Fall 11-16-2020

Exploring Language Endangerment and Language Change in Tohono O'odham

Keiko F. Beers
University of New Mexico - Main Campus

Keiko Beers

Candidate

Linguistics

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Rosa Vallejos, Ph.D., Chairperson

Melissa Axelrod, Ph.D.

Melvatha Chee, Ph.D.

Doris Payne, Ph.D.

Ofelia Zepeda, Ph.D.

EXPLORING LANGUAGE ENDANGERMENT AND LANGUAGE CHANGE IN TOHONO O'ODHAM

by

KEIKO FRANCIS BEERS

B.A., Art History, University of Minnesota, 1998
M.A., Linguistics, San Francisco State University, 2006

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Linguistics

The University of New Mexico
Albuquerque, New Mexico

December 2020

DEDICATION

To my big brother, Tom, who supported me when I began to ask questions and encouraged me when I wanted to explore.

ACKOWLEDGEMENTS

This dissertation led me through one of the most challenging adventures that I have yet to experience, and I feel certain that I was able to finish because of the unconditional support and encouragement from a number of people whom I would like to acknowledge.

To begin with, I owe a special thanks to all members – past and present – of the O'odham *Ñe'ok* Revitalization Project. My involvement in this project is where my work among the O'odham community and on the O'odham language began. This dissertation would never have been possible without the support of a handful of members of the O'odham community in particular. I am grateful and deeply indebted to Robert Cruz, who shared countless hours of his time with me working on the O'odham materials on which this work is largely based, and who always made time to teach me about his language, O'odham culture, and history. I also would like to thank the late Philip Miguel, who was always willing to share his expertise on O'odham grammar with anyone who was interested and who generously set aside time during his spring break to work with me on translations. Other community members who contributed to this work or supported it in one way or another are David Garcia, Raphael García, the late Mel Ortega, and Frances Benavidez. Sincere gratitude to you all. I also owe many thanks to Jacob Franco Hernandez who provided Spanish translations for several of the texts that I utilize in my dissertation and who spent many hours Skyping with me from Sonora to talk about O'odham structure. Additional members of the O'odham *Ñe'ok* Revitalization Project from UNM who directly contributed to processing some of the texts that I include in my dissertation are Melissa Axelrod, Serena Kavanaugh, Laura Hirrel, Ginnie Kerfoot, and Andrés Sabogal. To Lo:la (Laura), Jinisk (Ginnie), and Antpik (Andrés), in particular, beyond transcribing and translating that old, crackling wax-

cylinder recording with me, I have many fond memories of long (and often exhausting) drives to and from Arizona, co-organizing community events and presentations, and staying up late to meet project deadlines – nearly always with the three of you.

Beyond the project, I want to express my sincere gratitude to my committee members. First, thank you to my chair, Rosa Vallejos, whose Language Contact seminar inspired me to explore the primary topic that I address in my dissertation and who provided me with guidance at the very inception of this work. Thank you, Rosa, for pushing me a little harder than I would have pushed myself. Thank you to Melissa Axelrod for not only supporting me while working on my dissertation and as a PhD student over the years, but also for introducing me to Robert and the O'odham group during my first year in the department. Through your example I have learned about how to build caring and respectful relationships with communities as a linguist. Many thanks also to Melvatha Chee, whom I first knew as a fellow graduate student and was lucky enough to later have on my dissertation committee. I appreciate your meaningful comments on my dissertation and for providing perspectives on my own work that I had not considered before. Thank you also to Doris Payne for your kind, helpful, and detailed comments on my dissertation. I'm certain that your feedback greatly improved the quality of my work. Thank you to Ofelia Zepeda for offering invaluable insights on my dissertation as both a Native speaker and scholar of the O'odham language. Much of the work that you had already carried out on O'odham frequently guided aspects of my own work. Finally, I also want to thank all five of you – as women in science – for all that you did before me and continue to do to help transform the field of linguistics into one that is more equitable.

Many thanks also to the following people who provided feedback on topics directly related to my dissertation: Joan Bybee, thank you for your comments on the very earliest stages of my research on progressive constructions, and thank you also to Madeleine Mathiot for corresponding with me at length about progressives in O'odham. My sincere appreciation also to Naomi Shin and Jill Morford for your useful suggestions regarding some of my findings on demonstratives.

There are also a number of faculty in the linguistics department who enriched my learning experience during my studies in other ways. A warm thank you (again) to Jill Morford always being a source of encouragement in both my teaching and in my language work even though I was never a student in any of your classes. Thank you also to Caroline Smith for being a wonderful professor. I found your classes to be some of the most rewarding that I took while a student. Dawn Nordquist, thank you for asking me to collaborate with you on designing the online Ling 101 course. I learned many useful teaching methods from you. Sherman Wilcox, thank you for your friendship and mentorship over the years. Bill Croft, thank you for your many engaging ideas and for your useful feedback on my dissertation proposal.

I also have a number of colleagues and friends in the department – past and present – who contributed in significant ways to the many positive (and entertaining) experiences I had as a PhD student. There are too many individuals to list here, but I especially want to thank the following people:

Laura Hirrel, I don't think I would have survived without you. Nobody shared in both my pain and successes more than you, including at the very end of my PhD journey when I needed a final gentle push in the midst of all the unsettling events associated with the year

2020. Corrine Occhino, my fellow Midwesterner, thank you being a source of positivity at all times and for reminding me not to be so hard on myself. Ricardo Napoleão de Souza, you always put things into perspective in such a meaningful way, and I learned a lot from you about self-care and work-life balance. Susan Brumbaugh, you were the first person I met in the department when I had not yet decided which graduate program I would attend. You were warm and inviting and were a large determining factor in my decision to attend UNM, as well as a good friend once I arrived.

Other friends in the department who have played a significant and important part in my graduate school experience: Logan Sutton, Shelece Easterday, Huichieh Hsu, Jalon Begay, Sara Siyavoshi, Benjamin Anible, Evan Ashworth, Chris Peverada, Motomi Kajitani, Ljubica Kostović, Deborah Wagner, Daven Hobbs, Daejin Kim, Josefina Bittar, Sook-kyung Lee, Rebeca Martínez Gomez, Jackelyn Van Buren, Aubrey Healey, Pavlina Peskova, and Karol Ibarra Zetter.

Outside of the department, thank you to Kiyoko Simmons for providing me with the opportunity to mentor and work with several inspiring groups of young undergraduate scholars during my time as a project assistant with the Mellon-Mays Undergraduate Fellowship Program. Thank you also to Stephanie Sanchez for providing a safe and stimulating space for our small 'Women in Academia' group where we could both commiserate about and celebrate our experiences as we all worked towards the completion of our PhDs. I am also grateful to the following organizations who provided financial support for my dissertation: Andrew W. Mellon Foundation, the Latin American and Iberian Institute, and the Center for Regional Studies. Without this support, completing this work would have been much more challenging.

I have also been fortunate to make a number of important friendships outside of the university since moving to New Mexico: Heather Trost, Jeremy Barnes, Rosie Hutchinson, Drake Hardin, Candice Hopkins, Raven Chacon, Nani Chacon, Lauren Kostelnick, Claudia Valdes, William Fowler-Collins, Babak Shahsiah, Mark Rudd, and Marla Painter. All of you have been an inspiration to me in one way or another.

I also have my 'family of others' to thank for their love and support. With most of you, my relationships go back over two decades: Sylvia Algire, Vy Duong, Michael Carreira, Kate Fruchey, Nueng Julakasilp, Eric King, Jessica LaFlamme, Annie Lewandowski, Daniel Lewis, Yasi Perera, Benjamin Piekut, Conor Prischman, Sara Ryu, and Shannon Vogt; also, my dear friends from back home in Minneapolis, Eric Ostby, Michele Zoromski, Cullen Gallagher, Chad Popple, and John Lambe. And of course there are also the Deerhoofians, Satomi Matsuzaki, Ed Rodriguez, Greg Saunier and Deron Pulley. I am fortunate to be a part of such a solid community of loving friends who will always have each other's backs.

Thank you also to my immediate family: my parents, my brothers, Dave, Dale, and Tom, my sister-in-law, Sue, and my nieces and nephew, Hannah, Grace, and Sam, and also the Dieterich in-laws, Dan, Diane, Daniel T., Daniel S. and Peg. I love you all.

Finally, I have two more individuals to thank: John Dieterich, who has been full of patience and whose loving support never wavered while working on my PhD. I am the luckiest person to have you as my life partner. And last but not least, Ida Strawberry 'Tutsa' Sakumura, a.k.a. Rascal, who was literally by my side more than anyone while working on my dissertation. Life is always a little nicer because of you. Thank goodness for pups.

EXPLORING LANGUAGE ENDANGERMENT AND LANGUAGE CHANGE IN TOHONO O'ODHAM

by

Keiko Beers

B.A. Art History, University of Minnesota, 1998
M.A., Linguistics, San Francisco State University, 2006
Ph.D., Linguistics, University of New Mexico, 2020

ABSTRACT

Endangered languages are often assumed to undergo grammatical losses. Grammaticalization – which is concerned with how grammar evolves – has been largely overlooked in such contexts. In this dissertation, I address two questions: 1) Do grammaticalization processes in endangered languages look as they do in more robustly spoken languages, and 2) are there other changes that require explanations outside of grammaticalization?

The language that I consider is Tohono O'odham (Uto-Aztecan), spoken in Arizona (US) and Sonora (Mexico), which has become endangered over the last century due to coercive assimilationist policies. I conduct a comparative analysis of pre-existing oral data over approximately the last 100 years, focusing on progressive and demonstrative constructions.

The findings suggest that severe language shift has had little impact on the evolution of progressives and demonstratives. The structures change in predictable ways or remain stable. Changes that may be due to infrequent language use are within the lexical domain.

Table of Contents

List of Figures

List of Tables

List of Images

Abbreviations

1	first person
2	second person
3	third person
ABS	absolutive
ABST	abstract
ACC	accusative
ADV	adverb
ALIEN	alienable (possession)
ART	article
ASSERT	assertive
ASSUMP	assumptive
AUX	auxiliary
COMPL	completive
COND	conditional
CONC	concursive
CONN	connector
CONCR	concrete
CONT	continuative
COP	copula
DESID	desiderative
DET	determiner
DIST	distal
DISTR	distributive
DNMLZ	denominalizer
DM	discourse marker
DTM	determinative
DUB	dubitative
DUR	durative
EP	epenthetic
FACT	factual (evidential)
FUT	future
HES	hesitation
IE	informal ending
IMP	imperative
INCEP	inceptive
INDEF	indefinite
INST	instrumental
INTER	interrogative
INTS	intensifier
IPFV	imperfective
ITER	iterative
MED	medial
NEG	negative
NMLZ	nominalizer
OBJ	object
PART	particle
PASS	passive
PAST	past
PFV	perfective
PH	placeholder
PL	plural
POSS	possessive
PROG	progressive
PRON	pronoun
PROX	proximal
PUNC	punctual
QNT	quantifier
QUOT	quotative (evidential)
RECIP	reciprocal
REFL	reflexive
RL	relativizer
RMT	remote
REP	repetitive
RES	resultative
SBDR	subordinator
SPEC	specific
STAT	stative
SUB	subject
SUCC	successive
TRNSLOC	translocative
VBLZ	verbalizer

Abbreviated text titles used for examples

Historical texts:

Abbreviations	Full title and source
War	*The Story of War with Mexicans* (Dolores & Kroeber, 1909-1951)
Antonio	*Antonio Lopez's Account of Life as a Child on the Desert* (Dolores & Kroeber, 1909-1951)
Boy	*Narrative of the Boy Who Became a Warrior* (Dolores & Waterman, 1909)
JD	*The Reminiscences of Juan Dolores* (Dolores & Mathiot, 1991)

Midcentury texts:

Abbreviations	Full title and source
Coyote Skunk	*Story of Coyote and the Skunk* (Pancho & Mathiot, 1959)
Coyote Turkey	*Story of Coyote and the Turkey* (Pancho & Mathiot, 1959)
Coyote Quails	*Coyote and the Quails* (Pancho & Mathiot, 1960)
Coyote Chicken	*Coyote, the Cat, the Dog, and the Chicken* (Pancho & Mathiot, 1960)
Mathiot n.d.a	(Examples for Volume 1 of *Tohono O'odham-English Dictionary*)
Mathiot n.d.b	(Examples for Volume 2 of *Tohono O'odham-English Dictionary*)

Modern texts:

Abbreviations	Full title and source
Buzzard	*Hunter Turns into Buzzard* (Cruz, 2012)
Saguaro	*Albino Saguaro* (Tucker, Fitzgerald, Miguel, 2012)
##	(Refers to designated YouTube video number; see Table 20, §4.1.3)

Chapter One: Introduction to the dissertation

1.0 Introduction

This dissertation is concerned with grammatical change specifically in the context of a language that UNESCO classifies as "definitely endangered," meaning that few children learn the language as a mother tongue in their homes (Moseley, 2010). The language in this case is Tohono O'odham (Uto-Aztecan), a cross-border language spoken in Arizona (US) and Sonora (Mexico).

A language can potentially become endangered within a language contact situation in which a large segment of a population gradually shifts from one language to some other socially dominant language, resulting in decreased domains of language use as well as decreased intergenerational transmission. In other words, the language's vitality has become threatened, and the situation is ultimately one of "shifting bilingualism" (Grinevald, 1998, p. 258) and can include a wide range of speakers of varying fluency levels who "can be plotted on a continuum" (p. 259).

More often than not, *language endangerment* or the often-used term *language obsolescence* are synonymous in many people's minds with *decay*, *erosion*, *attrition*, *loss*, *simplification*, etc. They are usually associated with a "reduction of one kind or another…" (Dorian, 1978, p. 590). In some cases, this may be due to the fact that speakers of the language, including those who are fluent, have increasingly fewer opportunities to use it, using a socially dominant language more regularly. The situation may lead to various patterns of change.

Some examples of areas that previous studies have identified as vulnerable in these situations are the lexicon (Dixon, 1972; Schmidt, 1985), nominal classifiers (Nettle &

Romaine, 2000), pronominal systems (Bavin, 1991), and case marking systems (Dorian, 1981; Schmidt, 1985). Additionally, phonological systems may contract (Campbell & Muntzel, 1989; Tsunoda, 2005), analytic structures tend to be favored over more synthetic ones (Flores Farfán, 2004; Mithun, 1989; Romaine, 2010), and the syntax may also simplify over time (Hill, 1983; Schmidt, 1985; see Romaine, 2010 for synopses of several of the aforementioned studies). In other words, studies on the subject matter usually aim to identify which facets of language tend to be susceptible to "structural and stylistic attrition" (Romaine, 2010, p. 325).

A different topic that has received far less examination in the language endangerment literature are the processes associated with grammaticalization, whereby lexical items or, more specifically, lexical items within constructions gradually become increasingly more grammatical over time (e.g. Bybee, Perkins, and Pagliuca, 1994; Heine, Claudi, & Hünnemeyer, 1991; Hopper & Traugott, 1993; Kuryłowicz, 1975; Lehmann, 1985). This topic has been overlooked possibly because studies of shifting languages tend to focus on the issues listed above whereas grammaticalization is concerned with how languages develop or evolve; however, this domain of study could offer valuable insights into how severe language shift may or may not impact the development of grammar rather than focusing solely on what may be lost. At this point, we have at our disposal a number of cross-linguistic, typological studies which provide hypotheses about how particular constructions tend to evolve over time based on both synchronic and diachronic evidence. If we have a good understanding of – or at least clear hypotheses about – how specific constructions are expected to evolve in more robust languages, we can then utilize the same framework to examine to what degree less vibrant languages align with those expectations.

Based on this premise, in this dissertation I address the following broad **research questions**: 1) Do grammaticalization processes in scenarios of severe language shift (i.e. in endangered languages) look as they do in more robustly spoken languages or do they differ to some extent and, 2) are there other types of changes that might require explanations outside of grammaticalization?

In order to address these questions, I conduct a usage-based examination of patterns of grammatical change in Tohono O'odham (O'odham) over the span of approximately the last 100 years. It is during this period that the O'odham on the US side of the border (along with all other Native American communities) have been the target of specific coercive measures leading to the gradual replacement of their language with English. More specifically, I utilize pre-existing oral data from three different points in time – the early 1900s, the late 1950s to the early 1960s, and the early 2000s. This allows me to carry out a longitudinal "real-time" comparison (Labov, 1972) of how the grammatical constructions under examination are used in the language. All of the materials are produced by speakers who are assessed to be fluent speakers of O'odham, including those from the early 2000s (see §4.2). I made this decision in order to eliminate effects of partial acquisition from the study and to focus on fluent speakers who have increasingly fewer opportunities to use O'odham in comparison to those at the turn of the twentieth century. Beginning around the 1950s, while children at that time likely learned O'odham as a first language at home, they frequently learned English once they attended boarding schools (see §2.4). Today, on O'odham reservations, English is heard far more frequently at meetings and conferences, in businesses, on the street, in schools, etc. even if O'odham speakers are present.

The two specific areas of grammar that I have chosen to examine are **progressive constructions** and **demonstratives**. I have selected these because both have been broadly examined in previous cross-linguistic, typological studies which in turn have established hypotheses regarding their evolution within the framework of grammaticalization theory. Their developments have not, however, been explored in detail as topics of study in their own right over a specific period of time during which a population has shifted from their original community language to another more socially dominant language.

Bybee, Perkins, and Pagliuca (1994) propose that as progressive constructions evolve and expand in use, they can come to express "an ongoing action or state of affairs" (p. 136). In the case of demonstratives, Diessel (1999) offers an account of how they may proceed down multiple pathways, developing into a range of more grammatical functions such as third-person pronouns, specific indefinite articles, connectives and markers of relative clauses (i.e. determinatives), among others.

I make no *a priori* assumptions about what the grammaticalization of these constructions will look like in O'odham; however, I propose a few possible hypotheses below regarding what the examination could reveal.

One possibility is that progressive and demonstrative constructions may show signs of changes that are disruptive in nature to the grammaticalization processes; that is, a given construction may display types of changes that are not in line with what we currently know about how progressives and demonstratives in this case tend to evolve in vibrant languages. This hypothesis is based on previous findings which suggest that languages used in decreased domains tend to show losses of one type or another, for example, the levelling of verb paradigms (Campbell & Muntzel, 1989; Dorian, 1976, 1978), the loss of noun case marking

and classification systems (Schmidt, 1985), and the decreased use of embedded clausal constructions (Hill, 1978).

It may also be the case that progressive and demonstrative structures will evolve as expected and/or appear stable over time. This would suggest that some constructions may be more impervious to the effects of severe language shift.

Another possibility that must be taken into consideration given the nature of contact is that some constructions may suggest a diffusion of patterns from English, a topic that previous contact studies have addressed, including in some case studies of endangered languages (see §1.2 below).

I will present more specific hypotheses in the context of the two studies on progressive constructions and demonstratives themselves.

My final goal for this dissertation is more pragmatic in nature. In collaboration with O'odham speakers/linguists, a final outcome of this dissertation is to make the language records, particularly from the early 1900s, more accessible to the O'odham community, including members who do not speak O'odham, on both sides of the US-Mexico border. In these culturally rich materials, O'odham speakers describe important events and personal memories from around the turn of the 20th century. These materials have up until this point remained in language archives and/or in difficult-to-use formats such as hand-written records or, in one case, a wax-cylinder recording. They have also, to my knowledge, never been fully translated. A concrete outcome of this dissertation is a collection of fully processed materials. They have been transcribed where needed, typed, converted into the Alvarez-Hale writing system, and translated into both English and Spanish (for more details about the collaborators and how the language records were processed, see §4.3.1).

Lastly, a comment on terminology: throughout this work, I refrain from using the term *obsolescence* to refer to endangered languages except within direct quotes. In my view, *obsolescence* is strongly associated with the deterministic view that the community *will* at some point cease speaking the language entirely. By definition, this term describes something that is in the midst of becoming obsolete. This belies the fact that many communities are working to reverse language loss trends by creating a new generation of first-language learners and/or second-language learners. The term *endangered* or even *threatened* is more apt in simply providing a description for the current status of a language which communities can address in the manners that they judge to be most appropriate.

The organization of this dissertation is as follows: in the remaining sections of this chapter, I present a brief overview of grammaticalization theory, while the subsequent sections deal more directly with the issues addressed in the literature surrounding language endangerment and shift. In Chapter 2, I describe in greater detail some background information on O'odham, including its genetic classification and language vitality. I also provide a synopsis of the historical events that have led to the current status of the O'odham language, and I establish a language shift timeline which is based on various historical figures' observations and statements regarding language use among the O'odham community at different points in time. Chapter 3 provides an overview of the structure of O'odham in order to familiarize the reader with the basic typological profile of the language. Chapter 4 is dedicated to the data methodology. I describe the materials on which this study is based, including information about the types of discourse that are included in this study, what is known about the speakers, and how the materials are processed in terms of glossing, producing translations, etc. Chapter 5 includes the first of two areas of grammar that I

examine – progressive constructions – while Chapter 6 addresses my analysis of demonstratives. Finally, in Chapter 7, I bring the findings together to offer some final concluding remarks regarding the extent to which language shift has or has not impacted these grammaticalization processes.

1.1 Background on the theoretical framework of grammaticalization

To begin with, I briefly describe the primary tenets of grammaticalization theory, the framework on which this study is based. Grammaticalization theory offers a detailed account of how and what aspects of language may change over time (e.g. semantic content, and phonological form) as grammar evolves. The hypothesized principles for grammatical change guide my own analysis of language change in O'odham. More specific details with respect to the grammaticalization of progressive constructions and demonstratives – the two areas of grammar that I am examining in this study – will be discussed in greater detail in Chapters 5 and 6.

From a theoretical standpoint, grammaticalization is concerned with the cross-linguistic processes involved in the creation of grammar, viewing grammar as "highly susceptible to change and highly affected by language use" (Bybee, 2003a, p. 145) rather than treating it as an autonomous and closed off system. Theoretically, grammaticalization gives explanatory power to the argument that it is shared human cognition and experiences in the world that are the driving forces behind the creation of language structure. If universal patterns in the evolution of grammar can be identified, as proponents of grammaticalization theory argue, then this evidence provides major support for a usage-based view of language.

Grammaticalization as a process entails the gradual development whereby lexical items or a group of lexical items become increasingly more grammatical over time (Bybee,

2003a, p. 146; Bybee, Perkins, & Pagliuca, 1994, p. 4; Heine, 2003, p. 575; Heine, Claudi, & Hünnemeyer, 1991, p. 3; Heine & Kuteva, 2003, p. 529; Hopper & Traugott, 1993, p. 2; Kuryłowicz, 1975; Lehmann, 1985, p. 303; Meillet, 1912, p. 131). The process involves a "long chain of developments" (Bybee et al., 1994, p. 5) where lexical morphemes gradually become "more frequent and general in meaning, gradually shifting to grammatical status, and developing further after grammatical status has been attained" (p. 5). Central to this definition is the added observation that lexical items do not enter into grammaticalization as isolated entities but in the context of specific constructions (Bybee, 2003a, 2003b; Heine, 2003; Lehmann, 1992; Traugott, 2003).

The original lexical source out of which a grammatical morpheme evolves tends to be general or basic in meaning (Bybee et al., 1994, p. 5; Hopper & Traugott, 1993, p. 97). For example, verbs such as *go* or *come* are more likely to grammaticalize than verbs with more specific and narrow meanings such as *march, plod, sprint, stumble*. After a lexical item enters into a grammaticalization path, more concrete meaning is lost while more abstract meanings remain (Bybee et al., 1994, p. 6; Heine, 2003, p. 579; Hopper & Traugott, 1993, pp. 3-4). Put another way, lexical items undergo **semantic reduction** (also referred to as **desemanticization** and **bleaching)** in which concrete meanings that are associated with particular forms are reinterpreted as having more abstract meanings. Semantic abstraction in turn allows a morpheme to be used in a wider range of contexts. The implication is that more concrete meaning will be associated with earlier stages in the process and more abstract meanings with more advanced stages.

In addition to semantic reduction, **inference** also plays a crucial role in semantic change during grammaticalization.

> In this type of change, a gram [i.e. grammatical morpheme] that often occurs in an environment in which a certain inference may be made can come to be associated with that inference to such an extent that the inference becomes part of the explicit meaning of the gram (Bybee et al., 1994, p. 25).

That is, not only does grammaticalization entail semantic reduction but also the addition of inferred and sometimes abstract meanings (also see Heine, 2003, p. 579; Hopper & Traugott, 1993, pp. 3-4; Traugott & König, 1991).

Semantic reduction is also frequently mirrored by **phonological reduction**; the more general a morpheme becomes in meaning, the more frequently it will be used as it evolves, and the more phonological reduction the morpheme will tend to undergo (Bybee & Pagliuca, 1985, p. 60). In fact, it has been noted that grammatical forms tend to be shorter in segmental length than lexical items (Bybee et al., 1994, p. 19). Furthermore, highly frequent forms tend to be the shortest (Zipf, 1935 cited in Bybee et al., 1994, p. 19). Finally, once a morpheme has semantically and phonologically reduced, it may also begin to fuse with other surrounding material, possibly becoming a clitic or an affix.

This evolution from lexical to more grammatical is also argued to be **unidirectional** in that once a word or a collection of words evolves towards more grammatical or abstract meanings the reverse is rarely observed (Bybee et al., 1994, p. 13). The grammaticalized item is unlikely to regain either its previous semantic content or the phonological segments that were lost along the way, and phonological fusion is unlikely to reverse its course. Hopper and Traugott (1993) summarize the unidirectionality hypothesis in the following manner: "The basic assumption is that there is a relationship between two stages A and B, such that A occurs before B, but not vice versa" (p. 95). The process, however, is not deterministic. Once

a word or phrase within its construction begins to grammaticalize, it need not come to complete fruition, arriving at the endpoint of its grammaticalization pathway (p. 95). In the context of threatened languages, one question to ask is whether decreased use of a language is one possible explanation for why a construction may not reach the end of the grammaticalization pathway down which it travels.

Next it has been observed based on cross-linguistic observations that similar lexical sources or constructions tend to follow similar **universal paths** of development in the world's languages (Bybee et al., 1994, p. 14). Hopper and Traugott (1993) describe these paths as **clines** "along which forms evolve, a kind of linguistic 'slippery slope' which guides the development of form" (pp. 6-7). Relatedly, it is hypothesized that the original meaning of the lexical source that enters into grammaticalization determines its likely evolutionary course and its resultant grammatical meaning, a concept which Bybee et al. (1994) refer to as **source determination** (p. 9). Source determination and unidirectionality combined "predict that there will be some cross-linguistically similar paths for the development of grammatical meaning" (p. 14); see also Bybee, 2003a, p. 145). While unique, language-specific grammatical developments are entirely possible, the cross-linguistic tendency is that similar lexical sources tend to grammaticalize and display similar paths of change (Bybee et al., 1994, pp. 14-15). The implication of these cross-linguistic trends is that such universal patterns are driven by shared human experience and cognition.

Finally, given the cross-linguistic evidence that grammaticalization occurs along universal pathways beginning with similar source constructions, it is common to find traces of those original sources or a **retention of earlier meaning** in grammaticalized forms (Bybee et al., 1994, pp. 15-16; see also Hopper & Traugott, 1993, pp. 6-7). The various stages

leading from lexical to grammatical are gradual enough that their development will be synchronically detectable within a language. If grammaticalized forms include traces of earlier source meanings, these clues can shed light on the possible sequences of change along a grammaticalization pathway "as though they were links on a chain, one having given right to another" (Bybee et al., 1994, p. 17). Trace meanings can assist in understanding the different stages of development from a lexical item with concreate meanings on one end of the continuum to more abstract and grammatical meanings on the other end, for example.

The recognition of the gradualness of grammatical change, the various clues that are available in languages for detecting change (semantic, phonological, etc.), the unidirectionality of grammatical development and the hypothesized universality of how grammar evolves in languages will inform my own analysis of progressive constructions and demonstratives in O'odham. Grammaticalization offers a rich and detailed framework on which to carry out an exploratory study of language change in O'odham and to examine the possible impacts of severe language shift and language contact on grammaticalization processes that have been attested in previous typological studies.

1.2 Language change in the context of a shifting linguistic landscape

As noted earlier, this dissertation examines the ramifications of decreased domains of use and severe language shift on the ongoing evolution of grammar. More specifically, it examines progressive and demonstrative constructions and whether or not they display patterns of change that would be expected based on previous cross-linguistic observations in grammaticalization studies. Given the relatively recent history of the O'odham community, this work is ultimately concerned with a language in an intense situation of contact. The O'odham community has contended with various infiltrating groups in the region over the

last few centuries; however, over approximately the last 100 years, the greatest external force for O'odham people residing in Arizona in particular has been English-speaking colonizers of the region (see §§2.3.4 and 2.3.5). This situation of contact and colonization has been the primary driving force behind the shifting linguistic landscape from O'odham to English. With this in mind, different possible factors, including those that have previously been observed in contact scenarios, must be considered with regard to language change in O'odham. For example, a fair amount of information is available on contact phenomena such as diffusion of grammatical patterns from one language to the next and the gradual loss of language structures. While the identification of these types of change is not the primary goal of this study, they must be considered as *possible* explanations for the types of changes that may appear. In the following sections I provide an overview of topics that both language contact and language endangerment studies frequently address in the literature.

