field-river-?	Al711c	56v	7
34	Field	quinaa quego techina	quinaa-quego-techi-na sown.field-river-back/behind[18]-?	Al711c	56r	25
35	Field	quinaa queco tedoo	quinaa-queco-tedoo sown.field-river-?	Al697	62v	9
36	Field	quinaa queco ti huini quero pichina	quinaa-queco-tihuini-quero-pichina sown.field-river-?-squash[19]-deer[20]	Al697	62v	11

[18] Ticha (Lillehaugen et al. 2016): "texi…back, shoulders. behind."
[19] Ticha (Lillehaugen et al. 2016): "queto…squash"
[20] Ticha (Lillehaugen et al. 2016): "bichina…deer"

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
37	Field	guina guego xaquee	guina-guego-xaquee sown.field-river-?	Al686b	69v	2
38	Field	quinaa quego yaa	quinaa-quego-yaa sown.field-river-?	Al711c	56r	23
39	Field	guina guego yaga	guina-guego-yaga sown.field-river-wood	Al686b	69v	3
40	Field	quinaa que goyalaa	quinaa-quego-yalaa sown.field-river-?	Al697	62v	18
41	Field	quinaa queco yoo huea	quinaa-queco-yoo-huea sown.field-river-land-?	Al697	62v	10
42	Field	guinaa gueco zana	guinaa-gueco-zana sown.field-river-?	Al649	36r	20
43	Field	guinaa guecozoo	guinaa-gueco-zoo sown.field-river-?	Al649	36r	20
44	Field	guina g[ue]la chee	guina-guela-chee sown.field-cornfield-?	Al686b	69r	25
45	Field	guina g[ue]la robi	guina-gueco-hueca sown.field-river-?	Al686b	69r	21
46	Field	quinaa que la xa loo	quinaa-quela-xaloo sown.field-cornfield-?	Al700	48v	3

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
47	Field	quinaa quela zee	quinaa-quela-zee sown.field-cornfield-?	Al697	62v	21
48	Field	guinaa gueguee	guina-gueguee sown.field-?	Al686b	69v	1
49	Field	quinaa quee ti la	quinaa-quee-tila sown.field-?-?	Al697	62v	6
50	Field	quinaa quexo cuii	quinaa-quexo-cuii sown.field-?-?	Al711c	56v	9
51	Field	quinaa queeyahuee	quinaa-queeya-huee sown.field-high.place-bend	Al711b	46r	8
52	Field	guinaa gueyazaa	guinaa-gueya-zaa sown.field-high.place-?	Al649	36r	21
53	Field	quinaaquia[?]tela	quinaa-quia-tela sown.field-high.place-?	Al697	62v	29
54	Field	guinaa guiya yana	guinaa-guiya-yana sown.field-high.place-corn	Al642	40r	17
55	Field	quinaa qui[loo] huea	quinaa-quiloo-huea sown.field-?-?	Al697	62v	8

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
56	Field	quinaa quique quelaxohueza a	quinaa-quique-quela-xo.huezaa sown.field-?-cornfield-?	Al711c	56v	6
57	Field	quinaa raqueeya	quinaa-raqueeya sown.field-?	Al711c	56v	9
58	Field	quinaa roci	quinaa-roci sown.field-?	Al697	62v	8
59	Field	quinaa ruabiacee	quinaa-rua-biacee sown.field-edge[21]-?	Al711c	56r	27
60	Field	quinaa sati laaqui	quinaa-sati-laaqui sown.field-?-?	Al697	62v	19
61	Field	quinaa sato laga	quinaa-sato-laga sown.field-?-same[22]	Al697	62v	13
62	Field	quinaa sego	quinaa-sego sown.field-?	Al697	62v	13
63	Field	quinaa sego quia	quinaa-sego-quia sown.field-?-high.place	Al697	62v	19
64	Field	quinaa see hui[u?]	quinaa-see-huiu sown.field-?-?	Al697	62v	20

[21] Córdova 059r: “borde de otra cosa…tohua”
[22] Córdova 174r: “entonces mismo…laagaa”

