May Helena Plumb **Personal Statement**

Introduction. My first taste of linguistics came in high school Spanish class. I learned that object pronouns in Spanish come *before* the verb rather than after. Why? I wondered, and how? This was just the first mystery. Many Spanish words and phrases, my Spanish teachers explained, were borrowed from Arabic during the Muslim rule of the Iberian Peninsula. These tidbits encouraged me to explore the vast diversity of the world's language structures. I read about Warao, where the object comes first in the sentence, and the subject last. I learned about regular sound changes, language endangerment, dialectal variation, and ergative case marking. I was hooked. In my freshman orientation survey at Haverford College, I wrote that my dream was to do language documentation in a foreign country.

To my surprise, that's exactly what I did. During my first college class (Introduction to Syntax), Brook Danielle Lillehaugen mentioned that she was looking for a research assistant to study Zapotec languages in Oaxaca, Mexico. I think I might have tripped over desks running up to her after class. Everything I learned about Zapotec fascinated me. The historical record—there is an extant corpus of over 400 Valley Zapotec documents from the colonial period—was particularly interesting. Five years later, I continue to be intrigued by linguistic diversity in Oaxaca and how study of the colonial documents can inform our understanding of the modern linguistic landscape.

Past & Current Research.

Colonial Valley Zapotec. Since May 2013, I have collaborated with Lillehaugen on the Ticha Project, a digital explorer for Colonial Zapotec texts; I am now a coauthor on this project.[1] Ticha's primary goal is to make digital copies of Colonial Zapotec documents (and related materials) freely available online. I have played an active role in many stages of this project, including digitization, linguistic analysis, and XML-encoding. For my undergraduate thesis, I conducted an independent study of conjunction in Colonial Valley Zapotec. I presented this research at two conferences,[2,3] and my article on the topic has been accepted for publication in the *International Journal of American Linguistics*, the top journal for linguistic research on languages of the Americas. My recent research includes a study of resumptive pronouns in Colonial Valley Zapotec relative clauses.[4]

Modern Valley Zapotec. I conducted my first research in Oaxaca as part of a Linguistics Field School run by K. David Harrison, Brook Danielle Lillehaugen, and Theodore B. Fernald (funded by NSF REU Site Grant #1461056). In June 2015, I was a undergraduate participant, and I returned as a student coordinator in June 2016. The primary goal of each field school was the creation and growth of online dictionaries; I am a coauthor on dictionaries for three Valley Zapotec varieties.[5,6,7] We also recorded over 15 ethnographic interviews relating to language ideologies and Zapotec identity. Since 2016, I have supported the Voces del Valle project, an initiative to encourage Zapotec speakers to write their languages on Twitter.[8] I participated in a panel about Voces del Valle at the Coloquio sobre Lenguas Otomangues y Vecinas in April 2018.[9]

I began my independent fieldwork this summer in San Jerónimo Tlacochahuaya (SJT), a town about 21 kilometers east of Oaxaca City. During this trip, I added over 100 recordings to the existing online dictionary[5] and recorded and translated about 20 monolingual descriptions of local traditions and celebrations. The topics of these recordings were based on the interests of my Zapotec teachers; for example, Antonio García Cruz, a farmer who returned to SJT after many years in the United States, offered to describe traditional agricultural practices. My primary Zapotec teacher, Moisés García

1

Guzmán, is particularly interested in understanding the colonial documents. In collaboration with García Guzmán, I created a modern SJT Zapotec translation of a Zapotec testament written in SJT in 1675. I will present this last research at the 2019 meeting of the Society for the Study of the Indigenous Languages of the Americas.

Research Interests & Future Plans. My proposed dissertation research is a synchronic and diachronic description of word order in San Jerónimo Tlacochahuaya (SJT) Zapotec. It is a relatively narrow proposal that encapsulates my primary interests: syntax and historical linguistics. My dissertation project, however, represents only the initial step in my larger research plan. Below I discuss my broader goals and research philosophy.

Descriptive linguistics. I am excited to continue studying Valley Zapotec in both its modern and colonial forms. The colonial documents can serve as a valuable resource in understanding the modern Valley Zapotec languages, and I will make use of this historical record in my dissertation research and beyond. I favor a "bottom-up" approach to language description—that is, I focus on the specific language and community I am working with and I look to naturalistic text corpora to identify important features of the grammar. I recognize the importance and efficacy of involving Zapotec speakers not simply as consultants but as teachers (of Zapotec), students (of linguistics), and co-authors. In my recent fieldwork, I worked collaboratively with speakers to study the language in a way that was meaningful to them. I will continue to adapt my research to reflect the interests and priorities of my Zapotec teachers.

Language documentation. I believe in a holistic approach to language documentation. What does this mean? I begin with the notion that the syntax of a language is best studied with an appreciation of the sociolinguistic context, and vice versa.[10] Furthermore, language documentation is not simply a preliminary step to linguistic description; instead, it should be understood as a separate endeavor, namely the collection of primary data in and about a language, for the benefit of numerous scholarly and lay communities.[11] The documentation I will complete during my dissertation research is only a small drop in this bucket. My longterm goal is to create a comprehensive, integrated description of the language, considering all areas of grammar as well as the connections to culture and history. That is to say, I will reach for the stars. Here is how I will start.