1.2.1 Implications of language endangerment due to gradual shift

Previous studies suggest that the types of changes that are observed in an endangered language depend on how advanced language shift is in the community. This discussion is relevant especially in scenarios of gradual shift which take place over the course of multiple generations. In these cases, the endangered language "typically goes through a period when it is not used for all the functions and purposes it was previously" (Romaine, 2010, p. 324), which has specific ramifications in terms of language change. In this section, I briefly summarize the possible stages of gradual shift based on previous descriptions in the literature. Next, based on these guiding principles, I asses the stage of language shift that the O'odham community appears to be in today. Finally, I discuss the implications of gradual shift as discussed in the literature in terms of language change and how these implications

support my motivation for looking at the speech production of fluent speakers of O'odham rather than partial speakers or what Dorian (1973) first referred to as "semi-speakers" (p. 417) and described as persons who never fully acquire a given language.

To begin with, a language endangerment situation can include a wide range of speaker types from those who are monolingual in their original community language to those who may only know a handful of vocabulary words. (A clear example of this type of spectrum is described in concise detail in Schmidt, 1985, for Dyirbal). In other words, language endangerment which develops gradually over time is best described as a continuum with different stages which may or may not result in the complete decline of the language (i.e. the outcome is not necessarily deterministic). Campbell & Muntzel (1989) observe that gradual-shift situations usually involve an intermediate stage in which "the dominant language comes to be employed by an ever increasing number of individuals in a growing number of contexts…. This situation is characterized by a proficiency continuum determined principally by age (but also by attitude and other factors)" (p. 185).

Batibo (1992) offers a detailed description of different phases of language shift that a community may experience in language contact scenarios. The phases are based on the author's first-hand observations of a number of languages spoken in Tanzania. While no two situations are alike in all regards, Batibo's observations provide useful insights that can loosely be applied to other language shift contexts.

(1) Phases of gradual language shift (based on Batibo, 1992, pp. 90-93)

Phase 1: Monolingualism in the community language is the most pervasive (i.e. pre-shift stage).

Phase 2: Emerging signs of bilingualism. The given community feels that they must learn or have been coerced into learning a particular socially dominant language for specific circumstances while continuing to speak their own language within their community, at home, etc. (i.e. the situation is diglossic). There are few signs of "code switching, interferences and linguistic erosion" (p. 90) at this stage.

Phase 3: Bilingualism becomes increasingly more common and more speakers employ the socially dominant language as their primary mode of communication. At this stage, speakers utilize more code-switching, code-mixing, and borrowing, and more elements of the socially dominant language begin to enter into the mother tongue.

Phase 4: The socially dominant language is used in a majority of domains, and a large group of partial-speakers emerges.

Phase 5: The community language is completely (or nearly completely) replaced with the socially dominant language.

While the phases are presented as discrete categories, in reality they represent a continuum and common trends in scenarios of language shift and loss. As Batibo (1992) points out, it may sometimes even be the case that, "some languages may be mentioned as belonging to one phase when actually they belong to another" (p. 90), suggesting the fluidity between phases given that no two language communities are alike in all respects. Along this

continuum, a community may fall more towards one end of the spectrum or somewhere in the middle based on the range of fluency levels of the speakers who reside in the community.

In relation to the O'odham community, relatively recent reports on the status of the O'odham language (see discussion in §2.2) suggest that the community is currently situated somewhere between Phases 3 and 4 on Batibo's (1992) language shift continuum summarized in (1) above. The community has a number of predominantly middle-aged bilingual speakers of both O'odham and English (with some also having knowledge of Spanish). Code-switching and borrowing of terms from English is pervasive in today's speech (see examples (58), (63), (99)-(101), (103)-(104) in Chapter 6) and at least impressionistically, more often than not English is the primary mode of communication in most settings, including on O'odham reservations, in schools, during meetings, etc. This assessment is based on personal observations and conversations that I have had with various community members. Those who fall in the younger age-groups (30 and below) tend to be either partial-speakers or know basic phrases and vocabulary (Tohono O'odham Department of Education & O'odham *Ne'ok* Revitalization Project, 2013, p. 7). Many would probably be regarded as monolingual English speakers.

Batibo's Phase 3 described above has particular significance in that the types of speakers that define it tend to be more susceptible to changes in their original community language. Some of these changes are likely to be due to increased exposure to and the use of a socially dominant language. Linguists concerned with change in endangered languages frequently note borrowing and coding-switching, as Batibo does, as well as a diffusion of grammatical patterns. Winford (2003) comments that as speakers rely more on the socially dominant language for a wider range of their communicative needs, they are likely to end up

being "the chief agents of the diffusion" (p. 260) not only of lexicon but also structure. Similarly, Grinevald (1998) observes that scenarios of gradual language loss which span across generations are most likely "to be accompanied by linguistic change" (p. 259) with the bilingual generation of speakers producing "a somewhat changed form of the language" (p. 259) when compared with the elder and predominantly monolingual speakers in the community, but one "which is still accepted by the whole community" (p. 259). Thomason (2010) additionally notes of fluent bilinguals,

> [T]he first and predominant interference features are lexical items belonging to the non-basic vocabulary; later, under increasingly intense contact conditions, structural features and basic vocabulary may also be transferred from one language to the other (Thomason, 2010, p. 36).[1]

In this study, it is this particular group of speakers that I am concerned with – those who are fluent bilingual speakers of O'odham and English. The data that I utilize provide samples of speech which spans from the early 1900s, to the mid-1900s, and up through the early 2000s. Over this timeframe, bilingualism becomes increasingly more common and the use of English becomes gradually more pervasive across domains of use (see timeline presented in Figure 5, §2.4). I hypothesize that the fluent, bilingual speakers from the 2000s will show the most signs of change in their use of O'odham, changes which cannot be attributed to partial acquisition of the language. What impact this context will have on how the language changes or does not change will be determined through a detailed analysis (see Chapters 5 and 6).

[1] The exception to this, Thomason (2010) points out, would be among communities where borrowing is culturally prohibited or avoided.

1.2.2 Establishing the type of contact scenario

Next, in a study dealing with a language contact setting one must also establish the sociolinguistic status of the different languages involved in the situation in recognition of the fact that "a language does not exist in a vacuum but is the means for communication within a group of people, with a certain political and economic system" (Dixon, 1997, p. 1). It is necessary to determine whether the different groups in the contact setting and in turn the languages that they speak are on equal or unequal footing within the larger community as a whole. The type of contact situation between two or more groups as well as the "intensity of contact" (Thomason, 2010, p. 36) is likely to be relevant in the types of changes that may be observed in the languages. This is not to imply that there exist a set of clear-cut, deterministic criteria with which one can predict the degree or type of changes that will develop, but that social factors clearly play a part and cannot be ignored, and in some cases there are some recognizable tendencies or trends (Aikhenvald, 2010; Brown, 1999; Thomason, 2001; Thomason & Kaufman, 1998; Trudgill, 2010). After describing two basic types of language contact scenarios, I then discuss the most relevant type in relation to the O'odham community.

To begin with, in one type of contact scenario each language in the setting may exist in relative linguistic equilibrium with one another, each language holding relatively equal social status and neither language viewed as more socially dominant than the other. In this type of setting, each community is more likely to be able to maintain their respective languages while at the same time being proficient at some level in the other language in the contact setting (Dixon, 1997); that is, bi- or multilingualism among all parties is more probable. In terms of language change, the linguistic outcome of this scenario is more likely

to involve **multilateral diffusion** (Aikhenvald, 2010, p. 9), meaning that the languages involved may be more prone to equally influence one another and, over time, linguistic patterns may be shared across languages leading to possible convergence.

The other possibility (at least in terms of extremes) is for one language to be associated with more societal power and in turn more social prestige. For example, in circumstances where there are clear "hierarchies of prestige (castes, etc.) and relations of dominance (social and/or political) between languages or dialects" (Aikhenvald, 2010, p. 9), the social stratification is known "to influence the direction of borrowing and diffusion" (p. 9). In cases of unequal power relations between communities, "borrowing from a prestige into a non-prestige language" is more likely to occur; however, it should be pointed out that a rejection of the prestige is also possible precisely as a way to repudiate the socially or politically dominant populations and what they represent. In other words, the language attitudes of the speakers and how they interpret/react to situations of power can also influence the degree of contact-induced change (Thomason, 2001, p. 77).

Nevertheless, this situation of linguistic equilibrium often may result in language shift – the less dominant group gradually adopting or being forced to adopt the socially dominant language in place of their own community language over time out of economic, political, etc. necessity while the socially dominant group is more likely to only speak their own language (Dixon, 1997, p. 110). This process may be gradual, occurring over many generations, or it can be relatively rapid, developing over a couple of generations, as can be the case, for example, with immigrants moving to a new country where a different language is spoken. Language shift is also often observed among Native populations, as in the US where colonizers have infiltrated much of the continent and claimed territories that others already

inhabited. This process of language shift can be observed in the O'odham community in which large segments of the population are either bilingual in O'odham and English or monolingual in English, while few persons today are monolingual in O'odham.

In terms of language change, the outcome is more likely to be **unilateral diffusion** (Aikhenvald, 2010, p. 9), meaning that linguistic changes will tend to move in one direction. Patterns in the socially dominant language, in this case English, are more likely to appear in O'odham for example, but the opposite is less likely to be true or at least not to the same extent.

In the case of O'odham, it is clear that the community is undergoing severe language shift given the fairly recent statistics (see §2.2). It is also evident given the colonial past of the US and their treatment of Native American groups across the country that the situation is not one of equilibrium (see §2.3). Hypothetically then, unilateral diffusion of patterns from English into O'odham is more likely to be observed.

1.2.3 Externally versus internally motivated changes in situations of contact

The previous sections have alluded to changes in contact situations that may be ***externally-motivated*** – changes driven by external sociolinguistic factors which may lead one language to be influenced by another language. In terms of what may be transferred from one language to another, according to Thomason & Kaufmann (1988), contact studies suggest that "any linguistic feature can be transferred from any language to any other language" (p. 14). Language change, however, is additionally driven by internal factors, and how to distinguish between the two motivators of change is often not clear cut. In fact, it may also be the case that external and internal factors work in collaboration to move a particular change forward

(e.g. Dorian, 1993; Heine & Kuteva, 2003, 2005, 2010). In this section I briefly address some issues in the literature regarding externally and internally motivated changes.

Hickey (2012) describes ***internally*-motivated** changes as "[a]ny change which can be traced to structural considerations in a language and which is independent of sociolinguistic factors…" (p. 388). Language-internal processes of language change have sometimes been characterized as *natural* phenomena, "while other changes arise in circumstances where they would not be expected" (Thomason & Kaufman, 1998, p. 22). Mufwene (2007) argues that the view that internal changes are *natural* is highly problematic in that this casts changes stemming from external sources as *unnatural.* In truth, according to Mufwene, all changes in a language are in some manner "external to its structure, lying in the communicative acts of speakers…" (p. 65). Speakers make adjustments to their speech in order to increase their chances of communicating successfully with one another and also in order to meet the constantly changing needs of speakers to express new concepts.

With regard to grammaticalization in particular, traditionally it has been argued that this is "an internal process, and that grammaticalization contrasts with language contact, which is viewed as providing an alternative explanation for language change processes" (Heine & Kuteva, 2010, p. 87). Heine and Kuteva (2010) propose that internally- and externally-motivated, contact-induced changes need not be mutually exclusive processes. They contend that the two motivators "may work in conspiracy with each other" (p. 87) and both are natural drivers of language change.

Heine and Kuteva (2010) refer to grammaticalization processes that occur in one language due to the influence of another language as **contact-induced grammaticalization** (p. 88; see also Heine & Kuteva, 2003, 2005). More specifically this is the process whereby

speakers of one language (subconsciously) create a grammatical category in their own language and with their own phonetic material based on an observed category in another language. That is, contact-induced grammaticalization entails the creation of grammatical meaning in one language based on its existence in some source language. This is in opposition to borrowing a morpheme wholesale with both its phonetic form and meaning intact (Heine & Kuteva, 2010, p. 86). A language may develop a grammatical category based on its existence in another language via universal strategies of grammaticalization or also "by replicating the process that they observe in the model language" (p. 89; see also Heine & Kuteva, 2003, 2005).

Aikhenvald (2010) suggests that contact-induced grammaticalization or what the author refers to as **grammatical accommodation** is one characteristic type of change in shifting, endangered languages (p. 249). In these instances, the presence of a more socially dominant language "enhances the processes of areal diffusion" (p. 249). Aikhenvald makes the additional important observation that diffusion shows us that while endangered languages may show losses, they often also show clear gains as new structures develop in the language. Sasse (1992) also states that at least in the context of fluent bilingual speakers, the original community language often may be altered by the development of "an identical category system, and … structurally similar means to mark these categories" (p. 64) as exists in the language which serves as the source.

That said, it may be difficult to determine which factors motivate a particular example of grammatical change in a contact setting. In circumstances of heavy contact, it is imperative to recognize the possibility that changes that may appear at first glance to be externally motivated (i.e. contact-induced) may in fact be the result of internal processes that

take place in all languages or that the two processes have come together to promote some type of change, as Dorian (1993) points out:

> There is always the possibility that internal pressures within a language structure were lining up in favor of such change already, and the still more likely possibility that internal pressures have combined with external influence from the other language, in the contact setting, to produce the change (Dorian, 1993, p. 136).

Dorian suggests that a multitude of factors may play different roles in the changes seen in contact situations and, furthermore, that "language structures are complex, and the forms which linguistic influence takes may be correspondingly so. There is a danger in looking for cross-linguistic influence in terms of waxing or waning of rather obviously parallel structures" (p. 132). In other words, dichotomies tend to be misleading and promote strictly either/or judgments.

In contact scenarios there is the added difficulty that no two situations are identical in terms of the languages and language types that are involved, the history of the contact situations, the social circumstances of the speakers, as well as the language attitudes of the speakers. As a result, the extent to which previous studies may inform researchers' predictions about what may motivate a particular change is somewhat minimized. Thomason (2001) notes that scenarios and outcomes of contact are, in reality, "very messy" (p. 60), with a multitude of possible factors simultaneously contributing to language change.

Given the presence of possible obfuscating factors, the question remains as to how one is to accurately identify instances of contact-induced, externally-motivated change. Thomason (2010) proposes the following definition for this type of change: "Contact is a source of linguistic change if it is less likely that a particular change *would* have happened

outside a specific contact situation" (p. 32; original emphasis). Thomason reiterates that this is a definition rather than a clear criterion for contact-induced change. However, deductions in favor of contact-induced changes, I argue, can be made in some cases. In the context of grammaticalization as described in §1.1, certain paths of development and change have been well established through cross-linguistic, typological comparisons. These universal paths do not preclude the possibility, however, of "language-specific or even unique instances of grammaticalization which arise by the idiosyncratic selection of source material" (Bybee et al., 1994, p. 15). Instances in which two or more languages in contact show developments that could be labeled as unique to those languages might suggest a clear instance of externally-motivated language change. For example, if English displays a grammatical development that appears to be an anomaly in comparison to what has previously been found in typological studies, and another language in contact with English, for example O'odham, shows the identical anomalous pattern, we can speculate that the change may be due to external factors rather than being internally motivated.

In the present study, while my primary goal is to examine the degree to which specific structures in O'odham have grammaticalized, and to what degree those developments are in line with proposed universal paths of development, I also consider the possibility that some observed changes could be externally motivated. Given the extreme situation of language shift in the community, however, other possible outcomes that are associated with decreased domains of language use must also be considered.

1.2.4 Characteristics of language change in endangered languages

In this section, I discuss previous characterizations of language change in threatened languages. The first characterization is the *rate* at which the changes occur, and the second is the *types* of changes that may occur.

Endangered languages in many regards do not change in manners that differ from vibrant languages (Aikhenvald, 2010; Campbell & Muntzel, 1989; Dressler, 1977; Sasse, 1992; Thomason, 2001). Words can be borrowed, structures manipulated, and morphology may be created or abandoned. In other words, endangered languages may show losses in grammar, but they can also show gains (see Aikhenvald, 2010, pp. 249-250). Given that this is the case, Aikhenvald (2010) argues that the primary difference between changes in endangered and vibrant languages is the rate at which the changes take place:

> The difference between language change in 'healthy' and in endangered or obsolescent languages very often lies not in the SORTS of change, which are typically the same [Campbell and Muntzel 1989]. Rather, it lies in the QUANTITY of change, and the SPEED with which the obsolescent language changes (Aikhenvald, 2010, p. 249; original emphases).

For example, Schmidt (1985) observes in Dyirbal (Pama-Nyungan) of Australia that "vast amounts of change are compressed into a short time span of about 25 years" (p. 213; quoted in Aikhenvald, 2010, p. 249).

Aikhenvald (2010) similarly notes accelerated change in Tariana (Arawakan), a language spoken in the Brazilian Amazon. The speakers live in an increasingly "destabilized" (p. 249) situation of multilingualism where Tucano (Tucanoan) is growing as the more dominant language of the area. Aikhenvald notes, "The more Tucano people speak at home,

the more Tucano-like patterns occur in their Tariana" (p. 249). For example, verbs have grammaticalized into aspect markers in Tariana based on patterns in Tucano over just a few generations. This development of verbs into aspect markers is not in and of itself unusual, but heavy contact and language shift may have played a role in accelerating "intertranslatability between Tariana and Tucano at a morpheme-by-morpheme level" (Aikhenvald, 2010, p. 250). The types of changes Aikhenvald points to here are of the additive type in that new structures are created based on the source languages in the contact setting. These may at the same time co-occur with changes that are subtractive, which are described in what follows.

Partial speakers of a language – speakers who do not have the opportunity to fully acquire their ancestral, community language due to limited exposure to it – may not employ all of the original structures that more fluent speakers do. In this manner, features of the community language may be interpreted as being lost along the way. Fluent speakers of an endangered language due to decreased domains in which they can use it may also show certain kinds of losses. A few examples are discussed below.

It may be the case that specialized grammar that is associated with lesser-used genres of speech show signs of disappearing. This is sometimes referred to as "stylistic-shrinkage" (Campbell & Muntzel, 1989, p. 195). For example, Campbell and Muntzel (1989) at the time of writing note that very few speakers of Pipil utilize specialized literary devices or speech styles associated with highly ritualized contexts, and speakers furthermore do not produce the original passive constructions in the language (p. 195). Hill (1983) similarly notes that remaining speakers of Nahuatl (Uto-Aztecan) no longer utilize the honorific and reverential verbal marking systems (p. 265).

Other subtractive types of changes are also observed. Hill (1978) examines the use of subordinate clauses in Cupeño and Luiseño (Uto-Aztecan) due to language shift. Hill finds that, while speakers of these related languages can produce different types of subordinate clauses when asked to do so, in natural discourse they tend to avoid these clausal types, preferring separate, individual sentences with a single predicate rather than producing embedded clauses (p. 46). In this instance, loss is equated not necessarily with the speakers' inability to use the particular construction but with the decreased frequency with which they use them and their tendency to "avoid the special constructions" (p. 46). In contrast, speakers in recordings from just forty years prior produced subordinate clauses on a regular basis.

Finally, there are the speakers themselves who comment that while they are fluent in their language and the community accepts their command of the language as such, they feel that their version of the language has shortcomings in comparison to their own grandparents' or other community elders' use of the language – a type of "discourse of nostalgia" (Hill & Zepeda, 1998, p. 263). In this regard, the O'odham community is a case in point. I have heard several O'odham people comment that they can't speak the way their ancestors did a hundred years ago even though they are fluent in O'odham. While speakers may not know precisely what has been lost, they intuitively sense that the language has lost *something*.

Given these previous findings, the possible impacts of decreased language use on parts of O'odham grammar cannot be overlooked. In this case, it is necessary to consider in what manners severe language shift might manifest in the evolution of progressive and demonstrative constructions in the language. If a language is used less often than it was at one time, one must ask what the outcomes of this may be in the context of the ongoing evolution of particular grammatical constructions, in this instance progressive aspect and

demonstratives. It may or may not be the case that they evolve through the various hypothesized stages as expected. The most realistic scenario will most likely be a complex one.

1.2.5 Speakers of endangered languages as actors in language change

As the previous section suggests, the overall picture in most cases of language shift and contact likely consists of both losses as well as additions or gains. The outcome of these changes is likely the creation of a new version of a language that combines elements of all of the languages that serve as input. With this in mind, in this last section I describe two pervasive views in the literature towards the changes that can develop in endangered languages.

The first view focuses on speaker "competence" (Hill, 1983, p. 258) or speakers' knowledge of their community language. This view tends to focus on knowledge deficiencies in the endangered language, viewing the newer version of the language as '"broken-down' or as 'eroded' realizations of an ideal competence" (p. 258). The second view focuses on what speakers *do* with the language; that is, it treats speakers as (subconsciously) reacting to their surrounding social context and utilizing the linguistic tools that are available to them. The latter view is arguably more beneficial, especially in light of communities who are working to revitalize their languages and establish positive attitudes towards their respective languages.

From this vantage point, linguists treat "the study of usage as an accomplishment of speakers, inextricably intertwined with their sensitivity to social contexts … (Hill, 1983, p. 258), and the speakers manipulate "symbolic materials available from a wide range of codes in constructing a changing society" (p. 258). The alterations that develop in the language may

lead to a new version that differs – even dramatically – from how it was spoken in the past (e.g. Bilingual Navajo; see Schaengold, 2003, 2004) but, as Heine & Kuteva (2005) put it, "is not necessarily less coherent, less 'systematic,' or less complete …" (p. 34). In their study of language contact, Heine and Kuteva find more evidence that bilingual speakers are more aptly described as linguistic actors (pp. 34-35). The authors invoke Hagège's (1993) concept of "language building" where "users of language … build linguistic structures in order to meet the requirements of communication" (p. ix) during interaction with one another. Speakers manipulate the linguistic systems that are at their disposal in the manner that will benefit them the most based on their current sociolinguistic needs. While losses may be a part of this process, new additions are also likely to take place.

That said, the remorse that speakers of endangered languages often feel about the cultural knowledge that is being lost along with their respective languages cannot be minimized. Furthermore, the historical and societal pressures that nearly always lead to a language's threatened status should not be trivialized. Their impacts have been detrimental in a number of ways – both cultural and linguistic. However, the view that speakers are creative actors in the (re)construction of their linguistic systems may be more beneficial in terms of a community's linguistic psyche. This may especially be true for speakers who remain and continue to use their language, even if it is in fewer settings than in the past, for heritage speakers who have had the opportunity to only partially acquire their language, and finally for individuals who aim to learn their language as a second language.

1.3 Summary

In this chapter, I have briefly introduced the issues and questions that this study sets out to address: what the grammaticalization processes associated with the evolution of progressive

constructions and demonstratives look like in O'odham over the last approximately 100 years and what impacts, if any, severe language shift has had on those processes. Next, I have described the basic tenets of grammaticalization theory, which serves as the theoretical framework on which I base the two studies. Subsequently, I have addressed topics that more directly relate to language shift. I discussed the commonly observed stages of gradual shift, where O'odham fits on this shift continuum, as well as my motivations for considering fluent speakers of O'odham instead of partial speakers. I described the different types of possible language contact scenarios (e.g. languages with equal versus unequal statuses in a community) and what the linguistic outcomes of languages in those situations tend to be in terms of multilateral versus unilateral diffusion. These variables also factor into the likelihood that a community will be able to maintain their language or if they will instead experience gradual shift towards a more socially dominant language, as is the case among O'odham people. Next, I also have briefly addressed externally and internally motivated changes especially as they relate to grammaticalization, which is often argued to be a language-internal process. Others, such as Heine and Kuteva (2003, 2005, 2010) argue that the two processes can in some cases go hand-in-hand. Finally, I have addressed the types of changes that endangered languages may undergo. While they are more often described in terms of their losses, studies indicate that they may also show gains and new creations. In this regard, they do not differ from more robust languages. Instead, it may be the sped-up rate of the change that differs as well as the greater amount of change. While some types of losses may tend to be more strongly associated with endangered languages, such as "stylistic shrinkage" (Campbell & Muntzel, 1989, p. 195) and a preference for structures that may be easier to process (e.g. Hill, 1978), more often than not, the picture is much more complex,

involving gains as well as innovations. From this vantage point, speakers draw from and manipulate the language material that is available to them.

Chapter Two: Tohono O'odham background

2.0 Introduction

In this chapter I provide various types of background information on Tohono O'odham. Section 2.1 includes a brief description of its genetic classification and the regions where the language is spoken. Section 2.2 describes the current vitality of Tohono O'odham, and §2.3 provides a brief history of both contact between Tohono O'odham people and various outsiders and the socio-political policies affecting the Tohono O'odham and their ability to maintain both their language and aspects of their culture. In §2.4, I establish a language-shift timeline based on how different people describe Tohono O'odham language use in various historical writings at different points in time during the twentieth century. Section 2.5 addresses the current language policy and language attitudes of members of the Tohono O'odham community. The final section, §2.6, provides some brief concluding remarks.

2.1 Genetic classification and locations spoken

Tohono O'odham (formerly referred to by the exonym Papago), which translates as 'Desert People,' is one of two major dialects of the O'odham language. The other major dialect is Akimel O'odham (formerly Pima), meaning 'River People.' The two dialects are mutually intelligible. Tohono O'odham can be further broken into the subdialects Totoguañ, Kokololodi, Gigimai, Hu:hu'ula, Huhuwoṣ, and Kohadk (Saxton, Saxton, & Enos, 1983). The O'odham language as a whole belongs to the Southern branch of the Uto-Aztecan (UA) family, and more specifically to the Tepiman (Upper Piman) sub-branch (see Figure 1). UA as a whole includes a wide-reaching expanse of approximately 30 languages of the Americas extending from Idaho in the North down to El Salvador in the South, and from the coast of

Northern Uto-Aztecan

1. Numic
    a. Western Numic: Mono, Northern Paiute
    b. Central Numic: Panamint, Shoshone, Comanche
    c. Southern Numic: Kawaiisu, Ute (Chemehuevi, Southern Paiute, Ute)
2. Tübatulabal
3. Takic
    a. Serrano-Gabrielino
        i. Serranan: Serrano, *Kitanemuk
        ii. *Gabrielino (Gabrielino, Fernandino)
    b. Cupan
        i. Cupeño, Cahuilla
        ii. Luiseño
4. Hopi

Southern Uto-Aztecan

5. Sonoran
    a. Tepiman: Upper Piman (**Tohono O'odham**, Akimel O'odham), Pima Bajo, Northern Tepehuan, Southern Tepehuan
    b. Taracahitan
        i. Tarahumaran, Rarámuri (Tarahumara), Guarijío
        ii. Opatan: *Opata, *Eudeve
        iii. Cahitan: Yaqui, Mayo
    c. Tubar
6. Corachol-Aztecan
    a. Corachol: Cora, Huichol
    b. Aztecan
        i. *Pochutec
        ii. General Aztec: Pipil, Aztec (several varieties)

Figure 1. Uto-Aztecan language family (based on Haugen, 2008b; Miller, 1984, p. 21).

* = extinct

California in the West over to Oklahoma in the East.