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
65	Field	quinaasuici	quinaa-suici sown.field-?	Al697	62v	28
66	Field	quinaa tani hoxo	quinaa-tani-hoxo sown.field-mountain[23]-?	Al711c	56v	8
67	Field	quinaate [chela]	quinaa-teche-la sown.field-back/behind-?	Al697	62v	7
68	Field	quinaatecoo	quina-tecoo sown.field-?	Al700	48v	3
69	Field	quinaa teexii	quinaa-teexii sown.field-back/behind	Al711c	56v	8
70	Field	quina teyaa	quina-teyaa sown.field-?	Al707	76r	3
71	Field	guinaa xadiguee	guinaa-xadiguee sown.field-?	Al649	36r	20
72	Field	guina xala	guina-xala sown.field-?	Al686b	69v	6
73	Field	quinaa xaquee xilla	quinaa-xaquee-xilla sown.field-?-cotton[24]	Al711c	56v	4

[23] Córdova 273r: “montaña…tani”
[24] Córdova 021v: “algodon…xilla”

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
74	Field	guinaa xa to nee	guinaa-xatonee sown.field-?	Al686b	69r	22
75	Field	guinaa xi chayoo	guinaa-xi.chayoo sown.field-POSS.?	Al686b	69r	22
76	Field	guina xoochigaya	guina-xoo-chiga-ya sown.field-?-?-?	Al649	36r	19
77	Field	guinaa xohuela	guina-xo-huela sown.field-?-?	Al686b	69r	24
78	Field	quinaa xo guee	quinaa-xo-guee sown.field-?-?	Al697	62v	18
79	Field	guina xo guina	guina-xo-guina sown.field-?-sown.field	Al686b	69r	25
80	Field	guina zacii	guina-zacii sown.field-?	Al686b	69r	25
81	Field	quinaa za laaya	quinaa-za-laaya sown.field-?-land[25]	Al697	62v	8
82	Field	guina zeyo	guina-zeyo sown.field-?	Al686b	69v	1

[25] Ticha (Lillehaugen et al. 2016): “laya…sp. var. of layo ‘land’”

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
83	Field	guinaa zovita	guinaa-zovita sown.field-?	Al686b	69r	24
84	Field	quinnaa quitecho quicha	quinnaa-quite-cho-quicha sown.field-?-behind-?	Al707	76r	6
85	Field	rua quela quichi	rua-quela-quichi edge-cornfield-thorn[26]	Al711c	56r	28
86	Field	xolare quego hueza	xolare-quego-hueza house.plot-river-?	Al711c	56r	19
87	Field	xolare quego lace xaqueechiihua	[xolare]-quego-lace-xaquee-chiihua house.plot-river-?-?-?	Al711c	56r	21
88	Hill	lacha queya rico	lacha-queya-rico flat.land-high.place-?	Al707	76r	9
89	Land	chacho xilo guaa	chacho-xi.loguaa be.located.separate-POSS.forehead	Al711c	56v	3
90	Land	quego xollaana	quego-xollaana river-noble[27]	Al698	48r	6
91	Land	zacui	zacui	Al711d	74r	2

[26] Córdova 186v: “espina…quechi”
[27] Ticha (Lillehaugen et al. 2016): “xoana…noble”

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
92	Place	chacho xi loqua xichayoo	chacho-xi.loqua-xi.chayoo be.located.separate-POSS.forehead-POSS.?	Al700	48v	7
93	Place	cho xiloba	cho-xi.loba behind-POSS.?	Al707	76r	7
94	Place	escooa	escooa	Co721	3r	24
95	Place	lacha lapa quiiaa	lacha-lapa-quiiaa flat.land-flat.land-high.place	Al698	48r	11
96	Place	lachi guia guee	lachi-guia-guee flat.land-high.place-?	Co721	3v	10
97	Place	la chi guia guia	lachi-guia-guia flat.land-high.place-high.place	Co721	3v	15
98	Place	quero pichina	quero-pichina squash-deer	Al642	40r	15
99	Place	queyoo	queyoo	Al707	76r	10
100	Place	quiya yana	quiya-yana high.place-corn	Al642	40r	15
101	Place	rua echezacui	rua-eche-zacui edge-?-?	Al698	48r	10

ID	Category	Toponym	Morphological Analysis	Text Code	Page	Line
102	Place	yobiqui chi	yobi-quichi same[28]-thorn	Al714	72r	27
103	River	quego lace	quego-lace river-?	Al711c	56r	20
104	River	quego-quinaa teque	quego-quinaa-teque river-sown.field-?	Al711c	56r	26
105	Wood	yaca biza	yaca-biza wood-end[29]	Al714	72r	31

[28] Córdova 266r: “mesmo…yobia”
[29] Córdova 064v: “cabo o fin o termino…pizaa”