Naturalistic, monolingual texts are one of the most efficient types of documentation. A single narrative might record cultural knowledge, give insight into discourse structure, provide examples of interesting syntactic constructions, and demonstrate how speech sounds are realized in context. However, I will explore diverse documentation strategies, including elicitation, ethnographic interviews, and survey studies. As my research expands, I will seek connections with history, anthropology, and other allied fields.

Language in context. In addition to describing SJT Zapotec grammar, I'm interested in interactions between Valley Zapotec language and identity. For example, how does the dialectal variation in the Tlacolula Valley interact with speakers' town identity? Does this variation affect speakers' opinions of and relationship with people from other towns? Additionally, older speakers have intuitions about how the language has changed in their lifetime, including phonological, lexical, and sociolinguistic changes. How do speakers react to these changes? These are broad questions I plan to explore in my future research.

Language revitalization. The Zapotec languages are highly endangered; in SJT, there are no fluent Zapotec speakers under the age of 40. I can be a valuable ally to indigenous language activists, as academics have access to funding and technological support

unavailable to most endangered language communities. Linguists can also offer advice on orthography, lexicography, and grammar when designing dictionaries and pedagogical materials. I have previously collaborated on revitalization efforts, such as the Voces del Valle project discussed above, and plan to continue supporting community-oriented projects. Language activists in SJT have recently begun teaching SJT Zapotec in two grades at the primary school, and I am in communication with them on how to best support this endeavor.

Accessible research. Linguistics research is too often hidden behind the paywalls and jargon of academic journals. These barriers hinder the development of the field and prevent speakers from accessing valuable research on their language. Although publishing in top-tier journals is required to maintain a competitive academic profile, I will try to balance this where possible by publishing in Spanish and submitting to open-access journals. Using the tools of digital humanities, I plan to make a online "museum" of sorts to present my research in a way that is accessible to multiple stake-holding communities.

Data management. Contrary to the lofty goals of the field, language documentation materials tend to gather dust in offices, waiting to be analyzed. To best serve scholars, language activists, and other stakeholders, linguistic data should be responsibly archived, allowing for reproducibility, proper citation (in alignment with the Austin Principles[12]), and rigorous future research. I practice progressive archiving, that is I archive my raw materials as soon as possible and update records as new analysis is completed. My data will be publically available (with my teachers' permission) at the Archive of the Indigenous Languages of Latin America.

Broader Impacts. When I tell Zapotec speakers that Zapotec was regularly written in the 1600s, they are often astonished. They have been told that Zapotec languages have no grammar, that these languages cannot be written at all. One memorable day in July 2014, I visited a Oaxacan archive with two Zapotec women to view the 17th-century Zapotec testaments preserved there. They saw, for the first time, history of their ancestors, written in their language. Their excitement was tangible. The women immediately attempted to translate the documents, although they struggled to interpret the handwriting and adjust to the phonological changes that had occured in the past 400 years. Seeing their enthusiasm that day drove home for me the importance of my research. By studying Valley Zapotec, I support not only the advancement of linguistics, but the revalorization of indigenous languages and the repatriation of indigenous history.

[1] Lillehaugen, B. D. et al. 2016a. *Ticha: a digital text explorer for Colonial Zapotec, first edition.* [2] Plumb, M. H. 2014. La pragmática de las conjunciones en el zapteco colonial del Valle. COLOV. [3] Plumb, M. H. 2015. The pragmatics of conjunction in Colonial Valley Zapotec. SSILA. [4] Plumb, M. H. 2017. Resumptive subject pronouns in Colonial Valley Zapotec relatives. CILLA. [5] Lillehaugen, B. D. et al. 2018a. *Tlacochahuaya Zapotec Talking Dictionary.* Version 2.0. [6] Lillehaugen, B. D. et al. 2017. *Teotitlán del Valle Zapotec Talking Dictionary.* Version 1.2. [7] Lillehaugen, B. D. et al. 2016b. *San Lucas Quiavini Zapotec Talking Dictionary.* Version 1.0. [8] Lillehaugen, B. D. 2016. Why write a language that (almost) no one can read? Twitter and the development of written literature. *LD&C* 10, pp. 356–393. [9] Lillehaugen, B. D. et al. 2018b. Twitter y las lenguas de Oaxaca. COLOV. [10] Bender, E. M. 2005. On the boundaries of linguistic competence: matched-guise experiments as evidence of knowledge of grammar. *Lingua* 115.11, pp. 1579–1598. [11] Himmelmann, N. P. 1998. Documentary and descriptive linguistics. *Linguistics* 36.1, pp. 161–196. [12] Berez-Kroeker, A. L. et al. 2018. The Austin Principles of Data Citation in Linguistics. Version 1.0.