Tohono O'odham in particular is spoken on both sides of the US-Mexico border in southern Arizona and northern Sonora, Mexico. The Tohono O'odham community in its entirety comprises a single unified tribe residing throughout the Sonoran Desert with a shared history and identity. Preceding the international border, US reservations, and land allotment (see §§2.3.3 and 2.3.4), Tohono O'odham lived on an expansive territory referred to as the Papaguería (see Figure 2 below). It encompassed the regions between the Gulf of California to the west and San Pedro River east and Central Arizona down into Sonora, Mexico. It was not until the ratification of the Gadsden Treaty

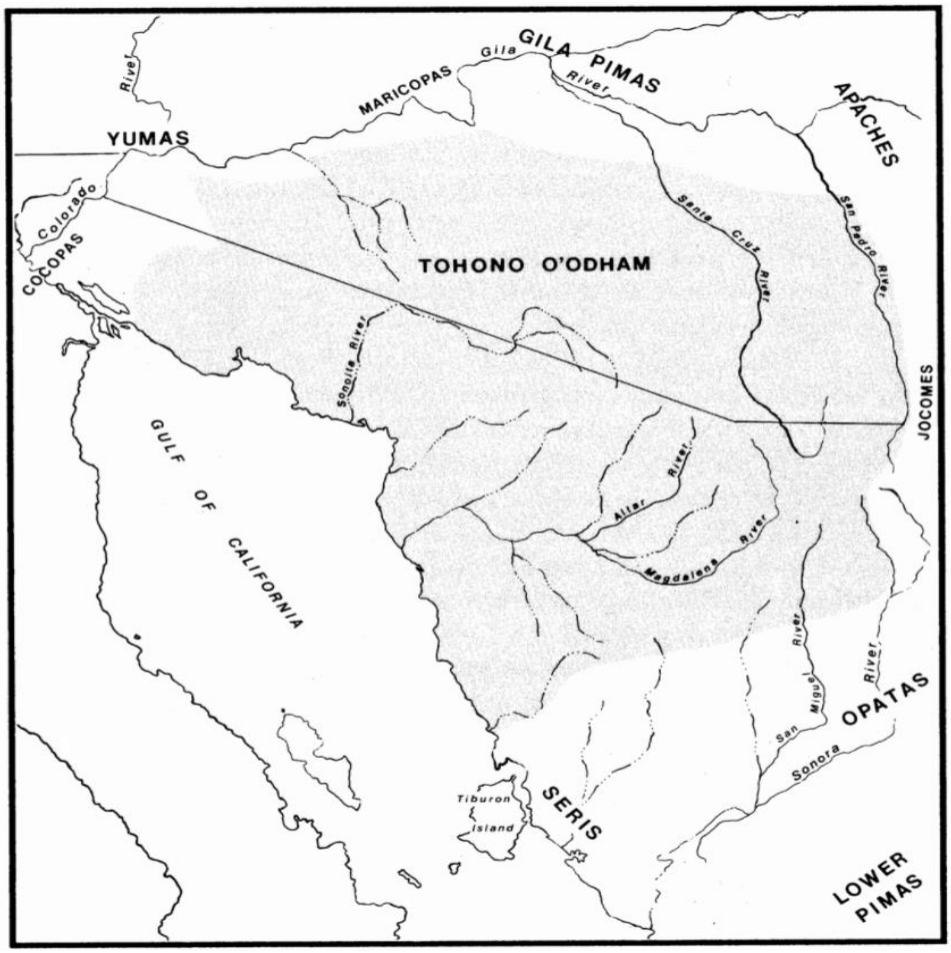


Figure 2. The Papaguería – Tohono O'odham and their neighbors, c. 1700 (Erickson, 1994, p. 17; used with permission, U of Arizona Press).

in 1853 (see §2.3.3), which resulted in the US purchase of present-day southern Arizona from Mexico, that the tribe was divided into North and South (Fontana, 1974, p. 184) as depicted by the line which cuts across the map in Figure 2 above. This land division, Madsen (2005) observes, has "resulted in vastly different experiences on either side of the border" (p. iii) and "being born on one side of the border has major social, cultural, and economic ramifications" (p. iii). Most obviously, while the Tohono O'odham in Arizona have shifted from being predominant speakers of O'odham to predominant speakers of English, the community residing in Sonora has shifted towards Spanish. Overall, the O'odham in the South have experienced far greater integration into Mexican society resulting in far fewer speakers of O'odham than in Arizona (see §§2.2 and 2.3.3).

It should also be pointed out that the Tohono O'odham spoken in the North versus the South do display differences in their grammars. To name just a couple of differences, Shaul (2012) states that speakers in Sonora show a more restricted use of the imperfective auxiliary *'o* (described in §3.3), and a past progressive marker has developed in Sonoran speech that is not used for the most part in Arizona. There are some phonological differences, as well. For example, Shaul points out that the vowel length contrasts (see §3.1.1) in the language may be disappearing in Sonoran speech. It may be the case that physical separation and prohibitions on movement across the Sonoran international border have promoted the beginnings of language divergence, but no study has established this with certainty. It is also plausible that some divergent changes were developing prior to this period. Whatever the cause or causes, these noted differences have led some to distinguish between the Tohono O'odham spoken north and south of the border, referring to the variety spoken in Mexico as Sonoran O'otam. (e.g. Moseley, 2010; Shaul, 2012). Regardless, many speakers of the latter identify as

Tohono O'odham, and those Tohono O'odham in Arizona also recognize them as such.

Today the Tohono O'odham Nation in the US is comprised of four federally recognized land bases in southern Arizona, the largest of which is the Tohono O'odham Reservation, followed by the San Xavier Reservation, and the San Lucy and Florence Village Districts (see Figure 3 below). The land bases are located roughly south of the Gila River and the Salt River tributary. While there are no federally recognized O'odham land bases in Mexico, various communities are scattered across northern Sonora, Mexico (see Figure 4 below), in Pitiquito, Caborca, Altar, Sáric, Sonoyta and Puerto Peñasco (Franco Hernández & Estrada Fernández, 2016).

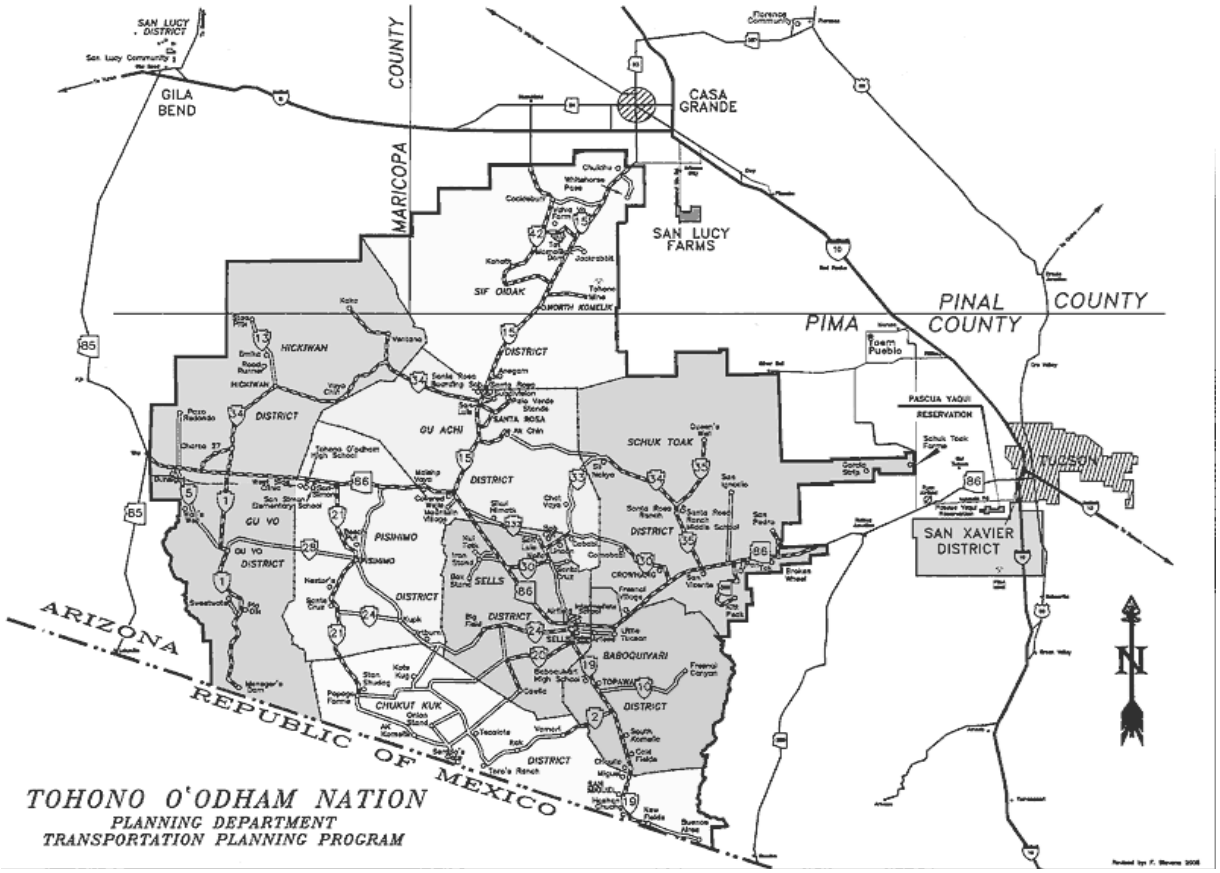


Figure 3. Tohono O'odham Nation Reservations (Tohono O'odham Nation, 2016, November 10).

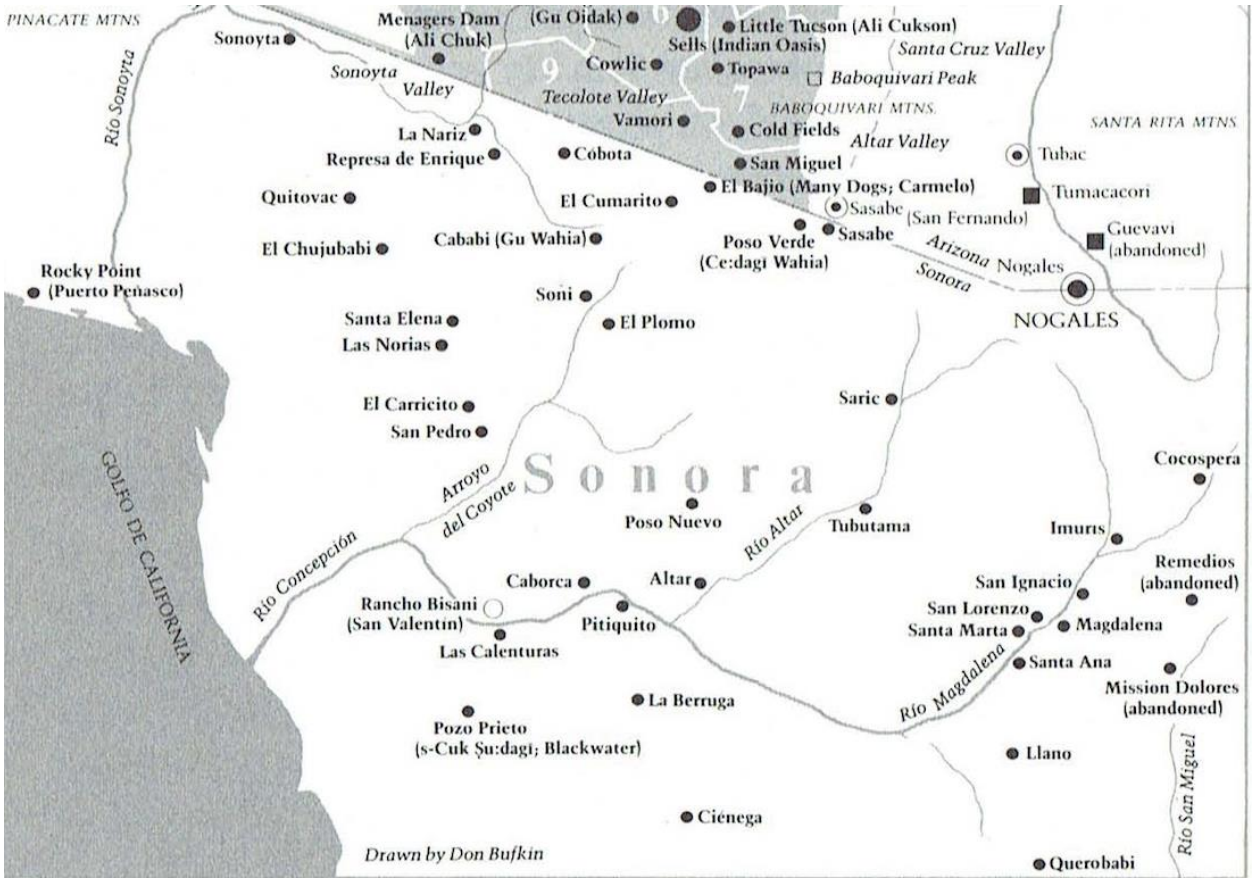


Figure 4. O'odham region of Sonora, Mexico (map sample from Fontana, 1989, p. ix; used with permission, U of Arizona Press).

2.2 Language vitality

According to Madsen's (2004) Tohono O'odham Nation Comprehensive Study, most speakers of Tohono O'odham in Arizona are above the age of forty (as cited in Fitzgerald, 2010), and inter-generational transmission is rare. Approximately "10 percent of 18 and 19-year-olds can speak O'odham, and Nation members aged 40 and older are the only age groups in which 50 percent or more can speak O'odham" (Madsen, 2004, p. 232 cited in Fitzgerald, 2010). An unpublished report based on a survey conducted in 2011 and 2012 found that out of 328 people surveyed, 34% were "very comfortable speaking and understanding O'odham" (Tohono O'odham Department of Education & O'odham *Ñe'ok* Revitalization Project, 2013, p. 3). Those who reported the latter were generally 50 years of

age or older (p. 7). Of those 20 years of age and younger, "This age group generally reports they cannot speak O'odham without resorting to using English and [are] unable to follow a conversation" (p. 7). These most recent numbers suggest that O'odham is currently in decline.

For all O'odham varieties in Arizona and Mexico including Akimel, UNESCO (Moseley, 2010) documents about 15,000 speakers and categorizes the language as "definitely endangered"; children rarely learn the language within the home. For Tohono O'odham in Arizona, "the 2000 U.S. Census reports 18,326 Tohono O'odham age 5 and older. Of those, 10,321 (56.3%) speak English only, and 8,005 (43.7%) speak a language other than English" (Fitzgerald, 2010, p. 232). While Tohono O'odham communities both north and south of the border have experienced language loss, the situation in Sonora is more dire with approximately 22 fluent speakers remaining according to a language survey conducted by Franco Hernández (Franco Hernández, 2009, cited in Franco Hernández & Estrada Fernández, 2016).

2.3 A brief history of Tohono O'odham contact and acculturation

Providing a historical backdrop for language change is particularly salient in circumstances where intense language contact and severe language shift have transpired, as has been the case for Native American communities across the US. Contact between Native American groups, Europeans and Anglos has led to various events that have placed enormous social pressures on Native Americans in the US to assimilate to a European lifestyle, including the adoption of the English language. "The entire fabric of Indian life has been affected by federal policies which have attempted to either assimilate or isolate Indian people. This attitude has discouraged, at least, and negated, at worst, Indian community development and cultural survival" (American Indian Policy Review Commission, 1976, p. 22). This is

certainly true of Tohono O'odham people in Arizona (hereafter O'odham unless a distinction between dialects/groups is needed). Acculturation of the O'odham community has included attempts at religious conversion, confiscation and restructuring of Native inhabited land, education in European values and language, and a gradual shift to a cash-based economy. All these factors have combined to marginalize O'odham culture, identities, and language – social pressures to which the community has responded – and are discussed below.

As a preface to the content in the following sections, it is important to state that what has been written in the literature about O'odham history should be accepted with some reservations for a couple of different reasons. First, for certain periods of O'odham history, a dearth of records is available. For example, Reff (2008) describes the records of the Spanish colonial period of the late seventeenth and early eighteenth centuries and those of the preceding eras as "fragmentary" (p. 358). Secondly, the historical records of this period are often problematic in that they have been "authored largely by non-Indians" (p. 358). This is true of later eras as well, during which O'odham (and all other Native American groups) have had very little role in shaping the histories that have been written on their behalf. Brenneman (2014) raises a similar issue pointing out that O'odham history has largely "been constructed on a Eurocentric template" (p. 208). Much of the material that I present here is based on other historians' interpretations of O'odham history based on that template. In an effort to create a slightly more (and probably insufficiently) balanced interpretation, in the sections that follow I interweave O'odham perspectives regarding their history where possible, particularly regarding how colonialism has impacted O'odham culture and language. For further readings on this specific subject matter, the reader is directed to the following O'odham authors: Cruz (2012) as well as Cruz's doctoral dissertation which, at the time of writing, is a work-in-

progress; Dolores & Mathiot (1991), particularly pp. 307-309; Lopez (1998); McCarty, Romero-Little, Warhol, & Zepeda (2009); Zepeda (1990, 1998).

In brief overview of what is included in the following sections, the first section provides a description of traditional O'odham existence to offer some sense of the changes to O'odham society that resulted from contact with Europeans and Anglos (§2.3.1). This is followed by a discussion of the Spanish colonial period during which the Catholic Church began its proselytizing efforts in the Southwest (§2.3.2), and the short-lived Mexican period that ended with the US Gadsden Purchase Treaty in 1853 (§2.3.3). Following Gadsden, US federal American Indian policies began setting in with reservations, allotment, and federal Indian education (§§2.3.4 and 2.3.5). Later, John Collier's Indian Reorganization Act and the Meriam Report leads to a gradual shift away from federal assimilationism (§2.3.6). The next section describes how the O'odham community's shift to a cash economy also promoted further contact with outsiders (§2.3.7). This is then followed by a chronological picture of language shift from O'odham to English in the twentieth century (§2.4). Finally, the last section addresses Tohono O'odham Tribal language policy, current discussions among O'odham in Arizona regarding language loss, language attitudes, and the O'odham community's efforts to maintain and revitalize the language (§2.5).

2.3.1 Traditional O'odham lifestyle

Fontana (1983b) states that it is not known how long O'odham people "have lived scattered throughout … the western two-thirds of southern Arizona and northern Sonora, Mexico" (p. 125). Others suggest that they have inhabited the area at least for two millennia, perhaps longer (Gladwin, 1957 cited in Smith, 1981, p. 127). O'odham state that they "have always

lived in their present country" (Gaillard, 1894, p. 296) and are the descendants of the ancient Hohokam (Teague, 2009).

Historically, O'odham were semi-nomadic or what Fontana (1983b) describes as "Two Villagers" (p. 131), migrating between two locations based on seasonal rain patterns. In the summer months, O'odham resided on plains between mountain ranges where they irrigated small subsistence farms with the washes that followed the summer rains, and during the winter months, O'odham moved to the mountain foothills where water was more readily available in the form of springs (Fontana 1983b, p. 131; Herrington, Woodward, & Alsberg, 1939; Underhill, 1978). In his own words, Cruz (2012), a member of the O'odham community, describes subsistence in the desert as follows:

> *O'odham* harvested the salubrious natural fruits of the Sonoran desert, grew food by flash flood farming in the alluvial desert plains, gave chase to wild game on their hunting grounds, and held their natural water resources which maintained them (Cruz, 2012, p. 99).

The O'odham economy was originally one based on trade, or more accurately, gift giving of goods such as foods, baskets and other woven materials, as well as labor (Underhill, 1939, 1978). O'odham did not bargain for the exchange of goods. Instead, products were viewed "somewhat like a ceremonial gift which was being reciprocated" (Underhill, 1939, p. 102) and "[g]iving became the regular Papago investment…" (p. 90), an economic ideology that all O'odham groups upheld. Regarding property, O'odham did not "hoard" land and "they did not quarrel about land boundaries" (Underhill, 1939, p. 90). Cruz (2017) describes traditional O'odham society in the following manner:

> Traditional O'odham villages worked together and senior elders sustained a leadership role through consensus decision-making for the benefit of the community. Everyone had a role to serve the welfare of the O'odham community. O'odham village people gathered to listen to one another and did for their villages what needed to be done based on the teachings of the core values. The activities may relate to communal hunting, ceremonials, village security or seasonal activities (Cruz, 2017, p. 8).

In time, many aspects of traditional life would be permanently altered as O'odham contact with various outsiders gradually intensified – namely with Spanish missionaries, Mexicans, and Anglos. The O'odham community also had contact with many other tribes living in the region such as their close relatives, the Akimel O'odham, Hia C'eḍ O'odham, and Sobaipuri, as well as the Yaqui, Piipash (Maricopa), Seri, Opata, Apache, and others, all of whom played important roles in shaping O'odham history; however, it was specifically Europeans and their descendants who had the explicit goal of acculturating O'odham and obliterating their worldviews.

2.3.2 The Spanish colonial period

While the study at hand is primarily concerned with contact between O'odham and English speakers in present day Arizona, the presence of Spanish speakers among the O'odham community and their role in acculturating O'odham cannot be ignored. The O'odham community likely first came into contact with Spanish explorers in search of gold around 1539 (Reff, 2008, p. 356). When they found no gold, they left the area. It wasn't until approximately 150 years later, towards the end of the seventeenth century, that the next group of Spaniards arrived. This time it was Spanish missionaries who came and situated

themselves among the O'odham people for the long term, setting in motion the marginalization of O'odham culture for centuries to come.

What is known of the first missionaries is that they settled throughout the Southwest "to convert the Natives and to establish communities centered around missions which taught the Indians religion and the agricultural way of life" (American Indian Policy Review Commission, 1976, p. 24). The first wave of proselytizing missionaries among O'odham began with the arrival of Jesuit Father Eusebio Kino in 1687 who "baptized many [O'odham] and gave them Spanish names" (Hagan, 1959, p. 139) during his 25-year residence. Kino established a few missions among O'odham, which eventually became inactive before the subsequent arrival of the Franciscan Friars (Underhill, 1978, p. 61). During the Jesuit period, Kino worked towards assimilating adults to Christianity and other aspects of Spanish culture, holding the belief "that only through respected men of the tribes could the the [sic] aborigines be educated to Christianity, and a better way of economic life" (Hagan, 1959, p. 12). Jesuit presence endured for the next 80 years, perhaps intermittently, and in 1767, the King of Spain expelled the Jesuits from the region and Franciscan Friars took up missionizing efforts in the region, who established their first mission at San Xavier del Bac on the present-day San Xavier Reservation. The friars were present up until the 1830s, at which point mission activity came to a temporary halt (Smith, 1981, p. 131).

The impacts of the missionary presence during this period are most likely grossly and inaccurately downplayed in much of the literature usually based on the low number of missionaries who were in the region at a time, as well as the short, sporadic duration of their stays. For example, it is estimated that no more than four missionaries were present at one time in the region (Wagoner, 1952), and missionizing efforts were concentrated in small

regions in the southern and eastern regions of the Papaguería which possibly means that many O'odham had little contact with Spanish missionaries (Fontana, 1983a, p. 139; Joseph, Spicer, & Chesky, 1949, p. 18; Kelly, 1963, p. 7; Underhill, 1985, p. 31). Based on a similar line of reasoning, Bruder (1977) maintains that "[o]nly in the period following 1860 when Anglo contact began, is there reason to believe that Papago life ways began to be seriously affected" (p. 250). Others also add that ongoing incursions of Apache into the region also deterred large scale missionizing efforts and "kept the Papago free from Spanish influence" (Joseph et al., 1949, p. 18). Ironically, Apache were attracted to the region precisely because of the livestock and plants that missionaries brought with them (Kelly, 1963, p. 7).

However, other facts cast great doubt on the claim that Jesuit missionaries minimally impacted the O'odham community. According to Jackson (1981), the Jesuits situated some of their missions among the neighbors of the Tohono O'odham, the Akimel O'odham, who inhabited the river valleys. Records indicate that at some point during the Jesuit presence, it became necessary to fill these missions with "heavy increments of Papago converts" (Dobyns, 1963, pp. 175-176). This was due to the fact that the Akimel O'odham population steeply declined. They had not only been introduced to the Jesuits' religion and aspects of Spanish life but also to "diseases to which Indians had no natural immunities" (Jackson, 1981, p. 243). Those Tohono O'odham who were later recruited to reside on the missions were also likely exposed to disease, although their numbers are uncertain. According to Dobyns (1963), there is strong evidence that the "total depopulation of the middle Santa Cruz Valley wiped out not merely the local northern Piman population" (p. 176), although the number of deaths among the other groups who were in residence at the missions is unknown. Some deaths were also attributable to battles between O'odham and Apache, as alluded to

earlier, but Dobyns (1963) argues that the evidence suggests "disease mortality was much higher than war casualties" (p. 176). If this is the case, then mortality was arguably the first variable associated with language loss for the Tohono O'odham and others in the region. As tribal member Cruz (2012) points out, "Language loss is inevitable when devastating epidemics kill pandemic-like, creating shock and panic leading to suicides, starvation, dehydration, and lethal secondary infections in the absence of care" (p. 99).

During this time, O'odham were also undoubtedly exposed to the Spanish language itself; however, as with many issues concerning this era, it is difficult to gauge with certainty the extent of that exposure. Guzman-Betancourt (1995) states that it was obligatory for Jesuits to learn and use the Native languages of the communities they resided among. In fact, according to Hagan (1959), the Spanish government criticized the Jesuits for their insistence in using Native languages in schools (p. 77). The Franciscan Friars also held the practice of learning Native languages rather than teaching Spanish to Native Americans, and similar to the Jesuits, this at times led to confrontations with the government of Spain: "The government wanted the friars to teach the Spanish language as an aid in political control, but the friars persisted in the use of Indian dialects" (p. 16). It can be deduced from this knowledge that while O'odham likely encountered the Spanish language during this era, few O'odham learned it to an extensive degree.

It is plausible, however, that some of the Spanish borrowings that exist in O'odham today were adopted during this period. The Jesuits and Franciscans not only introduced O'odham to Catholicism, which some O'odham incorporated into their belief system, but they also left their mark in the region with the introduction of wheat, horses, cattle, and a cash

economy (Fontana, 1983a, pp. 138-139), all objects that would not have had names in O'odham prior to Spanish arrival.

Throughout this period, the Jesuit and Franciscan missions combined were "systematically designed to bring about … total conversion, not only from 'paganism' to Christianity but also from being Indians to becoming loyal, tax-paying citizens of the Spanish Crown" (p. 55). The Spanish colonial period was the first of several involving outsiders who wanted to assimilate the O'odham people in all manners of life. The era represents the beginnings of major European efforts to alter O'odham belief systems. It marks the beginning of invalidating and marginalizing the O'odham way of life and encouraging O'odham to reject their own culture.

2.3.3 Borders in transition: The Mexican period and the Gadsden Purchase

In 1821, Mexico became an independent nation, no longer under the rule of Spain, and just 27 years later in 1853, O'odham people found themselves straddling a newly established US-Mexico international border. If there were, in fact, segments of the O'odham population that were able to avoid encounters with European outsiders during the previous era, this was no longer the case.

Following Mexican independence, Mexican farmers, ranchers and miners began to increasingly move into O'odham territory "with utter disregard for Papago rights" (Fontana, 1983, p. 139). The heightened encroachment was such that the O'odham community would eventually enter into a war with Mexicans in the region from 1840 until the eventual surrender of the O'odham in 1843 (p. 139; for a full account, see Hoy, 1994). The building intensity of contact also likely correlated with an increased exposure to the Spanish language.

In 1853, just ten years after the O'odham-Mexican war, the Gadsden Purchase Treaty was ratified and the US acquired present-day southern Arizona and the southwestern corner of New Mexico from Mexico. The region was now open to Americans who, like the Mexicans, were eager to find land for cattle ranching and mining activity (Kelly, 1963, p. 8; Herrington, Woodward, & Alsberg, 1939; American Indian Policy Review Commission, 1976; Smith, 1981). Fontana (1974) notes, "The taking of Papago lands had set in with a vengeance" (pp. 191-192) and O'odham saw the best of their lands being taken by outsiders. There was a "blurring of Papago boundaries by the steady encroachment of non-Indians on Papago lands" (p. 192) on both sides of the border. Encroachment was a trend that was taking place all across the US as European colonizers pushed westward (American Indian Policy Review Commission, 1976, p. 39).

The O'odham community was now also for the first time in their history divided by an international border, approximately two-thirds of the Papaguería in Arizona and one-third in Sonora, Mexico (Dobyns, 1972, pp. 40-41). The division would not impede travel over the newly created boundary for some years (Fontana, 1983a, p. 140; Madsen, 2005), but in time, the international divide was felt, and many O'odham in the US and Mexico began to have divergent experiences under two distinct international governments (see Madsen, 2005).

The ramifications of the division into North and South have, in the long term, been felt more severely by O'odham communities remaining on the Mexican side of the international border. As more Mexicans continued to push further into the region following the Gadsden Treaty, interactions and contact between O'odham and Mexican groups intensified, and many O'odham found themselves "closer to non-Papago neighbors with their traditional ways of life correspondingly being eroded…" (Madsen, 2005, p. 78). By the late

1890s, O'odham people in Sonora faced "an ever-burgeoning number of Mexican cattlemen, ranchers, miners, and farmers" (Fontana, 1983a, p. 141). In 1898, another battle between the O'odham and Mexican populations broke out at El Plomo, which resulted "in the death of some of the Indian attackers, and, ultimately, in a general exodus of Mexican Papagos to the United States" (p. 141; for a full description of the battle see Dolan, 1972).

For those who stayed behind in Sonora, their O'odham culture and identity rapidly became susceptible and vulnerable to change. This was due in part to a continued loss of land; no federally defined land base was established for the O'odham community in Sonora (Madsen, 2005, p. 79), as was eventually the case for O'odham residing in the North with the reservation system. Those remaining in Sonora began to intermarry with Mexicans who now occupied much of their land, becoming increasingly more integrated into Mexican society (Fontana, 1983a, p. 141; Madsen, 2005, p. 79).

Meanwhile, those O'odham who moved to the North following Gadsden because of war or other reasons, would, according to Kelly (1963), introduce "Spanish ways acquired during their longer and more intensive contact with non-Indians" (p. 8). The Spanish language may have been one of the things that O'odham people in Sonora began to acquire and in turn bring with them to the North, although the historical literature includes some conflicting statements regarding language use. By the onset of the Mexican period, if not earlier, Fontana (1989) claims, "Many Papago learned Spanish as a second language, and their own language was enriched by the infusion of Papago-pronounced words for new items in the cultural inventory" (p. 59), although it is not clear how the author arrives at this conclusion. Contrary to these claims, in the 1890s several decades following Mexican independence, anthropologist Gaillard (1894) conducted a survey of O'odham communities

in both Arizona and Sonora and reported that he met few O'odham who were able to speak Spanish (and none that could speak English) (p. 294). However, Gaillard also notes that Spanish was his only means of communication with the community, suggesting that some people were at least partially bilingual at this time. If there were O'odham people who were in fact bringing Spanish to the North, the long-term impacts this has had on the O'odham language in Arizona beyond the lexicon are not clear.

What is evident is that the events associated with the Mexican Period and the Gadsden Treaty permanently altered the cultural and physical landscape of the O'odham community, and the disruptions to O'odham life would continue to progress. In the long term, O'odham on both sides of the international border have grown to greatly resent the restrictions placed on their mobility. Today the Tohono O'odham Nation declares, "[T]he U.S.-Mexico border has become 'an artificial barrier to the freedom of the Tohono O'odham . . . to traverse their lands, impairing their ability to collect foods and materials needed to sustain their culture and to visit family members and traditional sacred sites'" (Tohono O'odham Nation, 2017b, December 17; no citation provided for internal quote).

2.3.4 The American Period: Treaties, reservations, and the General Allotment Act

While the O'odham community was contending with the Spanish and Mexicans, the US government was in the midst of the Treaty Making Period. During this time, the US government was moving towards an assimilationist agenda to deal with what was referred to at the time as the "Indian Problem." The Treaty Period led first to the reservation system, then to the General Allotment Act (also known as the Dawes Act) of 1887. Along with reservations and allotment, education, including English language instruction, eventually

came to be regarded as a central means to acculturating Native communities and to convincing them to reject their respective cultures.

Treaties between the US and Native Americans, according to the American Indian Policy Review Commission (1976), "clarified the relationship between the Indian nations and the United States Government, outlined their mutual responsibilities toward one another, guaranteed the tribes certain rights such as sovereignty and protection from encroaching white settlers, and negotiated trading privileges and land cessions" (p. 29) or at least, those were the intentions. In time, the reality proved to be that many settlers did not observe tribal sovereignty or tribes' rights to the land. At first, Congress attempted to control white settlers and traders who carried out illegal activities on tribal lands such as "using illegal means to obtain goods, and operating without appropriate licenses on Indian land" (American Indian Policy Review Commission, 1976, p. 30) by passing a set of temporary acts in the 1790s. Eventually, however, the government came to the conclusion that this was not enough and explicitly moved towards assimilationism as a solution to the problems between settlers and Native Americans. In 1802, Congress passed the Trade and Intercourse Act, which, according to the American Indian Policy Review Commission (1976), was the first law to explicitly express the need for "civilizing" Native communities (p. 30).

Education eventually became integral to the "civilizing" process. "As more treaties were negotiated, the provisions for education for the purpose of civilization became more frequent. These laws and treaties began to reflect a greater federal responsibility for Indian education" (p. 30). Alternatively, from the vantage point of a member of the O'odham community,

> American politicians wanted to "save" Indians from extinction by detribalizing them through education and making them citizens. Saving Indians through an educational process meant different strategies among the politicians who represented the many sectors of the white political world wanting Indian land (Cruz, 2012, p. 104).

Another major piece of the assimilationist agenda was the establishment of the reservation system, which worked hand-in-hand with Indian education. The reservation system, the federal government argued, was necessary for Native Americans to learn how to lead a sedentary existence and adopt a European lifestyle. The large expanses of land that many Native communities inhabited and their often nomadic or semi-nomadic way of life were diametrically opposed to Anglos' desire for individual, private ownership of property (Adams, 1995, pp. 5-7). The government was also aware that settlers were moving further into tribes' territories and a solution was needed that would appease both groups. In reality, however, what appeasement meant was that Native Americans were coerced into assimilating to an Anglo lifestyle. If tribes were confined to designated reservations, the reasoning went, they would eventually learn to live on and manage smaller plots of land as their new Anglo neighbors did, obviating their need for large expanses of land and opening up more space for settlers. By 1851, the first federal reservations were established in Montana (American Indian Policy Review Commission, 1976, p. 40). In 1869 it was recommended to President Ulysses Grant that the US abolish the treaty system altogether and that all "Indians be collected on reservations…" (Hagan, 1959, p. 21). In 1871, Congress passed the Indian Appropriation Act. All Native Americans were "now deemed to be wards of the government, a colonized people" (Adams, 1995, p. 7), falling completely under the jurisdiction of the US federal government.

The same line of reasoning which was being expressed across the nation was also applied to the O'odham community. Agents from the Office of Indian Affairs (today's Bureau of Indian Affairs) also recommended reservations for O'odham residing on what now was the US side of the US-Mexico border. Reservations, they argued, would allow the government to provide federal protections from the injustices caused by settlers moving in on O'odham territory (Fontana, 1974, pp. 188-189), but again, in truth what this amounted to was the assimilation of O'odham into white society. Without O'odham reservations, Agent Ruggles (1867) reasoned, the O'odham community "[i]n their present scattered condition …, located as they are, throughout a country of about 30,000 square miles in extent, it is impossible to give them such care and instructions as are necessary for that advancement in civilization" (quoted in Fontana, 1974, p. 190). In other words, confining O'odham people would make it easier for the federal government to forcibly acculturate them.

In 1874, the first O'odham reservation was established on San Xavier. This was followed by the founding of Gila Bend in 1882, and the Tohono O'odham Reservation in 1916 (Niswander, Brown, Iba, Leyshon, & Workman, 1970, p. 12). Finally, Florence Village would be added later in 1978 (Tohono O'odham Nation, 2017b, December 17). O'odham reservations would, however, fail "to provide federal protection for anything approaching the original territory of the Desert People in the United States" (Dobyns, 1972, p. 50).

It was also proposed that specialized education should complement all reservations for what the government considered O'odham advancement in a white society. Agent Ruggles (1867) argued that federal schools should be set up for the O'odham community and used to provide instructions "to a degree that will qualify them … to fulfil[l] all the obligations and perform all the duties of citizenship" (quoted in Fontana, 1974, p. 189).

This mirrored President Grant's Board of Commissioners' recommendations for federal schools around the same time (Hagan, 1959, p. 21). Furthermore, the Board suggested, in those schools, "the study of English [should] be introduced to the various tribes" (p. 21). At this time, President Grant also asked churches "to take up the civilizing and educating of these Indian people" (American Indian Policy Review Commission, 1976, p. 40), and the federal government provided funding for religious groups to do so (Hagan, 1959, p. 22). Among the O'odham, Catholics and Presbyterians served as the primary teachers in mission schools of this era (Fontana, 1983a, p. 143), for example, in the Tucson Indian Training School built by the Presbyterians. Many mission schools were heavily criticized, however, for emphasizing bilingual education (Hagan, 1959, p. 23). This criticism reflected the government's movement towards an English-instruction-only policy. Mission schools "were threatened with withdrawal of federal support, unless they complied with government regulations" (pp. 23-24). Some schools did, in time, comply, as is evidenced by the fact that the Presbyterians had produced students who refused to "speak O'odham anymore even though they could understand the language" (Cruz, 2012, p. 103), and had come to strongly associate themselves with the church.

The reservation system and education continued to go hand-in-hand, promoting the acculturation of Indigenous groups and teaching them how to live among their Anglo neighbors. While there may have been some so-called successes with these approaches, for the most part, they eventually were viewed as failures. In time, changes would be made to both the reservation and education systems.

The reservation system did not meet its goal of fully integrating and assimilating Native Americans. Instead the government felt that reservations had "perpetuated the Indians'

attachment to the tribal outlook and tribal institutions" (Adams, 1995, p. 16). Tribes continued to value a communal use of land and gift giving over trade which "reformers viewed as anathemas to the emergence of self-reliant individualism" (p. 16). Secondly, the reservation system was underfunded and had failed to meet people's basic needs such as adequate food supplies leading to turmoil. "Reservation life and insufficient rations placed the Indians in a chaotic situation. The Federal Program had been devastating…" (Hagan, 1959, p. 21).

A new solution was proposed. Reservations, it was argued, should be divided into parcels of land that would then be given to male heads of household and single persons as land allotments. In 1877, Secretary of Interior Carl Schurz envisioned that allotments would naturally lead Native American families to experience "the enjoyment and pride of individual ownership of property" (quoted in American Indian Policy Review Commission, 1976, p. 43), which would be "one of the most effective civilizing agencies" (p. 43).

In 1887, US congress adopted the General Allotment (Dawes) Act. Stuart (1977) describes the General Allotment Act as "the product of an alliance between western interests, intent on acquiring as much of the remaining Indian land as possible, and humanitarian interests, intent on integrating Indians into American society" (p. 454). US citizenship was also to be conferred upon any recipient of an allotment, as well as upon "any other Indians who had abandoned their tribes and adopted the habits of civilized life" (American Indian Policy Review Commission, 1976, p. 43). Among O'odham, allotment specifically affected those residing on the San Xavier Reservation.

Remaining surplus lands after allotment would to be sold off by the US government (Stuart, 1977, p. 452). Allotments were originally to be held in trust by the US for 25 years,

during which time the expectation was that men would learn how to farm and/or ranch, and women would learn how to be domestic workers and Americanized housewives (Marak & Tuennerman, 2013, p. 106). The government would use proceeds from sales to fund schools to promote such "progress" of Native Americans (Adams, 1995, p. 17). Education came to play a progressively more central part in these transformations, and the government would eventually make a major overhaul of how Indian education was conducted.

Kelly (1963) states that from the viewpoint of O'odham people, reservations, allotment, as well as American education had the devastating long-term effect of removing any sense of independence or freedom:

> The typical Papago feels most intensively both the loss of individual action in our highly institutionalized society, and the loss of almost complete self-dependence. The welfare agency, for example, will feed his children if they are hungry, but he must abide by their rules; the Indian-Service school educates his children, but he has no real control over what they are taught, or how they are taught (Kelly, 1963, p. 6).

Booth (2005), a member of the O'odham community, expresses a sentiment along the same lines regarding schools and reservations:

> [T]he OIA [Office of Indian Affairs] ... recruited, or kidnapped, O'odham children for government boarding schools; introduced meddle-some bureaucracy; and imposed on the tribe fences, unwanted projects, and increased police surveillance. In general, creation of the reservation meant loss of autonomy for the very independent O'odham villages (Booth, 2005, p. 376).

The federal government and its supporters had put policies into place that prevented O'odham people from living the way they had previously lived – hunting animals, moving between plains where they grew subsistence farms and the mountain foothills where springs provided water during the winter season. They were now dependent on the federal government and under pressure to give up all aspects of their culture.

2.3.5 US federal education

The assimilation of Native Americans following reservations and allotment continued to be problematic from the view of the federal government; not enough progress was being made. The US government's assimilationist agenda would reach new heights in the ensuing years; the developing philosophy was that Native American "savagery," as they viewed it, including Indigenous languages, needed to be wiped out and replaced if tribes were to become functional citizens of Anglo society. A new kind of education, the government concluded, was needed which included coercive English-only policies. It is the educational policies of this era and those who supported and implemented them that are often blamed for greatly accelerating Native Americans' loss of a secure sense of identity, loss of culture, and loss of language (e.g. McCarty & Watahomigie, 1998, 2009).

Colonel Richard Henry Pratt argued that reservations and allotment were only successful at colonizing Native Americans but not at assimilating them (Adams, 1996, p. 53). Additionally, education centers situated among or near community and family only perpetuated Native cultures, languages, and ways of life. Instead, children, who were young and impressionable, should be placed in schools away from home and completely cut off and separated from their families for a number of years in order to expedite assimilation into American society (Adams, 1995; Hagan, 1959; American Indian Policy Review

Commission, 1976). Indians, Pratt argued, were born as blank slates, not as "savages." Savagery, which included Indigenous languages, was something that was learned and could be prevented. In Pratt's own words, "Left in the surroundings of savagery, he grows to possess a savage language, superstition, and life.… Transfer the savage-born infant to the surroundings of civilization, and he will grow to possess a civilized language and habit" (Pratt, 1892, p. 56; quoted in Adams, 1995, p. 52). Children should be situated among Anglos where they would learn American values, including English. In time, policy makers agreed with Pratt's belief that "[o]nly by attending boarding schools … could savage institutions, outlooks, and sympathies be rendered extinct" (Adams, 1995, p. 59). With these sentiments in mind, the off-reservation boarding school era commenced.

In 1879, Pratt, who had made it his mission to assimilate all Native American children, established and became the headmaster of the Carlisle Indian Industrial School in Philadelphia (Hagan, 1959, p. 186), the first off-reservation boarding school that would serve as an example for many others across the country. Pratt, who had a long history of military service, ran boarding schools as if they were regimented military units (p. 187). Off-reservation schools were to emphasize "character, industrial training, and individualism" (American Indian Policy Review Commission, 1976, p. 44).

Eventually, greater weight was placed on English language instruction. Teaching English to children was considered a vital method for dismantling children's connections to their respective heritages. In 1888, John D. C. Atkins (1888), the Commissioner of Indians, made the following argument in support of using English as a tool for assimilation:

> Schools should be established, which children should be required to attend; their barbarous dialects should be blotted out and the English language substituted. The

> object of greatest solicitude should be to break down the prejudices of tribe among the Indians; to blot out the boundary lines which divide them into distinct nations, and fuse them into one homogenous mass. Uniformity in language will do this — nothing else will (Atkins, 1888, p. 9).

In 1887, a school superintendent also argued that a Native child's "inability to speak another language other than his own renders his companionship with civilized man impossible" (United States Office of Indian Affairs, 1887 quoted in Adams, 1995, p, 21).

By 1890, a "no Indian languages" policy was in place. Schools were told, "Pupils must be compelled to converse with each other in English, and should be properly rebuked or punished for persistent violation of this rule" (United States Office of Indian Affairs, 1890 quoted in Adams, 1995, p. 140). While many strategies were devised for enforcing this rule, most teachers administered corporal punishment, and in fact in the 1890s corporal punishment was condoned by the US government, along with "confinement, deprivation of privileges, or restriction of diet" (United States Office of Indian Affairs, 1891 quoted in Adams, 1995, p. 121).

Around the same time that these language policies were being crafted and implemented, and O'odham children in Arizona were beginning to be forced into attending off-reservation boarding schools, evidence suggests that the O'odham language remained stable despite the mounting pressures on O'odham to assimilate. In 1894, Gaillard (1894) reported, "[N]o Papagos could speak English" (p. 294). However, in the decades following the boarding school era, signs of a shifting linguistic landscape would slowly begin to emerge.

The Phoenix Indian Boarding School was the first federal off-reservation boarding school to open in the vicinity of the O'odham reservation. It opened in 1891 and was the largest of its kind in the region (Fontana, 1983a, p. 143). It was modeled after Pratt's Carlisle Indian School where "[e]very activity was under military discipline" (Hagan, 1959, p. 193). The Phoenix school was to provide students with academic training, but in reality it enforced manual labor skills as did other boarding schools across the country (American Indian Policy Review Commission, 1976). The Phoenix school was expected to "Americanize" Native Americans living in the Phoenix area, because "to outsiders, their traditional culture seemed valueless, their lifestyle immoral and degraded" (Trennert, Jr., 1988, p. 13). O'odham children were also placed in other federally operated off-reservation schools across the nation, including a few at Pratt's Carlisle Indian School (Hagan, 1959, p. 187). Others attended the Industrial Training School in Grand Junction, Colorado, the Chilocco Indian School in Chilocco, Oklahoma, and the Santa Fe Indian School in Santa Fe, New Mexico (Fontana, 1983a).

O'odham tribal members describe the boarding school era and other forms of federally funded education as follows: Cruz (2012) states, "O'odham people were pressured to reject their language, religion and culture. In the boarding schools use of Indian languages were denied. Ceremonial practices were outlawed by the Indian Affairs Department" (p. 103). Furthermore,

> When U.S. Indian schooling began, O'odham were not able to teach their children their Indigenous ways in the schools, instead the children were forced to learn a culture devoid of O'odham subject matter. Compulsory U.S. government education laws are viewed with caution by O'odham as an attempt to make their children white

> by usurping O'odham language and resulting in the acquisition of negative European cultural practices (based on audio recording cited in Cruz, 2012, p. 101).

Students across the nation were placed into schools by means of "cajolery, threats, bribery, fraud, persuasion, and force" (Adams, 1995, p. 305), and in 1892, the Commissioner of Indian Affairs was given the authority "to enforce regulations to secure the attendance of Indian children of suitable age" (p. 190). The following year, "to aid in enforcement of compulsory attendance rations were withheld from families who refused to send their children to school" (p. 190). O'odham tribal member Cruz (2012) describes it another away, stating that despite the harsh conditions,

> Indian families were starving on Indian reservations because government food rations were not delivered due to government corruption in the administration of the reservation system. Indian boarding school children therefore were happy to stay and survive at the boarding schools (Cruz, 2012, p. 104).

Pratt also created the "outing program" which was to complement the off-reservation boarding school experience. During the regular school year, children received vocational training in boarding schools, and during the summer months, they were placed with white farming families where they could apply the skills they had learned. They would learn to live as white families did and, furthermore, they would be immersed in English (Adams, 1995, p. 55, 156-157). The outing program also obviated the need for children to go home during the summer, sustaining their separation from their families and communities sometimes for many years at a time, and thus decreasing the chances, the reasoning went, that children would revert to their respective tribal lifestyles.

The outing program was introduced to O'odham and other tribes of the Southwest. Through the program, children both learned skills for and participated in domestic servitude (Fontana, 1989, p. 76; Hagan, 1959, p. 193). According to Fontana (1983a), girls also learned English (p. 143).

As was the case in boarding schools, evidence suggests that O'odham children endured hardships in the outing program as well. Juan Dolores (Dolores & Mathiot, 1991) provides a firsthand account of his and other O'odham children's experiences in the outing system during the late 1800s. Dolores and other O'odham children are brought to the home of a man named Mr. Thompson's to do manual labor in his fields. Dolores describes the harsh treatment he and his O'odham companions received including physical punishment for their inability to understand English:

> *'Amjeḍ a 'i mat hihi 'u:hum g n-we:m ki:kam, ḍ mu'i ha'icu pi 'ap 'e-juñhim 'am n-ta:gio 'ñḍ aṣ pi has o n-do:da. Wes taṣ ḍ ha'i hihim 'am Mr. Thompson's ki: wui. Gḍhu daiw ḍ 'am hema wo ha-'a:gidad mo has kaij g Mr. Thompson. Ñ s-cegito 'i:da Miligan. Ḍ o ñeokc o ṣa s-bagam kaijid. Ñ hi s-'e:bid waṣaba pi hebai wo n-do'ibia. Ḍ hekid hema ha'icu pi wo 'amic, ḍ wa hekaj 'an o bei na:kd 'an k 'in has 'ep o wanckwuhid. Ḍ o ñeokad ḍ hega'i jioṣ 'am o wa s-kaidagk. 'Amjeḍ 'aṣp has 'ep 'i t-'a:g cḍ pi 'amicud, waṣ hi wa s-ma:c mat bagat. Ṣp hems kaij, mat hema taṣ o t-kokda wes, mac pi ha'icu wehejeḍ s-'ape* (Dolores & Mathiot, 1991, pp, 261-263; converted into Alvarez-Hale orthography).
>
> The moment my parents went home many bad things happened to me. I was - completely helpless. Everyday some kids would go to Mr. Thompson's house…. I remember this white man. He always seemed angry when he spoke. I was afraid of him, but I could not escape. If someone did not understand something he would grab him by the ear and would yank him back and forth. When he spoke we could only

> recognize the word "God". He would say things to us that we did not understand but we knew that he was angry. Maybe he was saying that one day he would kill us all, that we were no good for anything (Dolores & Mathiot, 1991, pp. 307-308).

According to Cruz (2012), O'odham children were placed in the outing program not solely to teach them skills for servitude, but also because locals sought after their cheap labor:

> A few weeks after P.I.S. [Phoenix Indian School] opened, local farms began asking the school superintendent for workers to help harvest their fields. When the superintendent realized that the sedentary and semi-sedentary living practices of O'odham could accommodate the capitalist industrial economic system growing in the Southwest, he began to hire out the Indian students. Young Indian girls were placed in white family homes to perform domestic work in exchange for their board, care and instruction. This training was called the "outing system" to promote Anglo standards of social behavior and marriage guidance (Cruz, 2012, p. 105).

The outing system essentially imprisoned children in order to turn them into servants as well as to force them to learn how to live like Anglos.

By 1901, off-reservation boarding schools along with the outing system were declared a failure just as reservations and allotment had been. The goal of training Native Americans to manage their allotments as functional, independent landholders had not been met (Adams, 1995, p. 307). The notion that off-reservation schools could transform a generation of children within a few years and that those children would then return home and alter their respective societies was, the government concluded, unrealistic. Furthermore, a

new racist philosophy had emerged during this period, according to which certain races were inherently primitive; therefore you could not expect a race to change within one generation (p. 310). The cruelty of off-reservation schools was also recognized – pulling children away from their families in some cases for years at a time and placing them into a completely foreign environment (p. 311).

In the Annual Report of the Commissioner of Indian Affairs of 1905 (United States Office of Indian Affairs, 1905), then Commissioner Francis Leupp declared this approach to education as "a great mistake" (p. 2) and later argued that rather than bring the Indian to civilization, the approach that Pratt favored, civilization should be brought to the Indian (Adams, 1995, p. 318).

It was recommended that off-reservation schools gradually be closed and replaced with Day Schools on or near reservations, and that schools should incorporate aspects of tribes' cultures, while at the same time acculturating them. For example, Commissioner Leupp promoted the idea of assigning children with the task of retelling traditional stories or describing home life in written English (described in Adams, 1995, p. 316). While there was a small shift in sentiment about how Native American children should be treated and taught in schools, in reality, many boarding schools remained open and many schools did not change; children continued to endure long periods of separation from family and extreme discipline within school centers. In the 1920s, four-fifths of Native American children who received education were still attending boarding schools (p. 331). Due to funding shortages, nutrition was severely lacking, living conditions were unsanitary, dormitories were overcrowded, and many children were ill (pp. 331-332). At the Phoenix Indian Boarding School in particular where O'odham and other Native American children attended, children

contracted measles, smallpox, spinal meningitis, as well as trachoma and tuberculosis (Trennert, Jr. 1988, p. 102). While the severity of off-reservation boarding schools was recognized, and day schools were promoted, the goal of education in day schools continued to be assimilationist in their agenda. Students attended class during the day learning English, as well as simple math and industrial skills. Afterwards, it was hoped that when children returned home, they would introduce what they had learned to their families. In this manner, Native Americans would eventually learn the ways of white society.

During the early 1900s, several O'odham day schools were established in relatively close proximity to where O'odham families resided – Tucson, Sells, San Miguel, Gila Bend, where, earlier in 1882, a reservation for the O'odham residing there was federally recognized, as well as other locales in the region (see Fontana, 1983a, p. 143). Around the same time, there was also growing support for incorporating Native American children into public schools (Adams, 1995, p. 319); however, according to Hagan (1959), this development did not impact Southwest tribes until decades later (p. 192). The church also continued to be involved in the education of O'odham children, and several Catholic day schools were also established in the region (Fontana, 1983).

2.3.6 Against assimilationism: John Collier, the Meriam Report, and the Indian Reorganization Act

By the 1920s, criticism of the US's harsh treatment of Native Americans began to grow. Although tribes had already experienced irreparable damage to the entire fabric of their societies, an important, albeit slow, shift away from assimilationism was developing in the thinking of some Americans. This shift took shape in reforms to Indian policy led in large part by John Collier who "waged a relentless war against the government's Indian policy" (Adams, 1995, p. 330), heavily criticizing "the failure of the Indian Office to protect native

landholdings … and its determination to destroy Indian culture in the name of assimilation" (p. 330). Mounting pressures by Collier and others eventually led the Secretary of Interior, Hubert Work, to put together a team to survey the situation on reservations across the nation (p. 331). The investigation culminated in a document entitled, *The Problem of Indian Administration* (Institute for Government Research, 1928), more commonly known as the *Meriam Report*. The report revealed the harsh and unhealthy living conditions of boarding schools, and the severely impoverished state of reservations (Adams, 1995, p. 331).

Collier, who was eventually made Commissioner of Indian Affairs under President Franklin D. Roosevelt, and his associates drafted the *Wheeler-Howard Act* or what is better known as the *Indian Reorganization Act*. The Act brought an end to the further redistribution of tribal land as allotments "and enabled tribes to organize themselves as governments and as corporations for purposes of economic development" (Stuart, 1977, p. 453). Additionally, "[o]ther provisions … enabled the Secretary of the Interior to restore unsold surplus lands to the tribes, extend the trust period of allotments indefinitely, and provide for the purchase of lands to be added to the reservations" (p. 453). Collier viewed the Indian Reorganization Act as a means to restoring and preserving Native American communal life and culture as well as improving their economies. Furthermore, the Act was meant to promote bilingualism rather than the English-only policy that had been in place (Erickson, 1994, p. 148). According to Stuart (1977), however, for the most part "the commitment to assimilation remained strong" (p. 456).

The Phoenix Indian School, which O'odham and other Native American children attended, exemplified such a commitment. While the school received severe criticisms during the era of Collier and the Meriam Report – criticisms that led to changing the school's

"internal structure, subjecting it to embarrassing publicity, bringing to an end traditional ideas and programs, and nearly culminating in the school's closure" (Trennert, Jr., 1988, p. 182) – evidence suggests that some policies remained intact, particularly with regard to language. In 1998, Lopez (1998), a member of the O'odham community in his 50s at the time of writing, reminisces about his experiences as a second grader at the Phoenix Indian school – perhaps sometime between the mid 1950s and early 1960s: "When I went to Phoenix Indian School …, I wasn't allowed, I couldn't talk O'odham, and that was the only language I could speak" (p. 44).

Lopez (1998) also discusses similar experiences at another school: "Later on at St. John's High School, outside of Phoenix, the Indian boarding school there, again we were not allowed to talk our language for our privileges got taken away if we got caught talking our language" (p. 44).

While the movement away from assimilationism reflected by the Indian Reorganization Act marked the beginning of an important ideological shift in US colonial history, in reality, the Act proved to have its shortcomings (Stuart, 1977, p. 53). For example, even though the Act promoted Native communal life and culture, many schools continued to implement an English-only policy well into the twentieth century.

In conclusion, Zepeda (1998), an O'odham tribal member and linguist, describes the boarding school era as one during which "the O'odham people suffered severe punishments for speaking the native language at the boarding schools. The variations in degrees of language loss and marked native language attitudes still prevail for that generation of early boarding school students" (p. 49; referencing Crawford, 1995). The boarding schools have

had negative repercussions for the language and many other aspects of O'odham life and culture that continue to be felt today.

2.3.7 Cash economy

It is an oversimplification to state that reservations, allotment, and education were the only factors that led to the language shift O'odham communities have experienced. Fontana (1983a) points out that "[a]s early as the 1880s and 1890s Papagos began in growing numbers to sell their labor to mines and [cattle] ranches" (p. 145) as well as their goods, such as wood, produce, and crafts to people in southern Arizona who were not Native.

During the first quarter of the twentieth century, approximately 600 O'odham were employed by the mining industry in the region, which led to an increased exposure to "whites and their culture" (Smith, 1991, p. 132). Many O'odham in Arizona also left the reservation, temporarily or permanently, to work on cotton-plantations where they were exposed to "whites, Mexicans, Spanish-American[s], Negros and other Indian groups" (Smith, 1991, p. 132 referencing Dobyns, 1950).

Perhaps the most marked change to O'odham life in terms of integrating into a cash economy was their eventual participation in the cattle-ranching industry which was booming in the region. Fontana (1983a) argues that "[a]s important and lasting as the impact of schools on the Papagos may have been, just as important has been the impact of cattle" (p. 144). Cattle-ranching became central to the O'odham economy and a major industry that they controlled, some O'odham raising as many at 1,000 cattle specifically for cash profit. They ran their cattle businesses emulating "the model of Anglo cattlemen" (p. 144), and by the 1930s, O'odham "had become irreversibly tied to non-Indian cash economy for their livelihood" (p. 145).

During World War II, "Papago were drafted, and both men and women began to leave the Reservation in large numbers to work in cotton fields, ranches, mines, railroads, and airplane factories" (Joseph, Spicer & Chesky, 1949, p. 27). "[T]he conscription of Papago during the Second World War and the availability of wage labor in defense plants further promoted contact of Papago with non-reservation members after 1940" (Smith, 1981, p. 135). Their employment in industries related to WWII exposed many youths to the European version of economic prosperity, which was "obtainable through participation in the job opportunities and the economy of the non-Indian world" (Kelly, 1974, p. 14). All of these changes combined promoted cultural shift that directly resulted from contact.

As interactions increased with non-O'odham people in spaces such as those just discussed, so too did the need to communicate in a language other than O'odham.

2.4 Language shift in the twentieth century

In this section I construct a chronological sketch of O'odham language shift, focusing primarily on the twentieth century. The first reason for constructing this sketch is to offer a more nuanced picture of the various stages of language shift among the O'odham community. The second reason for doing so is to allow me to situate possible structural changes in the context of corresponding periods in time throughout the study. The periods can roughly be broken down into the following: **early-shift period** (early 1900s), **middle-shift period** (mid 1900s), and the **late-shift period** (late 1900s to the early 2000s). The period preceding the twentieth century I refer to as the **monolingual period** (pre-1900s).

The sketch that I provide in the subsequent sections is based on the observations of O'odham community members, schoolteachers, missionaries and nuns, historians, and other scholars either in first-hand accounts or cited in the various secondary historical sources that

I have consulted for this portion of the study. While many of the sources are not specifically concerned with the impacts of colonialism on the O'odham language, many include explicit comments on the status of the O'odham language at different points in time as well as information regarding bi-/multilingualism in the community. These observations allow me to construct an estimation of how language shift has and continues to evolve among the O'odham community.

The details in the different sources suggest that, up through the first half of the twentieth century, a majority of O'odham people continued to predominantly speak O'odham in Arizona. This is the case despite several centuries of European outsiders pressuring O'odham to assimilate. It isn't until well into the 1900s that shift towards English in Arizona noticeably accelerates following decades of federal policies and programs including the reservation system and Indian education. This suggests that while all of the preceding colonial eras clearly marginalized and damaged many aspects of O'odham culture, US federal Indian policies correlate most strongly with language loss. Of those federal policies, the enforcement of English-only in federal schools was likely the greatest force leading to the language loss observed today in combination with centuries of previous attempts at convincing the O'odham community that their culture was inferior to a European lifestyle.

2.4.1 Early-shift period (early 1900s)

Beginning first in the early 1900s, clues suggest that the O'odham language continued to be the primary language of a majority of O'odham children, although some children could also speak English. McCarthy (1985) reminisces about his experiences at what he refers to as the Phoenix School, (probably the Phoenix Indian Boarding School) sometime around 1906. He states, "School was tough for us. None of us could speak English, and the boys did not try to

learn the white man's talk – instead, they always talked in their own languages outside the classroom" (p. 29). Later in 1911, at which point McCarthy had transferred to a school in Santa Fe, New Mexico, he notes that a friend of his named Pat could "speak pretty good English" (p. 29), suggesting that at least some children were learning English at this time.

In 1911, a priest named Father Bonaventure O'Blasser began working within the O'odham education system based on his belief that schools should be central in Christianizing O'odham (Hagan, 1959, p. 149). Father Bonaventure arrived with some prior knowledge of O'odham which made him "well prepared for his work…" (p. 149), suggesting that he instructed his students in O'odham out of necessity. It is also noted that Father Bonaventure required interpreters "who aided him in understanding the more complicated discussions with the various village chiefs during his early work with the Papago" (pp. 149-50), indicating that at least some adults did not speak English at this time.

By the 1920's, there are some indications that O'odham children required preparatory classes in "Americanization" which were "needed to prepare the youngsters for the classwork and catechism, which were in the English language by the Sisters" (Hagan, 1959, p. 154). Although not entirely clear, this may suggest that the Sisters assessed that some English language instruction was required before the children could enter into the regular classroom.

2.4.2 Middle-shift period (mid 1900s)

Jumping ahead to the late 1940s, a majority of O'odham still do not speak English; however, a clear increase in the number of O'odham who know some English becomes visible. Segundo (1948) reports, "Less than forty percent of the Papago people speak English, less than twenty percent can read or write" (quoted in Hagan, 1959, p. 215). This also suggests

that some literacy in English is also developing at this time. Additional statistics for the late 1940s reveal a similar trend:

> The great majority of Papago children know at most only a few words of English when they enter school. Only a third of the adults speak it; perhaps a tenth can read and write it, and a negligible number do so habitually. In the two western districts when this study was made, no Papago over forty spoke English; only ten of the parent generation did so, and, of these, four subsequently moved off the Reservation. As a general rule, even parents who know English seldom use it in the home, hence the child's first years of schooling must be mainly devoted to learning the languages in which he is taught (Joseph et al., 1949, p. 98).

In other words, by approximately 1950, O'odham continues to be the primary language for O'odham adults and children and a majority of O'odham continue to primarily speak O'odham, but the number that speak some English has increased.

As the push for O'odham children to learn English continues, more O'odham students acquire it with each passing decade. In the late 1950s, a teacher reports "that each year a greater percentage of her pupils enter the first[-]grade Indian room knowing English and are better prepared to follow the regular curriculum" (Hagan, 1959, p. 219).

While more children are speaking English in schools at this time, O'odham remains the language spoken at home. In 1960, Fontana (1960) conducts a census study of language use on the San Xavier Reservation that reflects as much. According to Reeves (2006), the census data leads Fontana to conclude, "Tohono O'odham was still spoken in each household even if English was also spoken" (p. 47). Similarly, in 1963, Kelly (1963) observes that "the Papago dialects [are] still spoken in the vast majority of Papago homes" (p. 1).

2.4.3 Late-shift period (late 1990s-early 2000s)

A couple of decades later in the 1980s, a sample of 100 adults who are at least 55 years of age reveals, "With the exception of one or two speakers, all were more comfortable using the O'odham language than English. Of this group, approximately half were considered monolingual speakers of O'odham" (Zepeda, 1998, pp. 49-50). Children who speak O'odham on a regular basis on the other hand, have become less common. Fontana (1989) makes the following observation in 1989: "[O]ne listens in vain to hear Papago being spoken. The children playing in nearby yards are speaking only English" (p. 141). While children may have continued to speak O'odham at home, English appears to have now become their primary language even when in the company of other O'odham children.

By 1998, Zepeda (1998) states that O'odham in their twenties and teens "are, in general, the weakest native language speakers; among this group the number of bilinguals in both O'odham and English diminishes" (p. 50), meaning many of this age group are now monolingual English speakers. Zepeda notes that while there are some exceptions, overall,

> [L]anguage loss is most rampant among young teenagers and children. Many O'odham children and teenagers know only a few vocabulary words and cannot carry on a simple conversation. The conclusion, then, is this: very few O'odham children are learning O'odham as a first language (Zepeda, 1998, p. 50).

Around the same time, Lopez (1998), an O'odham teacher, laments the state of the O'odham language among the younger generations stating, "Even today when I meet with the kids in my village I have to speak English because they can't understand the O'odham language" (p. 44). He also notes that in some cases, children's proficiency in O'odham depends on where they live on the reservation: "[Y]ou go to the western part [of the

reservation], a lot of kids ... , in my village, tiny little kids, and they speak O'odham better than I do. The rest of the kids, it's English" (p. 44).

Finally, by the early 2000s, as discussed in §2.2, the 2000 US Census reports that 56.3% of those who participated in the census are monolingual in English (Fitzgerald, 2010, p. 232). Of those who are 18-19 years old, only ten percent can speak O'odham (Madsen, 2004, p. 232 cited in Fitzgerald, 2010, p. 232).

2.4.4 Timeline summary

These language shift trends are summarized in the timeline below in Figure 5.

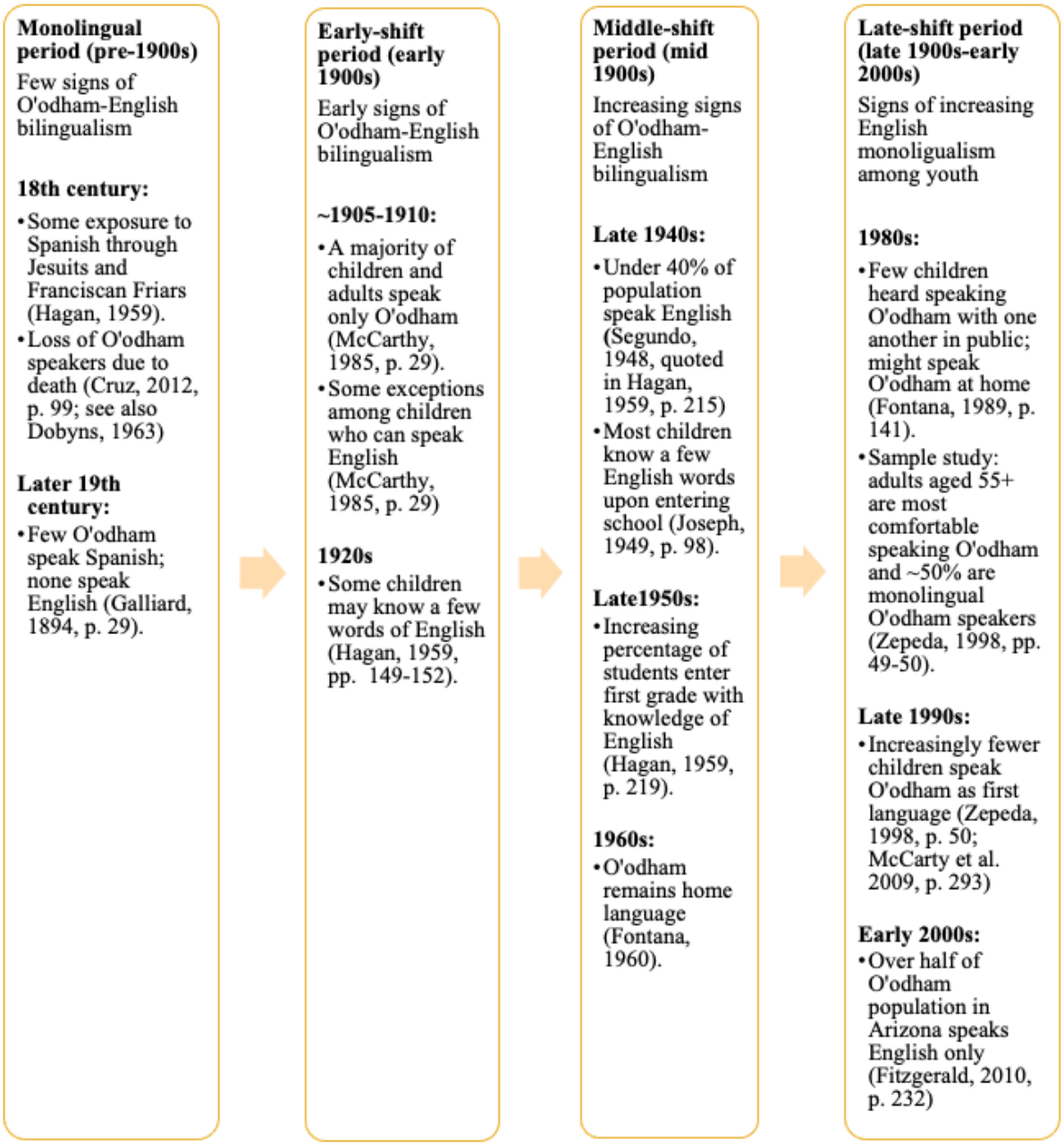


Figure 5. Timeline summary of language shift from O'odham to English.

While it is clear that colonialism has taken its toll on the O'odham community and their language when looking at the overall picture over the last approximately 100 years, the situation is not without hope. There remains a significant (yet dwindling) number of speakers who could potentially transmit O'odham to a new generation.

2.5 O'odham Tribal language policy today

The O'odham community is aware that a majority of children are no longer learning their language and are aware of its endangered status. The Tohono O'odham Tribal Council has declared it its mission to work towards the revitalization of O'odham. On the Nation's official government run website it is stated, "We must revive and protect our language including our traditional ceremonies. Our traditional ceremonies and songs recognizes, acknowledges and gives appreciation to our environment – the plants, animals, birds, mountains, water, air/wind, sun and rain with the inten[t] of cleansing the earth. We cannot lose our *Himdag*" (Tohono O'odham Nation, 2017a, December 17; italics added).

In 1986, the Tohono O'odham Tribal Council declared O'odham as the official language of the Nation (Zepeda, 1990, p. 251), noting that the language is central to O'odham identity. According to Zepeda (1990), the Tribal Council states "that language is the gift from the creator and, '…no other tribe can claim it. It is what makes us Tohono O'odham. … our Tohono O'odham *himdag* 'way of life' demands respect and maintenance by all those who claim it'" (Tohono O'odham Tribal Council, 1986, p. 3 quoted in Zepeda, 1990, p. 251; Zepeda's omissions). The declaration also states,

> The Tohono O'odham language is to be used as the official means of oral communication at any and all tribal councils, and all district, village, committee and board meetings as well as in any and all tribal and community functions and activities

> throughout the Tohono O'odham Nation. Other languages may be used as deemed necessary (Tohono O'odham Tribal Council, 1986, p. 3 quoted in Zepeda, 1990, p. 251).

The Nation's policy also acknowledges that the language is specific to their culture, and that community members should first and foremost transmit O'odham and *Himdag* to future generations:

> Our language has been used throughout our people's existence to convey thoughts, ideas, and feelings. The Tohono O'odham Nation declares that all Tohono O'odham parents, guardians, elders, and adults are the first and initial teacher of the Tohono O'odham *Himdag*' (Tohono O'odham Tribal Council, 1986, p. 3, quoted in Zepeda, 1990, p. 252).

Furthermore, the policy declares that schools will implement bilingual and bicultural curriculum:

> It shall be the policy of the Tohono O'odham that all educational systems shall establish and operate educational programs using bilingual/bicultural practices (Tohono O'odham Tribal Council, 1986, p. 5 quoted in Zepeda, 1990, p. 253).

Despite the Nation's official language policy, in many cases the reality of the situation is different according to specific community members. For example, Cruz (2012) observes that although O'odham is recognized as the official language of the reservation, "Today official meetings on the O'odham reservation are conducted in English" (p. 103).

Furthermore, while O'odham teachers value the language and believe it is important to teach O'odham in the school system, educators such as Lopez (1998) lament some of the

challenges that they face in the school system: "A lot of the kids don't speak the language because they just do so much in the school system, especially in my class because I have to move on to other things" (p. 44).

When students from several different tribes, including the Tohono O'odham Nation, were surveyed about their exposure to their respective tribes' languages among family, they "indicated that their parents and grandparents speak the Indigenous language—but typically not to them" (McCarty et al., 2009, p. 294). Similarly, Garcia (n.d) expresses his opinion noting, "The youth are not taught enough about their traditional ways such that they seem foreign and so they think their cultural ways are funny and they laugh at them" (quoted in Cruz, 2012, p. 101). At the same time, some students are clearly aware and concerned about the status of their language: "'Right now, we're losing it,' a Tohono O'odham youth said, 'so it's very important for me to learn about it and to speak it' [interview, April 19, 2004]" (McCarty et al., 2009, p. 301).

The Nation also offers continuing education at the Tohono O'odham Community College (TOCC) where one of the primary goals is "to facilitate the preservation of Tohono O'odham culture and tradition" (Tohono O'odham Nation, 2018, January 17). TOCC also holds a *Himdag Policy*, requiring all students and staff to study O'odham language, history, and culture (Tohono O'odham Kekel Ha-Maṣcamakuḍ, 2018).

Overall, the community believes that it is important for them to know O'odham and teach it to members of the community who want to learn it. This is reflected in the results included in the unpublished survey report mentioned earlier (Tohono O'odham Department of Education & O'odham *Ne'ok* Revitalization Project, 2013). In one section of the survey addressing language attitudes, those surveyed were asked to indicate on a scale of 1 to 5 how

strongly they agree with the statements below (with 1 meaning *disagree strongly* and 5 meaning *agree strongly*).

1. It is important for members of our community to know their language.
2. Our community should make efforts to teach the language to people who want to learn it.
3. An important part of being O'odham is to know the language (Tohono O'odham Department of Education & O'odham *Ne'ok* Revitalization Project, 2013, p. 34).

Of the approximately 327 survey participants who answered these questions, a majority overwhelmingly stated that they *strongly agree* with each of the statements, the average answer scores being 4.8, 4.8 and 4.7, respectively. These findings indicate that the O'odham language continues to be highly valued as a part of O'odham culture, and the community wants their language to continue.

While the O'odham community on the one hand understands the difficulties and challenges facing them in reversing the current trend of language loss, they also continue to value their language as central to their culture, identity, and ways of knowing.

2.6 Summary

In this chapter I have provided a background overview of O'odham regarding where it is situated in the Uto-Aztecan language family and an assessment of the language's vitality. I have also described the historical social pressures and events that have coerced O'odham people to assimilate to their colonialist surroundings which includes the adoption of English north of the border and Spanish south of the US-Mexico border. I have also sketched an estimated timeline of the shift from O'odham to English in particular based on comments on language use in various sources. This not only offers a more detailed understanding of the various social pressures that led to language shift, but the timeline also allows me to contextualize potential changes in language structure within specific periods in time. Finally,

this chapter summarizes more recent language attitudes among the O'odham community about the value and importance of their language. Ultimately, the community's desire is to revitalize and maintain their language.

Chapter Three: Typological description of O'odham grammar

3.0 Introduction

This chapter presents an overview of O'odham structure. The components that I have chosen to describe are often included in typological profiles of a language; however, the description I present here is more detailed with the goal of providing the reader with a basic understanding of O'odham grammar and the overall typological characteristics of the language. This description may also be useful for linguists who are concerned with universal patterns in the world's languages as well as differences. Section 3.1 describes the basic phonological characteristics of O'odham. Section 3.2 briefly addresses the morphological typology of O'odham. Section 3.3 includes a discussion of how clausal participants are coded in the language. Next, §3.4 provides a partial overview of verbal morphology with respect to O'odham's aspectual system, and §3.5 describes various methods for expressing nominal and verbal number in O'odham. This is followed by a description in §3.6 of grammatical relation and alignment patterns, and §3.7 addresses word order of subjects, objects, and verbs, adpositions, and genitive constructions. Finally, §3.8 briefly discusses the locus of marking (i.e. head/dependent marking) as it relates to typological concerns. I provide more in-depth descriptions of the areas of grammar that are the focus of this study – progressive constructions and demonstratives – in Chapters 5 and 6.

3.1 Phonological profile

The following sub-sections present a brief overview of the phonemic inventory, prosody, and syllable structure of O'odham.

3.1.1 Phonemic inventory

The basic phonemic inventory of the O'odham language includes five vowels and 20 consonants, listed in Table 1 and Table 2. These are given in the Alvarez-Hale orthography which the Tohono O'odham have adopted as their official writing system. For additional orthographic correspondences, the reader is directed to Appendix A.

	Front	**Central**	**Back**
High	i	e	u
Mid			o
Low		a	

Table 1. Tohono O'odham vowel inventory in the Alvarez-Hale orthography.

	Labial	**Dental**	**Alveolar**	**Retroflex**	**Palatal**	**Velar**	**Glottal**
Plosives	p b	t d	ḍ			k g	'
Fricatives			s	ṣ			h
Affricates					c j		
Nasal	m		n		ñ	ŋ	
Flap				l			
Approximant	w				y		

Table 2. Tohono O'odham consonantal inventory the Alvarez-Hale orthography.

O'odham also includes geminates, most frequently used to indicate the distributive aspect (e.g. *ta:ttam* 'teeth.DIST,' *babbad* 'frogs.DIST,' *kokktoñ* 'shirts.DIST; see also §3.5). All word-initial vowels are preceded by a glottal stop (e.g. *'a:cim* 'we, us,' *'eḍa* 'inside, in,' *'u:s* 'wood, stick, board'). Glottal stops may also occur word-internally between two vowels (e.g. *do'ag* 'mountain,' *ga'a* 'roast.IPFV,' *'gigi'ik* 'eight'). O'odham also includes vowel length contrasts (e.g. regular *e* versus long *e:*)[2] which can be phonemic (e.g. *'ep* 'again' versus *'e:p* 'another,' *'i*

[2] Phonemically, /ɨ/ versus /ɨ:/.

'here' versus *'i:* 'drink.PFV'). Long vowels occur only within stressed syllables, which are also most typically word-initial; the exception to this pattern is with borrowed words. Both Zepeda (1983, 1984) and Mathiot (1973) additionally mark short vowels produced word-finally (e.g. *'uwĭ* 'woman,' *su:dagĭ* 'water,' *nowĭ* 'hand'),[3] however, this distinction is not phonemic. Additionally, the schwa [ə] is pervasive in O'odham, although it also is not phonemic. Its appearance has been argued to be the result of vowel reduction (Hill & Zepeda, 1992).

3.1.2 O'odham prosody and syllable structure

Primary stress is prominent on the initial syllable of independent words while secondary stress falls on alternating (i.e. odd) syllables (Fitzgerald, 2002, p. 256). Prefixes, however, cannot carry primary stress (see Hill & Zepeda, 1992, p. 255). In the presence of a prefix, the primary stress shifts to the second syllable (or, in other words, remains on the first syllable of the root). For example, *céggia* 'fight' in isolation carries stress on the first syllable. When an object pronominal is prefixed to the verb, the stress is on the second syllable *ha-céggia* 'fight them,' not **há-ceggia.*

Regarding syllable shapes, a wide range is possible; this includes function words, which can consist of a consonant (C), a vowel (V), the combination of CV, or even CC (e.g. *g* 'the/a,' *o* 'will/would,' *pi* 'not,' *kc* 'and,' respectively). Lexical (or content) words may consist of simple syllable shapes such as CV on one end of the spectrum and words comprised of very complex codas on the other end. Examples of various syllable shapes are:

[3] Mathiot's (1973) employs the diaeresis [¨] to mark word-final short vowels.

CV	***ku**ḍu*	/kudu/	‘troubled/worried’
CVV	*ki:*	/ki:/	‘house’
CVC	*ce**mit***	/ʧɨmit/	‘tortilla’
CVVC	*to:n*	/to:n/	‘knee’
CVCC	*gogs*	/goks/	‘dog’
CVCCC	*wopsk*	/wopsk/	‘paternal grandfathers’

(Beers, Hirrel, Kerfoot, & Sabogal, 2011, pp. 4-5)

The patterns above illustrate that while complex codas are possible and not unusual in O'odham, complex onsets are rarely present in lexical words that are native to the language (i.e. loan words are the exception to this rule).

3.2 Morphological typology

O'odham is a mildly synthetic language, allowing some affixation that is predominantly agglutinative with clear and easily discernable boundaries. UA languages overall can range anywhere from analytic to polysynthetic (Haugen, 2008a). Verbs in O'odham can potentially carry the greatest degree of affixation, but affixation is minimal relative to some other UA languages such as Huichol (Langacker, 1977, p. 157), the latter of which illustrates great morphological complexity. Verbs in O'odham indicate imperfective versus perfective aspects (see §3.4), but otherwise it is possible for intransitive verbs to carry no additional morphological marking and for transitive verbs to carry only the obligatory pronominal prefixes (see §3.3). If suffixes are attached to verbs, it may only be a single suffix such as the progressive marker *-him* in (1) below. In other cases there may be two suffixes, as shown in example (2) below. All morphological glosses throughout the present chapter are my own unless otherwise noted. All free translations are the original respective authors'.

(1) 'Ab 'o ha-ṣa:muda-**him**.
LOC 3.AUX.IPFV 3.PL.OBJ-herd-**PROG**
'He/She is herding them along.' (Zepeda, 1984, p. 102)

(2) Wagt-**da-ñ**! [4]
dig-**DUR-IMP**
'Be digging it!' (Mathiot, n.d.a, p. 59)

Other areas of the language may also employ some affixation. For example, a series of different genitive prefixes and suffixes may be marked on the noun:

(3) **ñ**-je'e
1SG.POSS-mother
'my mother' (Zepeda, 1983, p. 76)

(4) kotoñ-**ij**
shirt-**3SG.POSS**
'his/her shirt' (p. 76)

(5) **t-**haiwañ-**ga**
1PL.POSS-cattle-**ALIEN.POSS**
'our cattle' (p. 78)

Derivational morphological processes are also possible with nouns. For instance, O'odham and other UA languages are known to employ denominal morphology (Langacker, 1977; Haugen, 2007, 2008a, 2008b; Zepeda 1983) or suffixes which derive verbs from nouns as illustrated in examples (6)-(8).

(6) 'asugal-**mad**
sugar-**DNMLZ.add**
'sugaring' or 'add sugar to' (Zepeda, 1983, p. 90)

(7) wopo-**pig**'
fur-**DNMLZ.remove**
'removing fur from the skin of an animal' (p. 93)

[4] All data is provided in the Alvarez-Hale orthography for the sake of uniformity throughout this work regardless of the orthography employed in the original sources.

(8) *maḍ-**t***
child[of.a.woman]-**DNMLZ.make**
‘to bear offspring’; also ‘to bear fruit' (Saxton et al., 1983, p. 39)

This process is argued to be a historical remnant of noun incorporation (Langacker, 1977; Haugen, 2007, 2008a, 2008b).

Postpositions may also be marked for person and the number of objects they refer to with a series of prefixes:

(9) Mali:ya 'o 'am **ñ**-baic dahă.
Mary 3.AUX.IPFV LOC **1SG.OBJ**-in.front.of sit.IPFV
'Mary is/was sitting in front of me.' (Zepeda, 1983, p. 49)

(10) Ju:kĭ 'o 'ab **t**-wui him.
rain 3.AUX.IPFV LOC **1PL.OBJ**-toward come.IPFV
'The rain is coming (*literally* walking) toward us.' (Zepeda, 1983, p. 49; original emphasis)

(11) Klisti:na 'o 'an **ha**-hugidan dahă.
Christina 3.AUX.IPFV LOC **3PL.OBJ**-next.to sit.IPFV
'Christina is/was sitting next to them.' (Zepeda, 1983, p. 49)

In sum, while complex morphological patterns are not present in O'odham relative to what can be observed in other UA languages, some affixation is possible and pervasive in certain parts of the grammar.

3.3 Coding of participants

O'odham does not employ a nominal case system to indicate verbal arguments, and word order is also not a reliable device in establishing grammatical relations given its relative flexibility (see §3.7). Instead, O'odham utilizes other mechanisms – namely a complex auxiliary system and pronominal prefixes on the verb – to code participants.

Auxiliaries obligatorily mark subjects. They additionally reflect imperfective and perfective aspects. The auxiliaries are shown in Table 3 in both full and truncated forms (all

of which likely grammaticalized from the independent pronouns in the language given their resemblance in form; compare Table 3 with the independent pronouns in Table 4).

	SINGULAR		**PLURAL**	
	Imperfective	**Perfective**	**Imperfective**	**Perfective**
1st person:	*'añ/ñ*	*'ant/nt*	*'ac/c*	*'att/tt*
2nd person:	*'ap/p*	*'apt/pt*	*'am/m*	*'amt/mt*
3rd person:	*'o*	*'at/t*	*'o*	*'at/t*

Table 3. Auxiliaries indicating person, number, and aspect; full and truncated forms.

Independent pronouns are optional and are usually used for pragmatic purposes such as "focus, contrast, and other pragmatic effects"(Haugen, 2008b, p. 149). No distinction is made between subject and object pronouns. Table 4 below lists the independent pronouns in their full and truncated forms.

	SINGULAR		**PLURAL**	
1st person:	*'a:ñi/'añ*	'I' 'me'	*'a:cim/'a:c*	'we' 'us'
2nd person:	*'a:pi/'a:p*	'you'	*'a:pim/'a:p*	'you'
3rd person:	*hegai/heg*	'he, she it, that' 'him, her, it, that'	*hegam/heg*	'they, those' 'them, those'

Table 4. Optional independent pronouns; full and truncated forms (Zepeda, 1983, p.159).

Auxiliaries almost always appear in sentential second position (for more on word order, see §3.7). This is illustrated in examples (12)-(14) below, where the auxiliaries are given in their imperfective forms. Independent pronouns in parenthesis as in (12) and (13) are optional. If the independent pronouns in (12) or (13) were not explicitly stated, this would illustrate one of the few contexts in which the auxiliary would technically not be in second position.

(12) ('A:ñi) **'añ** s-ba:bigĭ ñeok.
(1SG) **1SG.AUX.IPFV** STAT-slow speak.IPFV
'I am/was speaking slowly.' (Zepeda, 1983, p. 18)

(13) ('A:cim) **'ac** ha-ñeid g ki:k.
(1PL) **1PL.AUX.IPFV** 3PL.OBJ-see.IPFV DET houses.
'We are/were seeing the houses.' (Zepeda, 1983, p. 33).

(14) 'Uwĭ **'o** ñeid g ceoj.
woman **3.AUX.IPFV** see.IPFV DET boy.
'The woman sees the boy.' (Zepeda, 1983, p. 62)

Examples (15)-(17) below show the perfective forms of the same auxiliaries.

(15) 'A:ñi **'añt** ñeo.
1SG **1SG.AUX.PFV** speak.PFV
'I spoke.' (Zepeda, 1983, p. 61)

(16) 'A:cim **'att** hihi 'am tianda wui.
1PL **1PL.AUX.PFV** walk.PFV LOC store to.
'We walked to the store.' (p. 63).

(17) 'Uwĭ **'at** ñei g ceoj.
woman **3SG.AUX.PFV** see.PFV DET boy.
'The woman saw the boy.' (p. 62)

Objects are obligatorily marked on transitive and ditransitive verbs with pronominal morphology as listed in Table 5.

	SINGULAR		**PLURAL**	
1st person:	*ñ-*	'me'	*t-*	'us'
2nd person:	*m-*	'you'	*'em-*	'you'
3rd person:	—	'him, her, it'	*ha-*	'them'

Table 5. Obligatory object markers prefixed to transitive and ditransitive verbs.

Obligatory object marking on transitive or ditransitive verbs is illustrated below in (18)-(20). In example (18), *ha-* indexes the third-person, plural object *mimstol '*cats,' in example (19), *m-* reflects the second person, singular, object marker, and in (20), *t-* marks the first-person plural indirect object on the ditransitive verb *ma:k* 'give.'

(18) Mimstol 'o **ha**-huhu'id g gogs.
cats 3.AUX.IPFV **3PL.OBJ**-chase.IPFV DET dog.
'The dog is/was chasing the cats.' (Zepeda, 1983, p. 34)

(19) Klisti:na 'o ('a:pi) **m**-cendad.
Christina 3.AUX.IPFV (2SG) **2SG.OBJ**-kiss[5]
'Christina is/was kissing you.' (p. 36).

(20) Husi 'o g daikuḍ ('a:cim) **t**-ma:k.
Joe 3.AUX.IPFV DET chair (1PL) **1PL.OBJ**-ma:k.IPFV
'Joe is/was giving a chair to us.' (p. 38)

Finally, third-person, singular objects are zero-marked.

(21) Cehia 'o g 'ali **Ø**-ñu:kud.
girl 3.AUX.IPFV DET child **3SG.OBJ**-take.care.of
'The girl is/was taking care of the child.' (p. 32)

3.4 Verbs and aspect

As alluded to, O'odham utilizes a rich aspectual system. Aspect has been described cross-linguistically as expressing "different ways of viewing the internal temporal constituency of a situation" (Comrie, 1976, p. 3). Aspect is "not concerned with relating the time of the situation to any other time-point," as is the case with tense distinctions, "but rather with the internal temporal constituency of the one situation" (p. 5). Dahl (1985) similarly describes aspect as having "to do with the structure of the things going on or taking place in the situation described by the sentence" (p. 24). It is non-deictic in contrast to tense, the latter of which describes when some event or state occurs in relation to the moment of speech. Croft (2012) defines it as the "domain of grammar for the lexical and grammatical realization of how EVENTS unfold over time" (p. 398; original emphasis; also see p. 30),

[5] Verbs that are not glossed for the imperfective or perfective utilize identical forms (e.g. *cendad* 'kiss' is the form for either imperfective or perfective).

where EVENTS refer to all lexical aspectual categories (p. 34). Functionally, Bybee (1985) states, aspect allows "the temporal dimensions of a situation to be described from different points of view depending on how the situation is intended to fit into the discourse" (p. 142). Aspect describes the "temporal contours of a situation" (Bybee et al., 1994, p. 317).

This section presents a few of the most basic ways that O'odham verbs inflect for aspect; namely, the imperfective, perfective, future imperfective and future perfective, as well as the progressive (see Mason, 1950; Hale, 1959; Saxton, 1982; Mathiot, 1973 for a fuller treatment of O'odham aspect).

Verbs in O'odham are, in a majority of cases, distinguished between either the imperfective or perfective aspects by a change in form. The imperfective versus perfective aspects are most commonly marked by either the presence or absence of word-final consonants, respectively (Zepeda, 1983, p. 60; 1984, pp. 109-110). That is, imperfective verbs are considered the full form of verbs, while the perfective forms are derived by "subtractive morphology" (Steele, 2008, p. 318). This applies to both singular and plural forms as can be seen in the list of example verb forms in Table 6 below.

SINGULAR				**PLURAL**			
Imperfective		**Perfective**		**Imperfective**		**Perfective**	
*hi**m***	'walking'	*hi:*	'walked'	*hihi**m***	'walking'	*hihi*	'walked'
*hi:n**k***	'barking'	*hi:n*	'barked'	*hihin**k***	'barking'	*hihin*	'barked'
*ñeo**k***	'speaking'	*ñeo*	'spoke'	*ñeño**k***	'speaking'	*ñeñeo*	'spoke'
*ñei**d***	'seeing'	*ñei*	'saw'	*ñei**d***	'seeing'	*ñeo*	'saw'

Table 6. Regular imperfective and perfective verb forms (based on Zepeda, 1983, p. 59).

In other instances, both the final vowel and consonant are dropped to reflect the perfective:

SINGULAR				PLURAL			
Imperfective		**Perfective**		**Imperfective**		**Perfective**	
*cepos**id***	'branding'	*cepos*	'branded'	*cecpos**id***	'branding'	*cecpos*	'branded'
*cipk**an***	'working'	*cipk*	'worked'	*cicpk**an***	'working'	*cicpk*	'worked'
*gegos**id***	'eating, feeding'	*gegos*	'ate, fed'	*gegos**id***	'eating, feeding'	*gegos*[6]	'ate, fed'
*hu:kaj**id***	'warming'	*hu:kaj*	'warmed'	*hu:kaj**id***	'warming'	*hu:kaj*	'warmed'

Table 7. Irregular imperfective and perfective verb forms (based on Zepeda, 1983, p. 60 & glossary).

How these "subtractive" patterns for the perfective in O'odham evolved is not entirely clear; however, Langacker (1977) offers a possible explanation for this pattern in various UA languages based on reconstructions for PUA. Langacker (1977) notes that it has been postulated that CV-reduplication was historically the marker for the perfective in PUA (p. 130). Reduplication in combination with stress patterns led to "truncation of final syllables in perfective stems" (p. 130). Based on this, Langacker further states, "[W]e can speculate that reduplication affected placement of accent, which in turn facilitated truncation and other phonological changes in the unaccented portions of the stem" (p. 130). It may be the case that the subtractive pattern observed in O'odham for the perfective is a remnant of this phonological development.

There are also cases in which perfective forms may be "very irregular" (Zepeda, 1983, p. 61). A few examples are given in Table 8.

[6] Singular and plural forms are identical for this verb.

SINGULAR				PLURAL			
Imperfective		**Perfective**		**Imperfective**		**Perfective**	
ko:ṣ	'sleeping'	*koi*	'slept'	*ko:ks*	'sleeping'	*ko:k*	'slept'
na:d	'making a fire'	*nai*	'made a fire'	*nanda*	'making a fire'	*nai*	'made a fire'
wupḍa	'roping, tying'	*wu:*	'roped, tied'	*wupḍa*	'roping, tying'	*wu:*	'roped, tied'
meḍ	'running'	*me:*	'ran'	*wo:po'ŏ*	'running'	*wo:p*	'ran'

Table 8. "Very irregular" imperfective and perfective verb forms (Zepeda, 1983, p. 61 & glossary).

Finally, Table 9 provides some examples where imperfective and perfective verbs are identical.

SINGULAR				PLURAL			
Imperfective		**Perfective**		**Imperfective**		**Perfective**	
gaswua	'combing'	*gaswua*	'combed'	*gagswua*	'combing'	*gagswua*	'comb'
ceggia	'fighting	*ceggia*	'fought'	*ceceggia*	'fighting'	*ceceggia*	'fought'
ka:	'hearing'	*ka:*	'heard'	*ka:*	'hearing'	*ka:*	'heard'
cicwi	'playing'	*cicwi*	'played'	*cicwi*	'playing'	*cicwi*	'played'

Table 9. Identical imperfective and perfective verb forms (Zepeda, 1983, p. 61 & glossary).

In order to illustrate the imperfective and perfective verbs in context, comparative sentences are provided in (22)-(24). The (a) sentences in each set include the imperfective forms, and the (b) sentences illustrate the perfective forms, which are expressed via truncation. Note that the imperfective and perfective aspects are also marked on the auxiliary as a type of agreement.

(22) a. Ceoj 'o **ñeok**.
boy 3.AUX.IPFV **speak.IPFV**
'The boy is/was speaking.' (Zepeda, 1983, p. 62)

b. Ceoj 'at **ñeo**.
boy 3.AUX.PFV **speak.PFV**
'The boy spoke.' (p. 62)

(23) a. 'A:pi 'ap **him** 'am ki: wui.
2SG 2SG.AUX.IPFV **walk.IPFV** LOC house to
'You are/were walking to the house.' (p. 62)

b. 'A:pi 'apt **hi:** 'am ki: wui.
2SG 2SG.AUX.PFV **walk.PFV** LOC house to
'You walked to the house.' (p. 62)

(24) a. 'A:cim 'ac **cicpkan** 'am ki: webig.
1PL 1PL.AUX.IPFV **work.PL.IPFV** LOC house behind
'We are/were working behind the house.' (p. 62)

b. 'A:cim 'att **cicpk** 'am ki: webig.
1PL 1PL.AUX.PFV **work.PLPFV** LOC house behind
'We worked behind the house.' (p. 62)

In the case that a verb utilizes the identical form for both aspects, the auxiliary clearly indicates which aspect is expressed by the verb, as in example (25) below.

(25) a. Hegai cehia **'o** **cicwi**.
that girl **3.AUX.IPFV** **play**
'That girl is/was playing.' (Zepeda, 1983, p. 64)

b. Hegai cehia **'at** **cicwi**.
that girls **3.AUX.PFV** **play**
'That girl played.' (Constructed example)

Marking of aspect on the verb may also combine with the future particle *o* to indicate the future imperfective and future perfective aspects. In these cases, the future marker "always comes immediately before the verb" (Zepeda, 1983, p. 63).

The more complicated of the two patterns is the future imperfective. In addition to the *o* future particle, future imperfective constructions also require the suffix *-ad/-d* to attach to the imperfective verb form (i.e. the full form); however, the auxiliary in this case occurs in the perfective aspect even though the construction itself is not perfective. That is, if the future particle *o* is present, then the auxiliary, by default is in the perfective form; it is the presence

of *-ad/-d* that signals that the sentence is in the future imperfective. For example, in (26) the third-person, singular, perfective auxiliary *'at* is used in combination with the future particle *o* which immediately precedes the verb. The future imperfective suffix, *-ad*, attaches to the verb, *ñeok* 'speak,' which is in its imperfective form (the perfective would be *ñeo*). The identical pattern is also represented in (27) and (28).

(26) Ceoj 'at **o** ñeok-**ad**.
boy 3.AUX.PFV **FUT** speak.IPFV-**FUT.IPFV**
'The boy will be speaking.' (Zepeda, 1983, p. 72)

(27) Hegam kekel 'at **o** si ñe'e-**d.**
Those old.men 3.AUX.PFV **FUT** INTS sing.IPFV-**FUT.IPFV**
'Those old men will really be singing." (p. 72)

(28) N-apt 'i:ya **o** cicpkan-**ad** 'a:pim?
INTER-2SG.AUX.PFV here **FUT** work.IPFV-**FUT.IPFV** 2SG
'Will you be working here?' (p. 72)

Formation of future perfective constructions is comparatively straightforward. The future particle *o* precedes the verb in its perfective (i.e. subtractive) form. The auxiliary is also in the perfective form:

(29) Ceoj 'at **o** **ñeo**.
boy 3.AUX.PFV **FUT** **speak.PFV**
'The boy will speak.' (Zepeda, 1983, p. 63)

(30) 'A:cim 'att **o** **hihi** 'am tianda wui.
1PL 1PL.AUX.PFV **FUT** **walk.PL.PFV** LOC store to
'We will walk to the store.' (p. 63)

(31) 'Uwĭ 'at **o** **ñei** g ceoj.
woman 3.AUX.PFV **FUT** **see.PFV** DET boy
'The woman will see the boy' (p. 63)

Verbs may also inflect for the progressive aspect via the suffix *-him* (Hale, 1959, p. 72; Mason, 1950, p. 53; Saxton, 1982, p. 216; see discussion in Mathiot, 1973, pp. 57-58 on

the "interruptive marker"). In Zepeda's (1984) dissertation, the suffix is described as specifying "an exten[s]ion of the action…" (p. 104), which Zepeda frequently translates as 'going along.' This choice is likely due the fact that *-him,* sometimes also appearing in the form *-im,* is a grammaticalized form of the verb *him* 'walk.IPFV' (Mason, 1950, p. 52). I provide a more detailed description of progressive aspect and a number of illustrative examples in Chapter 5.

When the suffix attaches, the verb is in the imperfective (i.e. full) form as illustrated below in examples (32)-(34).

(32) 'Ab 'o ha-ṣa:mud-a-**him.**
LOC 3.AUX.IPFV 3PL.OBJ-herd.IPFV-EPEN-**PROG**
there AUX them:herding:along[7]
'He/She is herding them along.' (Zepeda, 1984, p. 102)

(33) Husi 'o si gikuc-**him** c 'ab him.
Joe 3.AUX.IPFV INTS whistle.IPFV-**PROG** CONJ LOC walk.IPFV
Joe AUX very whistling:along and there walking
'Joe is walking along and is really whistling.' (p. 104)

(34) 'Ab 'o him hegai c ñeok-**im**.
LOC 3.AUX.IPFV walk.ipfv 3SG CONJ talk.IPFV-**PROG**
there AUX walking him and talking:along
'He is walking along talking.' (, p. 103)

Beyond what I describe here regarding verbal morphology, the reader is directed to Hale (1959), Mason (1950), Mathiot (1973), and Saxton (1983).

3.5 Number

The primarily means that O'odham uses to express plural number for nouns and verbs is reduplication. Phonologically, the reduplicative process is quite complex (for an in-depth

[7] Third line in (32)-(34) is Zepeda's original glossing.

discussion, see Beers, Cruz, Hirrel, & Kerfoot, 2014; Fitzgerald, 2003; Hill & Zepeda, 1998), but in oversimplified terms, the reduplicant consists roughly of a copy of the initial CV of the singular or base forms.

	SINGULAR		**PLURAL**	
Nouns	*goks*	'dog'	***go****goks*	'dogs'
	maḍ	'child (of a woman)'	***ma:****maḍ*	'children (of a woman)'
	tonk	'hill'	***to****tonk*	'hills'
	mi:stol	'cat'	***mi****mstol*	'cats'
	'uwĭ	'woman, girl'	***'u****'uwĭ*	'women, girls'
Verbs	*hi:nak*[8]	'barking'	***hi****hinak*	'barking'
	ñeok	'speaking'	***ñe****ñok*	'speaking'
	ba:ñimad '	'crawling'	***ba:****bañimad*'	'crawling'
	ko:ṣ	'sleeping'	***ko:****kṣ*	'sleeping'
	'eñigadad	'getting dressed (sg)'	***'e****'eñigadad*	'getting dresses'

Table 10. Base and reduplicated forms of nouns and verbs (Zepeda, 1983).

Steele (2008) additionally points out that there are many cases in which the plural is indicated with suppletive forms (Table 11) or identical forms for singular and plural (Table 12).

SINGULAR		**PLURAL**	
ṣoak	'crying (IPFV)'	*ṣoañ*	'crying (IPFV)'
meḍ	'running (IPFV)'	*wo:po'o*	'running (IPFV)'
ka:c	'lying (INAN OBJ, IPFV)'	*we:c*	'lying (INAN OBJ, IPFV)'
jiwa	'arrived (PFV)'	*dada*	'arrived (PFV)'

Table 11. Suppletive forms for plural (Zepeda, 1983).

Singular		**Plural**	
judum	'bear'	*judum*	'bears'
naw	'prickly pear'	*naw*	'prickly pears'
ñeid	'seeing (IPFV)	*ñeid*	'seeing (IPFV)
golon	'raking (IPFV)'	*golon*	'raking (IPFV)'

Table 12. Identical forms for singular and plural (from Steele, 2008: 316; based on Zepeda, 1983).

[8] All verbs in Table 10 given in imperfective forms.

In Steele's (2008) description of data in Zepeda (1983), Steele also draws attention to examples that are "lacking a 'plural' counterpart" (p. 317). Examples are included in Table 13 below.

SINGULAR	**PLURAL**
ṣu:dagĭ 'water'	—
wa:ga 'dough'	—
ku:bs smoke, dust'	—

Table 13. Words that lack a plural counterpart (from Steele, 2008, p. 316; based on Zepeda, 1983).

It is not explicitly stated on what grounds the "identical" singular and plural nouns like those in Table 12 are distinguished from examples that lack a plural altogether. Presumably, this is semantically based. The substances in Table 13 are commonly construed as mass objects lacking individuated, countable parts, while the plurals objects such as *judum* 'bears' and *naw* 'prickly pears' are construed as countable from the perspectives of Steele (2008) and Zepeda (1983).

Finally, in some cases, gemination within the reduplicated noun form is used to express distributive nouns. The morpho-phonological process "results in the initial syllable of the word being closed by a geminate" (Fitzgerald, 2003, p. 65). Examples are illustrated in Table 14.

SINGULAR	**PLURAL**	**DISTRIBUTIVE**
hodai 'rock'	*hohodai* 'rocks'	*hohhodai* 'rocks'
nowĭ 'hand'	*no:nowĭ* 'hands'	*nonnowiu* 'hands'
kahio 'leg'	*kakio* 'legs'	*kakkio* 'legs'
daikuḍ 'chair'	*dadaikuḍ* 'chairs'	*daddaikuḍ* 'chair'

Table 14. Singular, plural, and distributive forms (from Fitzgerald, 2003, p. 65).

Fitzgerald (2003) states that geminate formations can also occur in some reduplicated verb forms to express the repetitive aspect, described by Mathiot (1973) as indicating "several identical actions in a single locus" (p. 49).

UNITIVE	REPETITIVE	GLOSS
gew	*geggew*	*'to hit object'*
hehem	*hehhem*	*'to laugh'*
kow	*kokkow*	*'to dig object out of ground'*
ñea	*ñenna*	*'to look in a certain direction'*

Table 15. Gemination in repetitive verbs (from Fitzgerald, 2003, p. 67).

3.6 Alignment and grammatical relations

O'odham is described as a nominative-accusative language (Hale, 2002), and UA languages as whole are described as following this alignment pattern (Langacker, 1977). The most explicit indication of this alignment is various methods for marking the accusative, while nominative case is unmarked (see Langacker, 1977, pp. 82-85 for a fuller account). For example, the suffix **-*a* has been reconstructed as the accusative noun-marker for Proto-Uto-Aztecan (PUA). Overt accusative inflection on nouns remains present in primarily northern UA languages, while it "has largely been lost in the southern languages…" (p. 83). O'odham falls into the latter category. Instead, a nominative-accusative system is reflected namely via the auxiliaries and pronominal morphology mentioned earlier (§3.3) on the verb. Another area of O'odham grammar, however, reflects an ergative-absolutive pattern. This is evident when verbal stem number (expressed via reduplication) is considered. I describe both alignment patterns in the subsequent sections.

3.6.1 Auxiliaries and pronominal morphology: Subject and object indexation

As stated in §3.3, subjects are obligatorily indexed by the auxiliary system which also marks aspect; more specifically, auxiliaries always reflect the person and number of single arguments (S) of intransitive clauses and the most agent-like arguments (A) of agentive transitive clauses, as well as aspect. The auxiliaries from Table 3 are repeated below in Table 16.

	SINGULAR		**PLURAL**	
	Imperfective	**Perfective**	**Imperfective**	**Perfective**
1st person:	*'añ/ñ*	*'ant/nt*	*'ac/c*	*'att/tt*
2nd person:	*'ap/p*	*'apt/pt*	*'am/m*	*'amt/mt*
3rd person:	*'o*	*'at/t*	*'o*	*'at/t*

Table 16. Auxiliaries indicating person, number, and aspect (full and truncated forms).

The most patient-like arguments (P) of transitive clauses, on the other hand, are obligatorily marked on the verb in the form of pronominal prefixes. These reflect person as well as number as shown in Table 17 (repeated from Table 5).

	SINGULAR	**PLURAL**
1st person:	*ñ-* 'me'	*t-* 'us'
2nd person:	*m-* 'you'	*'em-* 'you'
3rd person:	— 'him, her, it; that'	*ha-* 'them; those'

Table 17. Obligatory object markers prefixed to transitive and ditransitive verbs.

Auxiliaries and object pronominals combined can be said to treat S and A the same and P differently, reflecting a nominative-accusative alignment. That is, auxiliaries always index the person and number of S and A identically in intransitive and transitive clauses, respectively, and pronominal marking on the verb indexes the P of transitive clauses. Examples (35)-(37) are intransitive sentences in which the auxiliaries index their Ss. The

transitive sentences in (38)-(40) show the same patterns with respect to the Ss and auxiliaries. (Note: Third-person, imperfective and perfective auxiliaries do not distinguish for number.)

Intransitive examples:

(35) Kawyu **'o** meḍ.
horse **3.AUX.IPFV** run.IPFV
'The horse is/was running.' (Zepeda, 1983, p. 13)

(36) 'A:ñi **'an** s-ba:bigĭ ñeok.
1SG **1SG.AUX.IPFV** STAT-slowly speak.IPFV
'I am/was speaking slowly.' (p. 19)

(37) 'A:pim **'am** 'i:ya cicwi.
2PL **2PL.AUX.IPFV** here play.IPFV
'You (pl) are/were playing right here.' (p. 19)

Transitive examples:

(38) Gogs **'o** hegam ha-huhu'id
dog **3.AUX.IPFV** 3PL 3PL.OBJ-chase.IPFV
'The dog is/was chasing them.' (p. 34)

(39) 'A:ni **'an** g haiwañ Ø-ceposid.
1SG **1SG.AUX.IPFV** DET cow 3SG.OBJ-brand.IPFV
'I am/was branding a cow.' (p. 32)

(40) 'A:cim **'ac** g ki: Ø-kegcid.
1PL **1PL.AUX.IPFV** DET house 3SG.OBJ-clean.IPFV
'We are/were cleaning the house.' (p. 32)

The sentences in (41)-(43) illustrate P marking on the verb in the form of pronominal prefixes for the following transitive examples.

(41) Hegai 'uwi 'o 'a:pim '**em**-ñu:kud
that woman 3.AUX.IPFV 2PL **2PL.OBJ**-take.care.of
'That woman is/was taking care of you (pl).' (Zepeda, 1994, p. 36)

(42) Kli:stina 'o ('a:pi)[9] **m**-cendad.
Christina 3.AUX.IPFV (2SG) **2SG.OBJ**-kiss
'Christina is/was kissing you.' (p. 36)

(43) Hegai 'ali 'o 'a:cim **t**-kuḍut.
that child 3.AUX.IPFV 1PL **1PL.OBJ**-bother.IPFV
'That child is/was bothering us.' (p. 36)

If the sentence is ditransitive, including both a direct and indirect object, then it is the latter that is expressed in the form of the pronominal prefix. In (44) *m-* represents the second-person, singular, indirect object, not the singular direct object, *ho'ok* 'legend.'[10]

(44) ('A:ñi) 'ant o ho'ok **m**-'a'agid-ad…
1SG 1SG.AUX.PFV FUT legend **2SG.OBJ**.-tell-FUT.IPFV
' I will be telling you a legend…' (Modified example from Mathiot, n.d.a, p. 290)

What this also suggests is that O'odham more accurately employs a primary-secondary object system with regards to the pronominal morphology, rather than a direct-indirect object system; that is, the pronominal verbal prefixes mark for primary objects (indirect objects of ditransitive verbs and direct objects of transitive verbs) while secondary objects (direct objects of ditransitive verbs) go unmarked. Put another way, O'odham pronominal prefixes treat primary objects the same while they treat secondary objects differently (for more on primary-secondary marking see Dryer 1986; also Payne 2013; Croft 2001).

To recap, the examples above illustrate a nominative-accusative alignment, where S and A are always indexed in the same manner within auxiliaries, while P is indexed differently; namely, via pronominal morphology on the verb. Additionally, with regards to

[9] Optional independent pronoun added to original example.
[10] *Ho'ok* carries this meaning when paired specifically with *'a'agid* 'tell' (Mathiot, n.d.b).

objects of multi-argument clauses, O'odham appears to make a primary-secondary object distinction.

3.6.2 Verbal number and alignment

Number expressed by the verb stem, itself, is another area of O'odham grammar that reflects grammatical relations and alignment. As discussed in §3.5, the verbal stem indicates number via reduplication of (roughly) the initial CV of the base form of the verb, suppletion, or, in some other cases, no change in form between the singular and plural (i.e. they are identical). With regard to verbal number, Steele (2008) discusses a pattern in O'odham in which intransitive verbs reflect the number of the S – either singular or plural, but in multi-argument clauses, verbal number aligns with the direct object, respectively. In other words, S and P are marked for number in the same manner, while A is marked differently. Verbs never align with the number of A. This has been described in the literature as reflecting an ergative-absolutive pattern (see Frajzyngier, 2011; also Durie, 1986 for a discussion on verb suppletion and number). In this case, number agreement is absolutive. Lindsey (2000) discusses a similar pattern with verbal number in Shoshoni, a Central Numic language of the UA family that is distantly related to O'odham.

Consider the following comparative examples from O'odham. The contrasting examples (45) and (46) are all intransitive. In (45)a and (46)b, the Ss, *gogs* 'dog' and *O'odham* 'person' are singular, and the verbs *hink* 'bark.SG' and *ñeok* 'speak.SG' are also in their singular (and imperfective) forms. In (45)b and (46)b, the Ss are plural: *gogoks* 'dogs' and *cecoj* 'boys,' and the verbs are also plural, *hihink* 'bark.PL' and *ñeñok* 'speak.PL,' reflecting the plural number of the Ss.

(45) a. Gogs 'o **hink**.
dog 3.AUX.IPFV **bark.SG.IPFV**
'The dog is/was barking.' (Zepeda, 1983, p. 13)

b. Gogogs 'o **hihink**.
dogs 3.AUX.IPFV **bark.PL.IPFV**
'The dogs are/were barking.' (p. 15)

(46) a. 'I:da O'odham 'o pi **ñeok.**
this person 3.AUX.IPFV NEG **speak.SG.IPFV**
'This person is/was not speaking.' (p. 9)

b. 'Idam cecoj 'o **ñeñok**.
These boys 3.AUX.IPFV **speak.PL.IPFV**
'These boys are/were speaking.' (p. 10)

Examples (47) and (48) below illustrate verbal number expressed by the stem for transitive constructions. In (47), note that the A in this case is plural (i.e. the independent first-person, plural pronoun, *'a:cim*) and the P, *wisilo* 'calf,' is singular. Note also that the verb stem, *ceposid*.SG 'brand' is in the singular form; instead of agreeing in number with the plural A, it reflects the singular P, *wisilo* 'calf.' Similarly, in (47), the A in this case is the first-person, singular, independent pronoun, *'a:ñi*, and the P, *wipsilo* 'calves,' is plural. The verb, following the pattern for number in transitive sentences, is in the plural form, *cecposid* 'brand.PL,' aligning in number with the plural P. The same patterns are found in (48)a and b; the verb forms *wakon* 'wash.SG' and *wapkon* 'wash.PL,' exhibit the number of the Ps *ma:gina* 'car' and *mamgina* 'cars,' respectively.

(47) a. 'A:cim 'ac g wisilo **ceposid**.
1PL 1PL.AUX.IPFV DET calf **brand.SG.IPFV**
'We are/were branding the calf.'

b. 'A:ñi 'añ g wipsilo ha-**cecposid**.
1SG 1SG.AUX.IPFV DET calves 3PL.OBJ-**brand.PL.IPFV**
'I am/was branding the calves.' (Zepeda, 1983, p. 33)

(48) a. Husi 'o **wakon** g ma:gina…
Joe 3.AUX.IPFV **wash.SG.IPFV** DET car
'Joe is/was washing the car…' (p. 37)

b. Cehia 'o ha-**wapkon** g mamgina
girl 3.AUX.IPFV 3PL.OBJ-**wash.PL.IPFV** DET cars
'The girl is/was washing the cars.' (p. 33)

Finally, in ditransitive sentences also, verbal stem number reflects the direct object or the P. In (49), the verb *ma:k* 'give.SG' is singular as is its direct object or P, *daikuḍ* 'chair,' and in (49) the verb for 'give.PL' is in its plural form, *mamk*, aligning with its plural P, *mimsa* 'tables'. The plural pronominal prefixes on the other hand, mark the Recipients, 'us' and 'them.'

(49) a. Husi 'o g daikuḍ t-**ma:k**.
Joe 3.AUX.IPFV DET chair 1PL.OBJ-**give.SG.IPFV**
'Joe is/was giving us a chair.'

b. 'A:ñi 'añ g mimsa ha-**mamk**
1SG 1SG.AUX.IPFV DET tables 3PL.OBJ -**give.PL.IPFV**
'I am/was giving them tables.' (p. 38)

In sum, while other parts of O'odham grammar reflect an overall nominative-accusative pattern as stated by Hale (2002) and as Langacker (1977) has argued for the UA language family as a whole, verbal stem number in O'odham patterns differently (Steele, 2008), reflecting an ergative-absolutive system.

3.7 Word order

The present section addresses O'odham word order, most specifically, basic sentence word order of the subject (S) of either single or multi-argument clauses in relation to the object (O) of transitive clauses, and the Verb (V). I also consider word order with regard to adpositions and genitive constructions in the sections that follow.

3.7.1 Basic sentence patterns

O'odham basic word order of S, O, and V in transitive sentences is variable, although Langacker (1977) states that the most "neutral" word order or "that which carries no special nuances or semantic value" (p. 24) for O'odham and other southern branch UA languages is VSO. However, on discourse pragmatic grounds, Payne (1987) argues that "no strongly preferred word order of subject, object, and verb can be identified. However, there is a strong order condition on types of information, and semantic factors play a role in the choice of syntactic role" (p. 785). Similarly, Hale (2002) remarks that "[t]he position of arguments and many adjuncts is … governed by principles of discourse" (p. 303). O'odham variable word order of S, O, and V is illustrated in examples (50) based on Zepeda (1983).

(50)

Huhu'id	'o	g	ban	g	cu:wĭ.
chase.IPFV	3.AUX.IPFV	DET	coyote	DET	jackrabbit.
V	(AUX)		S		O

Ban	'o	g	cu:wĭ	huhu'id.
coyote	3.AUX.IPFV	DET	jackrabbit	chase.IPFV
O	(AUX)		S	V

Cu:wĭ	'o	huhu'id	g	ban.
jackrabbit	3.AUX.IPFV	chase.IPFV	DET	coyote.
O	(AUX)	V		S

Huhu'id	'o	g	cu:wĭ	g	ban.
chase.IPFV	3.AUX.IPFV	DET	jackrabbit	DET	coyote.
V	(AUX)		O		S

Cu:wĭ	'o	g	ban	huhu'id.
jackrabbit	3.AUX.IPFV	DET	coyote	chase.IPFV
O	(AUX)		S	V

'The coyote is/was chasing the jackrabbit.' (Zepeda, 1983, p. 31)[11]

[11] Zepeda provides a single, straightforward translation for the different word orders most likely for the sake of pedagogical simplicity in her grammar which is intended for beginning O'odham language learners; however, on discourse pragmatic grounds, the possible word orders highlight different identifiability statuses of the arguments.

As illustrated in the examples above, the positions of the arguments and verbs can vary. Note, however, that the auxiliary is always in second position regardless of the ordering of S, O, and V. This is the case in declarative sentences.

The second position of the auxiliary continues to maintain consistency in other circumstances, as well. As one example, when a sentence includes the negative marker *pi*, it can either precede or follow the auxiliary, as illustrated in (51).

(51) 'I:da O'odham 'o pi ñeok.
this person 3.AUX.IPFV NEG speak.IPFV
S (AUX) (NEG) V

Pi 'o ñeok 'i:da O'odham.
NEG 3.AUX.IPFV speak.IPFV this person.
(NEG) (AUX) V S
'This person is/was not speaking.' (Zepeda, 1983, p. 9)

The verb, however, cannot occur in sentence initial position in the presence of *pi*, as shown in (52) below.

(52) *Neok 'o pi 'i:da O'odham.
speak.IPFV 3.AUX.IPFV NEG this person
V (AUX) (NEG) S
'This person is/was not speaking.' (Zepeda, 1983, p. 9)

In sum, while VSO is arguably the neutral sentence word order in O'odham, in reality, word order can be varied. The order of S, V, and O instead are driven by discourse factors (see Payne, 1987 for a full description). There do exist, however, various other components of O'odham grammar that are more rigid.

3.7.2 Adpositions

Zepeda (1983) describes O'odham as utilizing postpositions as illustrated in examples (53)-(55) below.

(53) 'Uwĭ 'o 'am ki: **ba:ṣo** ke:k.
woman 3.AUX.IPFV LOC house **in.front** standing.
'The woman is/was standing in front of the house.' (Zepeda, 1983, p. 46)

(54) Huan 'o 'ab 'Cuk Ṣon **wui** him.
Juan 3.AUX.IPFV LOC Tucson **to** walk.
Juan is/was walking to Tucson (away from the speaker). (p. 47)

(55) Husi 'o gaḍhu cipkan to:nk **we:big**.
Jose 3.AUX.IPFV over.there.DIST work.IPFV hill **behind**.
'Jose is/was working over there behind the hill.' (p. 52)

Zepeda also points out, however, that in some instances, adpositions can precede the noun (i.e. they are prepositional).

(56) Ma:liya 'o 'am **ba:ṣo** ke:k g ki:
Mary 3.AUX.IPFV LOC **in.front** stand.IPFV DET house
'Mary is/was standing in front of the house.' (Zepeda, 1983, p. 130).

(57) Ma:liya 'o 'am **ba:ṣo** g ki: ke:k
Mary 3.AUX.IPFV LOC **in.front** DET house stand.IPFV
'Mary is/was standing in front of the house.' (p. 131).

Saxton (1982), on the other hand, states that the default pattern for O'odham is prepositional: "The preposition is predicate, preceding its argument in neutral word order…" (p. 189). Mathiot (1973) refers to postpositions as the default when discussing O'odham syntactic patterns; however, the author does not elaborate on this decision (p. 31). Mathiot (1973) simply states, "All Papago postpositions have dual syntactic functioning" (p. 31) occurring either before or following a noun. In UA as whole, both patterns are found (Langacker, 1977, pp. 93-94).

3.7.3 Genitive constructions

In genitive constructions the possessor precedes the possessed object as in (58)-(61) below (Zepeda, 1983). These examples also illustrate that genitive constructions can be formed

solely by juxtaposition (i.e. no case marking) in cases where the possessor is in the third person, singular.

(58) **'Ali** **je'e** 'at o cipk si'alim.
child **mother** 3.AUX.PFV FUT work.PFV tomorrow.
'The child's mother will work tomorrow.' (Zepeda, 1983, p. 75).

(59) **Mali:ya** **nawaṣ** 'o 'an wo'o mi:sa da:m.
Mary **pocket.knife** 3.AUX.IPFV LOC lie.IPFV table on.top.of
'Mary's pocket knife is lying on top of the table.' (p. 75)

(60) 'A:pi 'apt o wako g **Husi** **li:wa**.
2SG 2SG.AUX.PFV FUT wash.PFV DET **Joe** **jacket**.
'You will wash Joe's jacket.' (p. 75)

(61) 'Am do'ag we:big 'o ke:k g **'uwĭ** **ki:**.
LOC mountain behind 3.AUX.IPFV stand.IPFV DET **woman** **house**
'The woman's house is standing behind the mountain.' (p. 75)

When additional genitive case markings are present (e.g. third-person, plural possessors), the order remains the same; the possessor comes before the possessed object. In (62) below, the third-person, possessive prefix *ha-* is present, and in both (62) and (63) the alienable possessive suffix, *-ga,* is present.

(62) 'A'al **ha**-gogogs-**ga** 'at wo:p 'am ñ-ki: wui.
children **3PL.POSS**-dogs-**ALIEN.POSS** 3.AUX.PFV run.PL.PFV LOC 1SG.POSS-house to.
'The children's dogs ran to my house.' (Zepeda, 1983, p. 81).

(63) Ma:liya ha:l-**ga** 'o 'an dahă wo'ikuḍ da:m.
Mary squash-**ALIEN.POSS** 3.AUX.IPFV LOC sit.IPFV bed on.top.of
'Mary's squash is/was sitting on top of the bed.' (p. 83)

3.8 Locus of marking: Head/dependent marking

Overt morpho-syntactic marking in O'odham tends to be located on the head of a phrase when the two most basic phrase types that are "most prone to follow universal tendencies"

(Nichols & Balthasar, 2013, section 3, para. 2) are considered. Those are: "possessive phrase with noun possessor … and direct or primary objects in the transitive clause" (section 3, para. 2).

Considering, first, possessive constructions in which an overt noun possessor is present (as opposed to the presence of a pronoun/possessive pronominal marker), O'odham clearly marks alienable possessed object nouns with the possessive suffix, *-ga* (inalienable objects are unmarked; see Zepeda, 1983, p. 78-79). The possessive affixes fall on the head of the noun.

(64) 'Ali **gogs-ga** 'at koi 'am wo'ikuḍ weco.
child **dog-ALIEN.POSS** 3.AUX.PFV sleep.PFV LOC bed under
'The child's dog slept under the table.' (Zepeda, 1983, p. 81)

(65) Husi **ma:gina-ga** 'at pi me:.
Joe **car-ALIEN.POSS** 3.AUX.PFV NEG run.PFV
'Joe's car did not run.' (p. 81)

(66) Huan **kawyu-ga** 'o huhu'id g gogoks.
John **horse-ALIEN.POSS** 3.AUX.IPFV chase.IPFV DET dogs
'The dogs are/were chasing John's horse.' (p. 82)

In cases when the third-person possessor is lexically expressed (as opposed to being represented by an independent pronoun), however, possession can be expressed via juxtaposition alone, where neither the head (possessed noun) nor the dependent (possessor) is morphologically marked, as discussed in §3.7.3. Examples (58)-(59) above are repeated as (67)-(68) below.

(67) **'Ali je'e** 'at o cipk si'alim.
child mother 3.AUX.PFV fut work.PFV tomorrow.
'The child's mother will work tomorrow.' (Zepeda, 1983, p. 75).

(68) **Mali:ya nawaṣ** 'o 'an wo'o mi:sa da:m.
Mary pocket.knife 3.AUX.IPFV LOC lie.IPFV table on.top.of.
'Mary's pocket knife is lying on top of the table.' (p. 75)

In verb phrases, verbs are marked for objects (and auxiliary verbs are marked for subjects; refer to §3.3). That is, head-marking is employed in verb phrases, illustrated below.

(69) Ceoj 'o 'a:ñi **ñ-ceggia**.
boy 3.AUX.IPFV 1SG **1SG.OBJ-fight**
'The boy is/was fighting me.' (Zepeda, 1983, p. 35)

(70) Hegai 'ali 'o 'a:cim **t-kuḍut**.
That child 3.AUX.IPFV 1PL **1PL.OBJ-bother.IPFV**
'That child is/was bothering us.' (p. 36)

(71) Hegai 'uwĭ 'o 'a:pim '**em-ñu:kud**.
That woman 3.AUX.IPFV 2PL **2PL.OBJ-take.care.of**
'That woman is/was talking care of you (pl.)' (p. 36)

3.9 Summary

This chapter has described the basic appearance of O'odham. First, a brief overview of O'odham phonology has been provided, including a description of its phonemic inventory, stress patterns and syllable structure, including possible word shapes. Next, the morphological typology of O'odham has been best described here as mildly synthetic, allowing some affixation on the verb and other parts of the grammar, such as nouns and postpositions. Regarding participant coding, this chapter has described O'odham as marking subjects on auxiliaries and objects on transitive and ditransitive verbs. Next, it has been illustrated that O'odham verbs indicate aspect most commonly by the presence or absence of word final consonants or vowels and consonants for the imperfective and perfective, respectively. Other patterns, such as suppletion or no change in form are also possible. Future imperfective and perfective constructions are also possible, as well as the progressive aspect.

The most common expressions for various types of nominal and verbal number – namely, reduplication and in some instances gemination in combination with reduplication – were described. This was followed by a discussion on S, A, and P alignment. While person marking on the auxiliary and verb in the form of prefixes falls along a nominative-accusative pattern, number marking indicates ergative-absolutive alignment. When word order is examined in O'odham, sentence word order is variable, determined by discourse factors. Adpositions can come in either pre- or post-positions, and in genitive constructions, the possessor usually precedes the possessed object. Finally, the locus of marking tends to fall on the head in possessive constructions and with regard to object marking.

Chapter Four: The data and how they are utilized

4.0 Introduction

This chapter describes the data that I employ for this usage-based study in order to explore the different areas of grammar in O'odham that are being investigated; namely potential change in progressive aspect and demonstrative constructions. The materials are entirely comprised of natural, oral discourse produced by native speakers of O'odham from three different points in time, as I describe in greater detail in the subsequent sections. The major strength of this type of study is that it allows me to examine authentic examples of language use and real examples of grammatical structure at the time of recording. They allow a comparative, diachronic analysis of language use to be conducted across these different points in time. This type of study is sometimes referred to as a 'real-time' comparison (Labov, 1972; see also Hill, 1983). Additionally, few studies concerning language change in endangered languages carry out this type of analysis, usually due to the paucity of materials that are available for a majority of under-described Native American languages. One goal of this study is to offer new and valuable insights into language change in contexts of severe language shift.

Section 4.1 includes background information on the materials – what they are comprised of, how and when they were collected, who collected them, who the speakers are, and how/why these particular texts were selected. Next, §4.2 provides greater detail on what is known about the speakers. Finally, §4.3 describes how I processed the data in order to produce the corpora which represent the three different periods in time.

Other methodological details, such as coding procedures that are specific to each of the areas of grammar that I examine in this study, are discussed separately. Refer to each of the methodology sections in their respective chapters (§§5.4-5.5 and 6.3-6.4).

4.1 The materials

As stated above, the data that I utilize for this study are entirely comprised of natural, oral discourse. All of the materials are publicly accessible either through libraries, language archives, or on the internet. The data come from roughly three different points in time – the early 1900s, the mid-1900s, and the early 2000s, with speakers' births spanning across at least 100 years from as early as the mid-1800s to as late as the 1960s. I refer to the materials from each of these periods as the **historical** (i.e. the earliest known recordings of native O'odham speakers), **midcentury**, and **modern data**, respectively, throughout this work. A comparative, diachronic analysis of these data is intended to shed light on how potential grammatical changes correlate with the early-shift, middle-shift, and late-shift periods outlined in §2.4. Whereas very few O'odham spoke any English during the early-shift period and were primarily monolingual O'odham speakers, today (during the late-shift period) over half of the population are monolingual English speakers (see timeline summary in Figure 5, §2.4).

4.1.1 Historical data

The historical data are comprised of oral materials that were recorded and, in some instances, produced by Juan Dolores who was a native speaker of O'odham. Although the years that Dolores collected the materials are in most cases not precisely known, Dolores met anthropologist Alfred Kroeber in San Francisco around 1909 while he was employed as a guard at the Museum of Anthropology of the University of California, Berkeley (Kroeber,

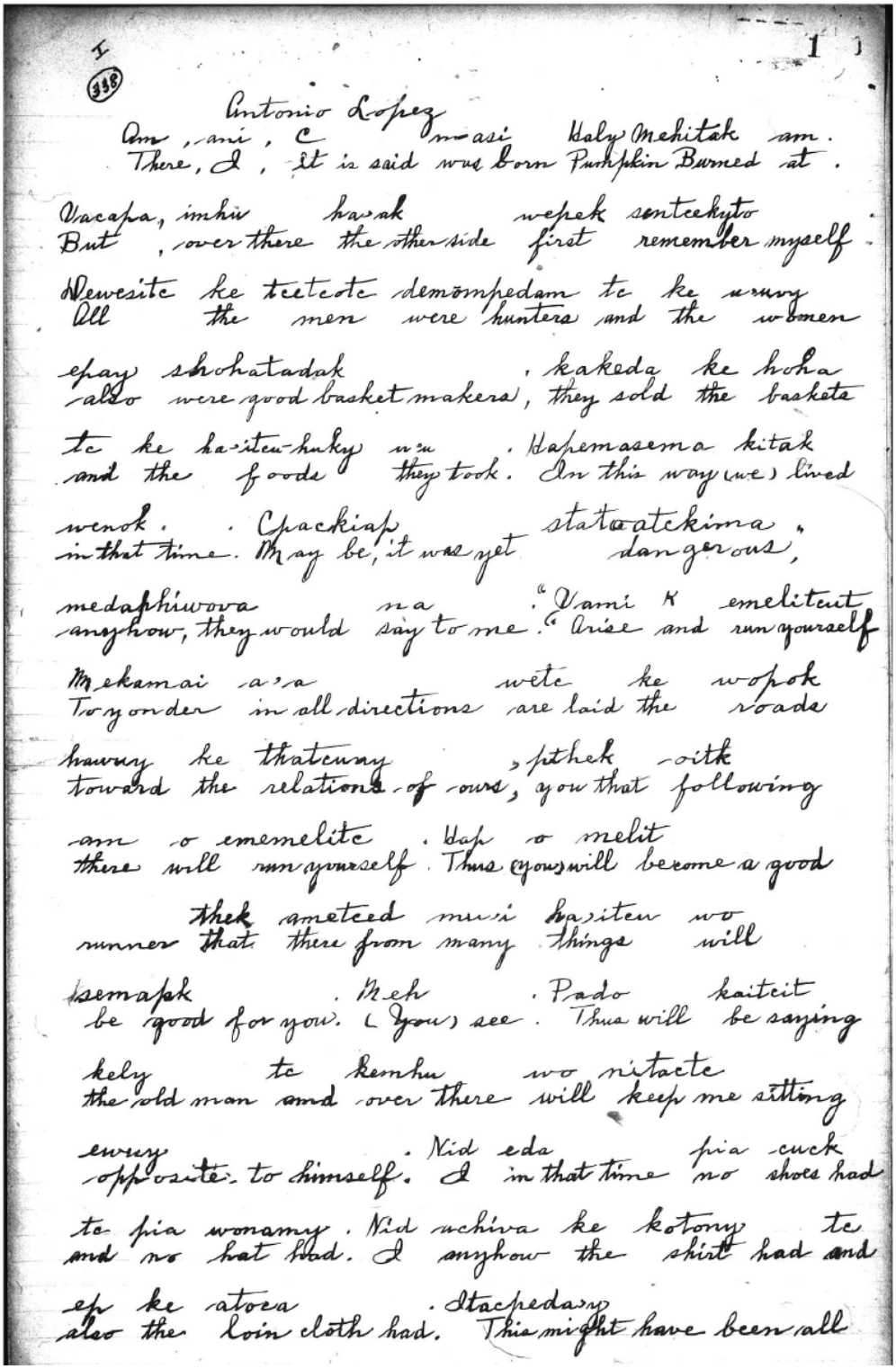
1950, p. 1). Over the course of the next seven or eight years upon meeting Kroeber, Dolores recorded data that would serve as the basis of linguist J. Alden Mason's (1950) grammar of the language. Eventually, Dolores returned to the Tohono O'odham Reservation in Arizona while serving as a Research Fellow in the University of California's Department of Anthropology between the years 1918-1919. During this period, according to Kroeber (1949), Dolores was devoted to collecting O'odham texts from various members of his community. In addition, Dolores "compiled and classified Papago lexicon, and lists of examples of affixes, thus doing much of the grammatical spade work on the language" (p. 96-97). Dolores also published descriptions of various aspects of the O'odham language, for example, *Papago Verb Stems* (Dolores, 1913*)* and *Papago Nominal Stems* (Dolores, 1923*)*.

While a large number of the language records that Dolores collected during the early 1900s were eventually published by Saxton and Saxton (1973) in *Legends and Lore of the Papago and Pima Indians* for a general audience, a handful of them remain in archival formats and have never been fully translated. The latter are the materials that comprise the historical data for this study, with the exception of Dolores's personal memoir (described below). I chose the unpublished materials specifically because various members of the O'odham community have expressed an interest in them being processed (i.e. transcribed and/or typed; see §4.2) and translated into English as well as Spanish for community members who do not speak O'odham but who want access to the culturally salient information contained in them. Many members on both sides of the border recognize that cultural knowledge is being lost along with the language. For the O'odham residing in Sonora, Mexico, in particular, few oral records of this historical nature have ever been translated into Spanish. Former Mexico Lieutenant Governor Jose Garcia laments, "There are

O'odham descendants [in Mexico] that say, ‘We know we’re O'odham, but we don’t have anyone to show us what it is to be O'odham,’…[a]nd we’re trying to revive and restore some of the traditional things that have been forgotten in Mexico" (Kilpatrick, 2014, *Facing Extinction*, section, para. 8). Translations of these O'odham texts into both English and Spanish can play a role in restoring some of that knowledge. Additionally, O'odham educators at the O'odham Culture Teacher Gathering in 2016 expressed an interest in these particular texts as potential resources for culturally relevant school curriculum. These texts are unique for a few different reasons described below.

The first two texts that I selected for processing are two different types of personal stories from the early 1900s, whereas the texts from the same time period later published in Saxton and Saxton (1973) are largely comprised of traditional oral stories, such as the O'odham creation story, numerous animal stories, etc. The first text that I have chosen is, in part, a **personal narrative**, or a telling of a specific event that the speaker was directly involved in. *The Story of War with Mexicans* (Dolores & Kroeber, 1909-1951) is narrated by two speakers. Approximately the first third of the narrative is told by Shawañ 'Uwe ('Smells Like Crow') who was a boy at the time of the events described in the story. The remaining two-thirds of the narrative is told by Dolores because, as he states, he also knew well the events of this story from his own grandfather. The speakers describe one of the many battles between the O'odham and Mexican people during the 1800s. The story is valuable in that it provides information on this battle from the O'odham perspective. The original handwritten version is 28 pages in length (2,534 words). The second text is *Antonio Lopez's Account of Life as a Child on the Desert* (Dolores & Kroeber, 1909-1951). It is a **personal memoir**, in which Antonio Lopez reflects on his childhood during the 1800s, providing a glimpse into

what daily O'odham life looked like before the US government's strict assimilationist policies had completely taken hold of the community. The original handwritten version of this text totals 11 pages (1,035 words). Both of these texts remain in handwritten form; however, scanned versions are available through the Bancroft Library at University of California, Berkeley (UC Berkeley). Dolores includes word glosses for both of the texts, as shown in the sample page in Image 1 on the next page from Antonio Lopez's memoir, but no free translations.



Antonio Lopez

Am , ani , c mi asi Haly Mehitak am .
There, I , it is said was born Pumpkin Burned at .

Vacapa, imhiv havak wepek sentcehyto
But , over there the other side first remember myself .

Newesitc ke teetcotc demompedam tc ke uruvy
All the men were hunters and the women

epay shohatadak , kakeda ke hoha
also were good basket makers , they sold the baskets

tc ke ha-itcu-huky uru . Hapemaseema kitak
and the foods they took . In this way (we) lived

wenok . Cpackiap statu atckima ,
in that time . May be, it was yet dangerous ,

medaphiwova na . " Vami k emelitcut
anyhow, they would say to me . " Arise and run yourself

Mekamai a'a wetc ke wopok
To yonder in all directions are laid the roads

havuy ke thatcuay , pthek oitk
toward the relations of ours , you that following

am o ememelitc . Hap o melit
there will run yourself . Thus (you) will become a good

thek ametced muvi ha-itcu wo
runner that there from many things will

hsemapk . Meh . Pado kaitcit
be good for you . (You) see . Thus will be saying

kely tc kemhu wo nitactc
the old man and over there will keep me sitting

ewuy . Nid eda pia cuck
opposite to himself . I in that time no shoes had

tc pia wonamy . Nid uchiva ke kotony tc
and no hat had . I anyhow the shirt had and

ep ke atoca . Stacpeda-y
also the loin cloth had . This might have been all

Image 1. Sample page from *Antonio Lopez's Account of Life as a Child on the Desert* (Dolores & Kroeber, 1909-1951).

A majority (if not all) of the texts that Dolores collected were first recorded onto wax cylinder and then later transcribed, a large number of which are held at the Phoebe A. Hearst Museum of Anthropology (Phoebe Hearst Museum) and also at the Berkeley campus as part of the *J. Alden Mason and Juan Dolores Collection of Tohono O'odham Sound Recordings*. This presumably was the case with these two texts mentioned above, as well, although the original recordings, to my knowledge, have never been located. Based on the audio recording dates of other materials from both Shawañ Uwe and Antonio Lopez (see California Language Archive, 2018), it is likely that Dolores collected both of these texts between 1918-1919.

The third narrative representing the historical era is one of two versions of a **traditional oral narrative**, or a story that has orally been passed down for generations. It is entitled *Narrative of the Boy Who Became a Warrior* also referred to as *Boy Who Revenges His Father* recorded in 1909 (Dolores & Waterman, 1909). It describes the events leading up to the transformation of a boy into an eagle. The version of the story used for this study is on wax-cylinder at the Phoebe Hearst Museum, from which I obtained a digitized copy. This telling of the story has been heard by few O'odham people, and the fact that an audio recording of it can be heard today adds to its value for the O'odham community. An alternate version of this story which was probably recorded around the same time is available in Saxton and Saxton (1973, pp. 189-206). The museum's catalog lists Juan Dolores as the language consultant. This text is approximately 10 minutes in length (1,202 words).

The final narrative comprising a part of the historical data for this study is another personal memoir, *The Reminiscences of Juan Dolores,* which was published decades after Dolores recorded it by Madeleine Mathiot (Dolores & Mathiot, 1991). According to Mathiot, Dolores recorded the memoir "during the first decades of this century" (p. 233).

Mathiot provides the English translation for the text, and also includes Dolores's original word-for-word literal glosses which also exist in the original hand-written version of the text (also available through the Bancroft Library at UC Berkeley). I included this text in this study because Mathiot conducted some of the leg-work needed in order to make the text more easily analyzable. More specifically, Mathiot conveniently marks morphological breaks in the O'odham. I added all morphological interlinear glosses where relevant. The text was also selected because it is of substantial length. Dolores's original handwritten version is about 55 pages (estimated around 5,500 words).

Information for all of the historical materials is summarized in Table 18 below.

Text title	**Data source**	**Text type**	**Date recorded**	**Speaker(s)**	**Word length**
The Story of War with Mexicans	Dolores & Kroeber (1909-1951)	Personal narrative	1918-1919	Shawañ Uwe ('Smells Like Crow') & Juan Dolores[12]	2,534
Antonio Lopez's Account of Life as a Child on the Desert	Dolores & Kroeber (1909-1951)	Personal memoir	1918-1919	Antonio Lopez	1,035
Narrative of the Boy Who Became a Warrior	Dolores & Waterman (1909)	Traditional oral narrative	1909	Juan Dolores	1,202
The Reminiscences of Juan Dolores	Dolores & Mathiot (1991)	Personal memoir	Early 1900s	Juan Dolores	~5,500
				TOTAL:	~10,271

Table 18. Information for historical texts.

[12] Shawañ 'Uwe tells approximately the first third of the story while Juan Dolores tells the remaining two-thirds.

4.1.2 Midcentury data

All of the midcentury data were collected by Madeleine Mathiot in the late 1950s and early 1960s. The first pool of data is comprised of four different *Coyote Stories* told by Jose Pancho and totals approximately 4,500 words. Digitized versions of each of the stories are publicly available online as PDFs along with recordings of Pancho narrating the stories (Pancho & Mathiot, 1959-1960; URL to online resources provided in References). The PDFs include a) phonetically transcribed versions of each story, b) O'odham versions in Mathiot's orthography, and c) free translations. I also augment the data in some cases with materials included in Mathiot's (n.d.a, n.d.b) two-volume *Tohono O'odham-English dictionary* which are also available on Mathiot's website in the form of two digitized, downloadable PDFs. These are revised versions based on the original 1973 publication, *A dictionary of Papago usage*, and were completed in the early 2000s (a precise date is not provided). Each lexical entry includes several illustrative examples of how a given word is used and also, where relevant, the various morphemes with which a given word can co-occur. An example is provided below (Mathiot, n.d.b, p. 433).

> **'a:g** Vtr [Neutr: def 'a:. Correl: 'a:g-k. Interr: 'a:ga-him]: to say something
> ex: Sbi:tovi mantp o 'a:. It stinks, I might say. — C has masma 'am 'a:gahim 'i:da'a kc uḍ hega'i ñkaidag. And the way (whatever it is) in which I have been saying this is the way I heard it (is my hearing). — Hema 'at 'am o 'a:gk gavul kaidam 'am o 'a: 'i:da'a ha'icu-'a:ga. Someone might tell this story differently.

The examples, according to Mathiot (personal communication, 2018), are largely, if not entirely, taken from various oral narratives. The narratives include several parts of the O'odham creation story as well as the *Coyote Stories*, noted above. Although some examples in the dictionary overlap with what already exists in the *Coyote Stories*, the original versions of the latter have the added strength of providing context for any examples of that are

selected for analysis, which is why I include them in this study, whereas the dictionary presents examples in isolation. The original versions of the texts also include phonetic details in the phonetic transcription that are not available in the dictionary. Given the overlap in material, however, I carefully avoid duplicates of examples appearing both in the actual texts and the dictionary for analysis. Mathiot (n.d.a) states that, for the dictionary, her "principal respondent was Jose Pancho" (p. 3). The two-volume digitized version of the dictionary totals 1,013 pages. Details for the midcentury materials are summarized below.

Text title	**Data source**	**Text type**	**Date recorded**	**Speaker**	**Word length**
Story of Coyote and the Skunk	Pancho & Mathiot (1959)	Traditional oral narrative	1959	Jose Pancho	~1,350
Story of Coyote and the Turkey	Pancho & Mathiot (1959)	Traditional oral narrative	1959	Jose Pancho	~1,200
Coyote and the Quails	Pancho & Mathiot (1960)	Traditional oral narrative	1960	Jose Pancho	~1,000
Coyote, the Cat, the Dog, and the Chicken	Pancho & Mathiot (1960)	Traditional oral narrative	1960	Jose Pancho	~950
				TOTAL:	~4,500

Table 19. Information for midcentury data sources.

******* Also, Mathiot's (n.d.a, n.d.b) *Tohono O'odham-English dictionary* which includes examples from both the *Coyote Stories* and parts of the *Creation Story* (1,013 pages).

While the midcentury data have the limitation of involving a single speaker, they strengthen this study by providing a sample of speech at a midpoint between the early 1900s and early 2000s. In terms of grammaticalization, the added data may provide more detailed

clues as to how the constructions under consideration change over time as English is used in progressively more contexts and by an increasing number of speakers in the community.

4.1.3 Modern language data

Samples of modern language use primarily come from six different public video recordings on the YouTube channel *Tohonotv* (Ortega, 2013-present), as well as one short narrative, *Albino Saguaro* told by Stella Tucker (published in Tucker, Fitzgerald, & Miguel, 2012). All six of the YouTube recordings were made during the early 2000s. In the first two videos, #1 and #2, the speakers provide a single, uninterrupted, **long answer** presumably in response to a question that was asked of them before the recording began. Videos #3, #4, and #5 are in **interview** format in which one speaker poses a question and the interviewee responds with an answer. Each interview includes large segments of continuous speech. The videos range from 5 to 20 minutes in length (or ~610 to ~3,225 words). The final YouTube recording is one short traditional oral narrative, *Hunter Turns into Buzzard,* that is about two minutes in length (~210 words), told by Robert Cruz (provided in Appendix B). The recordings were selected based on how comfortable the speakers were judged to be in using O'odham. The goal was to only include speakers who were fluent, meaning that they spoke O'odham with great ease and fluidity. In contrast, in one or two recordings, it was evident that a particular speaker in a given video recording frequently hesitated, searched for words, back-tracked, etc. Recordings of this nature were eliminated from the study for the reasons described earlier in Chapter 1. The primary person who judged the speakers' fluency was Robert Cruz, who is himself a fluent speaker of O'odham. The final sample of modern language use, *Albino Saguaro,* is described as "an example of spontaneous contemporary storytelling, defined as a person coming up with a story on the spot" (Tucker, Fitzgerald, & Miguel, 2012,

p. 417). The **personal narrative** is told by Stella Tucker in which she talks about finding a saguaro with white fruit instead of the usual deep pinkish-red. The story was recorded in 2002 by Collen Fitzgerald. Phillip Miguel provided an English translation of the narrative. It totals approximately 380 words.

Both male and female speakers were included in the modern data, although available materials of female speakers were limited in number. Male speakers are recorded in videos #1, #2, #3, as well as in *Hunter Turns into Buzzard.* Female speakers are included in #4 and #5, and in *Albino Saguaro.* It should be noted that while the interviewer in videos #4 and #5 was male (speaker #2), the only portions of these recordings that were transcribed and analyzed were produced by the female speakers. Speaker #2 actively conducts interviews for broadcast on *Tohonotv* (Ortega, 2013-present), and as a result, his speech is already represented in videos #2 and #3. His portions of the interviews with the female speakers were instead paraphrased into English for the purpose of having an overall understanding of the content of the interviews.

Information about the collection of modern data is summarized below in Table 20. With the exception of the narratives, *Hunter Turns into Buzzard* and *Albino Saguaro,* the recordings and speakers are referred to by number.

Text # /Title	Data source	Text type	Year recorded	Speaker	Female (F) or Male (M)	Word length
#1	YouTube Video[13]	Long answer	2000	#1	M	~410
#2	YouTube Video	Long answer	2016	#2	M	~1,650
#3	YouTube Video	Interview	2011	#3 & #2	M, M	~3,225
#4	YouTube Video	Interview	2010	#4 (& #2)	F, M	~1,300
#5	YouTube Video	Interview	2008	#5 (& #2)	F, M	~610
Hunter Turns into Buzzard	YouTube Video	Traditional oral narrative	2012	Robert Cruz	M	~210
Albino Saguaro	Tucker, Fitzgerald, Miguel (2012)	Personal narrative	2002	Stella Tucker	F	~380
					TOTAL:	~7,785

Table 20. Information for modern data sources.

4.1.4 Limitations and strengths of the data

My primary stipulation for the data for this study was that they be comprised entirely of naturally produced speech from different points in time. As pointed out earlier, the data that I located exist in language archives, in university libraries, or online as in the case of the materials Jose Pancho and Madeleine Mathiot produced and collected, respectively, as well as the YouTube video recordings. One limitation of the data is that they are comprised of different types of discourse rather than one uniform style from one period to the next. The archival materials may have been recorded after speakers were able to practice to some extent, thus resulting in more carefully crafted and formalized productions of oral speech.

[13] URL to YouTube channel is included in references (see Ortega, 2013-present.). More specific information about these recordings can be obtained from the author.

This may have also been the case with the production of the midcentury materials. In contrast, the YouTube video recordings are comprised of entirely impromptu speech styles. Furthermore, the modern data contrast most starkly to the other two data periods in that they consist of interviews and the "long answer" format, which lend themselves to more colloquial uses of language. The different discourse types could potentially influence certain aspects of this study and account for some of the patterns that emerge in the results. Where relevant, I will address these concerns in the studies presented in Chapters 5 and 6.

The most important advantage that the data offer, however, is that I am able to carry out a comparative, diachronic analysis of the structures based entirely on authentic speech rather than on contrived, elicited data responses. The naturally spoken data make it possible to capture details about how the speakers authentically use the structures in each of the time periods and how the structures have possibly changed – details that might otherwise be elusive. Overall, usage-based studies of O'odham are limited (a couple of examples are Payne, 1987, 1992). This study will contribute to and enrich our understanding of how progressive constructions and demonstratives are utilized in natural O'odham discourse.

4.2 The speakers

Information about the speakers of the historical, midcentury, and modern materials combined is quite variable. In some cases, very little is known but some information can be ascertained. In others, much more information is readily accessible. The goal here is to provide at least an estimation of the speakers' ages, language background, and dialects where exact details are not available.

4.2.1 Historical speakers

Information about the narrators of the historical narratives is limited with the exception of Juan Dolores. Dolores was born around 1880 probably in or very near the border of Sonora, Mexico (Kroeber, 1949; Dolores & Mathiot, 1991). His father eventually moved the family to the reservation near Tucson, Arizona, when Dolores was a child, at which point Dolores attended a Presbyterian boarding school. Dolores grew up among the O'odham community during a time when O'odham was the dominant means of communication. He spoke the Kokololodi (sometimes also written as Kolo:di) subdialect of Tohono O'odham. Dolores describes in his memoir that neither he nor any of his classmates could understand or speak any English upon entering school (Dolores & Mathiot, 1991, pp. 262; 307). During his years in various government schools, he eventually did learn to speak English as a second language, which he had reason to use intermittently during his adulthood for various off-reservation activities, including several different jobs away from home. Dolores, according to Kroeber (1949), also spoke Spanish.

While little has been written about the other narrators, Shawañ Uwe and Antonio Lopez, some information about them can be deduced. *The Story of War with Mexican* is told in part by Shawañ Uwe as well as Dolores, as noted earlier. During the story, Dolores states that he knew the story well from his own grandfather. This may suggest that Shawañ Uwe was from the same generation as Dolores's grandfather. Based on this information, a conservative estimate of Shawañ Uwe's year birth is around 1850, but possibly earlier. Shawañ Uwe also spoke the Kokololodi subdialect of Tohono O'odham. No information is available regarding whether or not he spoke any English or Spanish. Based on language

trends of the time during which Shawañ grew up (see §§2.3and 2.4), it is assumed that he was primarily if not entirely a monolingual O'odham speaker.

Similarly, little is known about Antonio Lopez. In Lopez's narrative, which Dolores recorded around 1918, he reminisces about what life was like when he was a child. Lopez also discusses being sent to a boarding school in Sacaton, Arizona, the first school in the area which was opened in 1871 (Hagan, 1959, p. 121), suggesting that he probably was younger than Shawañ Uwe. Lopez may have been born sometime around 1860. Although he was sent to school, Lopez reports that he feared it and returned home to live with various members of his family before ever attending the school. Based on this information, it is assumed that Lopez received little exposure to English. It is not known if Lopez spoke any Spanish. It is very likely, again based on the time he grew up, that he was predominantly a speaker of O'odham. Lopez, like Dolores and Shawañ Uwe spoke the Kokololodi subdialect of Tohono O'odham. Table 21 provides a summary for the historical speakers' language backgrounds.

Speaker's name	**Estimated date of birth**	**Tohono O'odham subdialect**	**Language(s) spoken**
Juan Dolores	1880	Kokololodi	O'odham, English, Spanish
Shawañ Uwe	~1850s (?)	Kokololodi	O'odham (others not known)
Antonio Lopez	~1860s (?)	Kokololodi	O'odham (others not know)

Table 21. Speaker background for historical data.

4.2.2 Midcentury speakers

According to Mathiot (personal communication, 2018), Jose Pancho was at least 50 when they worked together in the late 1950s and early 1960s. Based on this information, his year of birth is estimated to be sometime around the 1910s. Pancho was bilingual in O'odham and

English. It is not known if he spoke any other languages. Pancho spoke the Totoguañ subdialect.

Speaker's name	Estimated date of birth	Tohono O'odham subdialect	Language(s) spoken
Jose Pancho	~1910s (?)	Totoguañ	O'odham, English, (others not know)

Table 22. Speaker background for midcentury data.

4.2.3 Modern speakers

The modern speakers are estimated to have been born roughly during the third quarter of the 20th century approximately between the years of 1950 and 1975. This is based in part on personal knowledge about some of the speakers, on physical appearance of speakers in the YouTube videos or, in the case of Stella Tucker, based on what has been written about her in Tucker, Fitzgerald, and Miguel (2012). All of the speakers are known by Robert Cruz, the primary consultant for this study, to be fluent bilingual speakers of both O'odham and English. In some cases, the speakers' knowledge of other languages is not known. The subdialects of Tohono O'odham that are represented in the videos include the Totoguañ, Huhu'ula, and Kokololodi dialects. Stella Tucker, the narrator of *Albino Cactus*, was born around 1948 (Duarte, 2019) and was fluent in both O'odham and English. Whether or not Tucker spoke other languages is not known. Details for the modern speakers are summarized in Table 23.

Speaker	Estimated date of birth	Tohono O'odham subdialect	Language(s) spoken
#1	~1970-75	Totoguañ/ Huhu'ula	O'odham, English (others not know)
#2	~1951-2	Kokololodi	O'odham, English, some Spanish
#3	~1950	Kokololodi	O'odham, English, some Spanish
#4	~1970-1975	Totoguañ	O'odham, English (others not know)
#5	~1955-1960	Totoguañ	O'odham, English (others not know)
Robert Cruz	1949	Kokololodi	O'odham, English, some Spanish
Stella Tucker	~1948	Totoguañ	O'odham, English (others not known)

Table 23. Speaker background for modern data.

4.3 Data processing procedures

A majority of the data required various amounts of processing in order for them to be analyzable. This section will first address the specific processing details that I implemented for the historical data, which required the most extensive attention and time investment, as will be outlined below. This is followed by the processing procedures for the midcentury and modern data.

4.3.1 Historical data

The historical texts, as noted in §4.1.1, are comprised of one digitized version of a wax-cylinder recording, two handwritten narratives, and Juan Dolores's memoir. The different processing tasks that I carried out for the historical materials are briefly summarized below and then described in greater detail in what follows:

a. Transcribing wax-cylinder recording
b. Converting Dolores-Kroeber orthography into Alvarez-Hale orthography (hand-written texts only)
c. Typing materials and entering into FLEx program
d. Adding Dolores's word-for-word glosses into FLEx (hand-written texts only)
e. Discerning and adding morphological interlinear glosses
f. Adding English and Spanish free translations

The first steps for the wax-cylinder recording, *Narrative of the Boy Who Became a Warrior,* required transcription, as noted in the task-list above. I did this in collaboration with various members of the O'odham *Ne'ok* Revitalization Project,[14] especially Robert Cruz and Phillip Miguel, both native speakers of O'odham. For the hand-written texts, these needed to first be converted into a different orthography. In their original form, they are in an antiquated system referred to here as the Dolores-Kroeber orthography. The Tohono O'odham Nation's official writing system today is the Alvarez-Hale orthography as noted earlier, and given that one of the goals of this project is to make these historical texts more accessible to the community, it was imperative to carry out this step. In order to convert the texts, I first determined orthographic correspondences between the Dolores-Kroeber and Alvarez-Hale systems (provided in Appendix A), and then I implemented the conversion of the text into the Alvarez-Hale orthography. I also typed both the wax-cylinder recording and hand-written texts. After the first drafts of the texts were completed, Robert Cruz double-checked the O'odham for any errors. Final decisions, however, were made by me; I am solely responsible for any remaining errors.

[14] Members who worked on this text in some compacity in addition to myself are: (O'odham speakers) Robert Cruz, Rafael García, Phillip Miguel; (Other members) Melissa Axelrod, Jacob Franco Hernández, Serena Kavanaugh, Iphigenia Kerfoot, Laura Ruth-Hirrel, and Andrés Sabogal.

After these first steps with the wax-cylinder and hand-written texts were completed, I copied them into the FieldWorks Language Explorer (FLEx) freeware program (Summer Institute of Linguistics, 2017). The program expedites the glossing of language corpora and also includes search capabilities for language structures. FLEx is described in greater detail in §4.3.4. Once I entered the transcribed texts into FLEx, the next step with the hand-written texts in particular was to add Dolores's original word-for-word glosses into FLEx (see Image 1). The word glosses are rich with detailed information and ultimately provided guidance for the next steps. After Dolores's glosses were added into FLEx, I parsed the words into their component parts, morphologically glossed them, and added the English and Spanish translations. For the wax-cylinder recording, *Narrative of the Boy Who Became a Warrior,* I deciphered glosses and English and Spanish translations in collaboration with members of the *O'odham Ne'ok* Revitalization Project. For the hand-written texts, *The Story of War with Mexicans* and *Antonio Lopez's Account of Life as a Child on the Desert*, I determined English translations and morphological glosses through many long discussions with Robert Cruz and also Phillip Miguel, as well as Jacob Franco Hernández. I also frequently consulted grammatical descriptions of O'odham (Hale, 1959; Mason, 1950; Mathiot, 1973; Saxton, 1982; Zepeda, 1983, 1984) as well as O'odham dictionaries (Mathiot n.d.a, n.d.b; Saxton et al., 1983) for guidance particularly with the morphological glossing. Jacob Franco Hernández was largely responsible for providing the Spanish translations of all of the historical materials, excluding Juan Dolores's personal memoir (Dolores & Mathiot, 1991). All final changes to either the English translations or morphological glosses included in this work (including errors) are entirely my own.

Finally, Juan Dolores's memoir (Dolores & Mathiot, 1991) required limited processing. Mathiot conveniently indicates morpheme breaks and also provides an English translation of the O'odham text, as noted earlier. The publication does not, however, include morphological interlinear glosses. I deciphered and added morphological glosses which are crucial for a detailed understanding and analysis of the language's morphology. Additionally, Mathiot presents the text in the original Dolores-Kroeber orthography as well as her own orthography, both of which differ in many ways from the Alvarez-Hale system (correspondences provided in Appendix A). As a result, it was necessarily for me to also convert all text segments included in this study from Mathiot's orthography into the Alvarez-Hale system. However, it should be pointed out that I did not need to apply these additions and alterations to the entire narrative. For each of the areas of grammar that I analyze in this study, I first identified tokens of the specific constructions (see §§5.4 and 6.3 for more details) and then extracted the specific examples from the text. Once this was completed, I converted the examples into the Alvarez-Hale orthography and entered them into FLEx for morphological glossing.

4.3.2 Midcentury data

The midcentury materials, as described in §4.1.2, were originally collected by Madeleine Mathiot. The procedures that I applied to these materials are as follows:

a. Extracting relevant examples and entering into FLEx
b. Converting Mathiot orthography into Alvarez-Hale orthography
c. Adding morphological interlinear glosses

Mathiot provides all of the materials in the Mathiot orthography and also includes English translations. The procedures that I described for the Juan Dolores memoir in the previous

section were similarly followed for the *Coyote Stories* (Pancho & Mathiot, 1959, 1960) and examples from the online version of Mathiot's (n.d.a, n.d.b) two-volume dictionary. I identified, extracted, and converted relevant grammatical constructions from the Mathiot orthography into the Alvarez-Hale orthography, copied them into FLEx, and then morphologically glossed them.

4.3.3 Modern data

The modern data consist primarily of video footage available on the YouTube channel *Tohonotv* (Ortega, 2013-present) as well as Stella Tucker's *Albino Saguaro* (Tucker, Miguel, & Fitzgerald, 2012), all of which I describe in greater detail in §4.1.3. I elaborate on each of the steps listed below in what follows.

a. Extracting audio data and entering into Transcriber
b. Transcribing and typing recordings
c. Exporting typed transcriptions into FLEx
d. Adding morphological interlinear glosses
e. Adding English free translations

After I selected the videos for inclusion in the modern language corpus, the next steps were to extract the audio from the videos, convert the audio into WAV audio files, and then upload the files into Transcriber (Boudahmane, Manta, Antoine, Galliano, & Barras, n.d.), a free software program that is designed specifically for transcribing and segmenting speech. Once a recording was in Transcriber, I carefully transcribed each text in O'odham in close collaboration with Robert Cruz. After a text was completed, Mr. Cruz listened to the recordings again on his own time and double-checked the transcriptions for any mistakes in the typed text (again, any errors are my own). The next step was to enter the O'odham texts

into FLEx and, as with all other materials, add English translations and interlinear morphological glosses.

The only processing steps listed above that were needed for *Albino Saguaro* were copying the text into FLEx and identifying and labeling morphemes.

4.3.4 FieldWorks (FLEx) software

All of the data in some capacity were entered into FLEx (Summer Institute of Linguistics, 2017) either as entire narratives or, in the case of the midcentury data, as selected examples illustrating specific grammatical structures. The main strength of using this software is that is ensures that large amounts of text are glossed and analyzed consistently, as will be highlighted below. This section also describes other details of how I employed this tool to create the corpora I use for this study.

The first step in FLEx, as noted earlier, is to copy typed materials into the program. Once this is done, the text appears in the *Word* line as shown in the sample snapshot of one line from *Antonio Lopez's Account of Life as a Child on the Desert* in Image 2 below. The next step is to begin creating the interlinear text. In the case of the Dolores collection of texts, my first step was to add Dolores's original literal word glosses as illustrated in the fourth line or the *Word Gloss* line in Image 2 below. Next, I added morpheme breaks in the *Morphemes* line and glossed for each morpheme in the *Lex. Gloss* line. Once a specific morpheme is glossed, the program then automatically stores this information in its lexicon. The program will then automatically recognize and gloss any morpheme that has the identical form throughout the text, expediting the glossing process and, in the case of morpheme labels, making certain that they are consistent across all of the materials, which is crucial for studies involving morphological details. The program also has built-in features which allow one to

assign multiple possible meaning labels for a single form that can then be selected from (e.g. *t-* is glossed below as '1pl.poss' or first-person, plural, possessive, but another possibility is for it to be assigned the label '1pl.refl' for first-person, plural, reflexive). FLEx also enables the designation of word categories as I have done and as shown in the *Word Cat.* line below. Finally, I input English free translations and, in the case of the historical texts, Spanish free translations into the bottommost lines under the label *Free*.

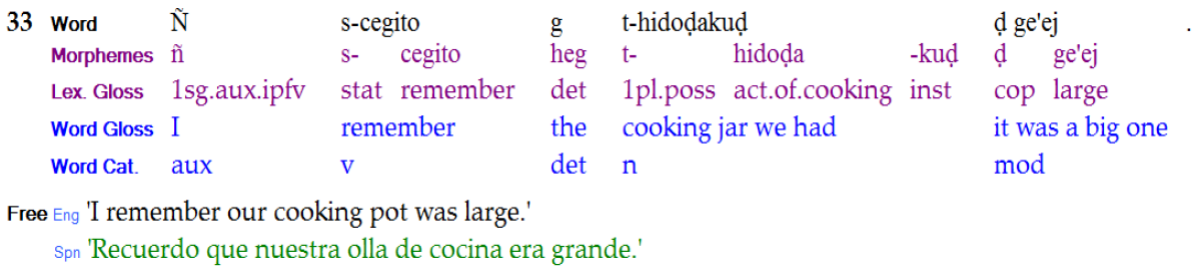


Image 2. Sample line in FieldWorks Language Explorer (FLEx).

For all of the data with the exception of the Mathiot's materials (i.e. Dolores & Mathiot, 1991; Pancho & Mathiot, 1959, 1960; Mathiot, n.d.a, n.d.b), I utilized FLEx for these purposes. For Mathiot's materials, I first identified tokens of each of the specific constructions under analysis and then entered these into FLEx for morphological parsing.

FLEx additionally includes search capabilities for specific morphemes (and any other feature such as specific words, word glosses, word categories, etc.) in the case that one wants to analyze specific constructions as illustrated in Image 3 below. For example, if all tokens of the progressive marker need to be identified and extracted, the tag *prog* is entered into the search box under *Specify Concordance Criteria*. Once the search has been entered, all examples of *prog* throughout the texts in the data base are shown under *Concordance Results*. A specific token of *prog* can then be selected (highlighted in light purple below), and

FLEx then brings you to that example, in this case the verbal construction *wipia-him* 'hunting' in line 34 of the given text, which is illustrated in the adjacent box to the right.

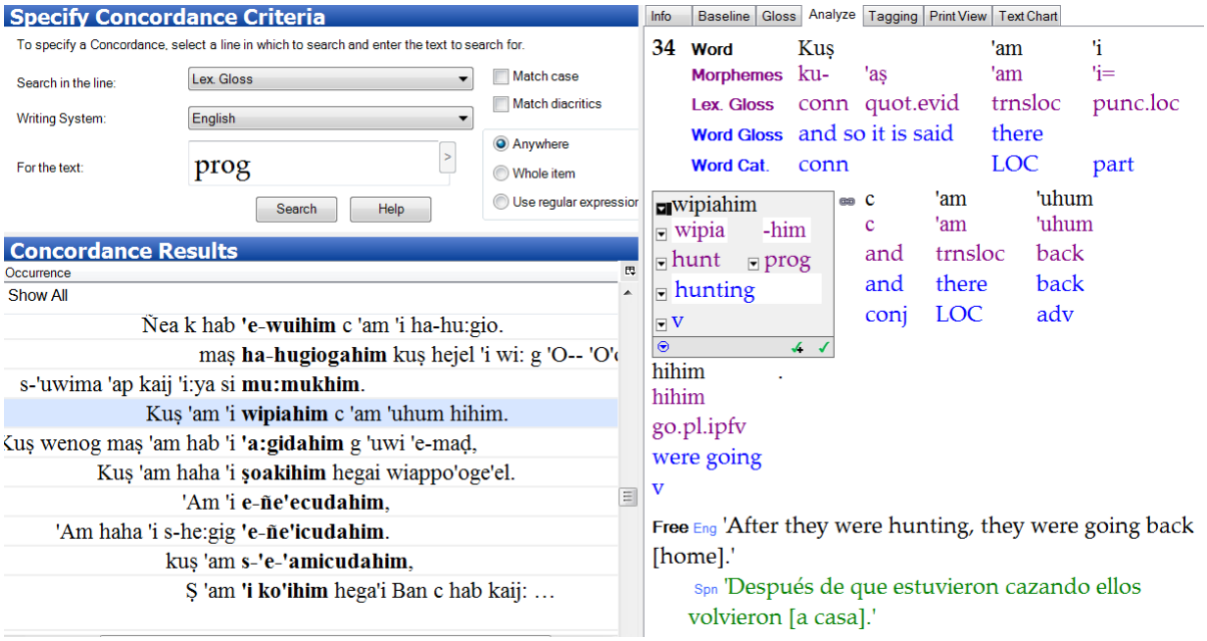


Image 3. FLEx search example.

This search method was the primary means that I used in many cases to identify the specific structures for analysis. I discuss any additional procedures that were needed separately in the respective methodology sections for each grammatical points that I investigate (§§5.4 and 6.3).

4.4 Summary

In this chapter I have described in detail the data that make up the corpora that I utilize for this usage-based study of language change in O'odham. The materials come from three different periods in time – the early 1900s, the mid 1900's and the early 2000s. They consist of personal memoirs, traditional narratives, interviews, etc. The historical data in particular have the added value of documenting historical events and ways of life from the O'odham perspective around the turn of the century. This chapter has also provided background

information about the speakers who produced the texts to the extent possible. Finally, this chapter has outlined the extensive processing that the texts required in order to make them usable for analysis, including transcription, orthographic conversions, morphological glossing, translations, entering data into the FLEx language processing program, etc. This work is in part the result of collaborations between myself, various members of the O'odham community, and/or members of the O'odham *Ñe'ok* Revitalization Project. The processing of the original, difficult-to-use formats of some of the historical materials in particular and the addition of both English and Spanish translations will increase the O'odham community's access to the culturally rich resources on both sides of the US-Mexico border.

Chapter Five: The progressive aspect

5.0 Introduction

O'odham utilizes a rich aspectual system, as described in §3.4, which includes the expression of the progressive aspect. The progressive marker in O'odham in its most basic form is identified as *-him*, on occasion appearing in the reduced form, *-im*. Its presence is described as extending some action in time (Zepeda, 1984, p. 103-104), and it has been suggested, based on synchronic evidence, that the motion verb *him* 'go, come, walk' is the original source for the progressive marker (Mason, 1950; Saxton, 1982; Fitzgerald, 2004; see also Langacker, 1977 on Uto-Aztecan). Various elements of shared, polysemous meanings between the motion verb and the suffix as well as their resemblance in form suggest that this is probably the case.

The primary goal of this usage-based study is to determine how the progressive may have changed in use over the three periods represented by the O'odham texts, which correlate with the early, middle, and late shift periods outlined in §2.4. Based on these findings, I will determine how the developments align with or differ from the hypothesized expected trends of change proposed primarily by Bybee et al. (1994). If they differ, to what extent might language shift be a contributing factor? Given the language context, it must also be kept in mind that language contact could be one explanation for some types of unexpected changes. Such cases would need to be considered with careful scrutiny.

In this chapter, I begin in §5.1 with an overview of how the progressive aspect is defined and the cross-linguistic grammaticalization pathways that have been proposed for the progressive, primarily in Bybee et al. (1994). This is followed by a more detailed summary in §5.2 of previous scholars' descriptions of the progressive marker in O'odham and in Uto-

Aztecan as a whole (which I briefly introduced in §3.4) – its meanings and function, its phonological shape and its hypothesized historical source. In §5.3, I provide an overview and more in depth study of the phonological shape of the progressive marker, and in §§5.4 and 5.5, I discuss data and token selection and the coding procedures that I implement in order to analyze the progressive marker in O'odham. I also include a number of examples in §5.5 based on my own descriptive analysis of the O'odham data in order to justify and illustrate the coding. Section 5.6 includes the quantitative results for the comparative analysis of the data across the three periods. Finally, in §5.7, I briefly discuss and interpret the findings.

5.1 Progressive aspect in grammaticalization theory

Bybee et al. (1994) state that the progressive aspect conveys an action that is "ongoing at reference time" (p. 126). A consequence of this meaning is that progressive constructions usually co-occur with dynamic predicates that depict active "situations" (p. 55) where there is a detectable "change or movement" (Bybee, 1985, p. 28). This is in contrast to stative predicates which represent contexts where no change can be detected over time (Croft, 2012, p. 34) or in which the "unchanging situation … will continue unless something happens to change it" (Bybee et al., 1994, p. 55). As a result, stative predicates do not usually occur with the progressive. Bybee et al. additionally point out that progressives probably also co-occur in their earliest stages with agentive subjects that are actively engaged in the event. Put another way, the dynamic predicates entail a subject with a great degree of 'control' over the verbal event (see §5.5.1 for more on 'control').

Based on Bybee et al.'s (1994) database, their cross-linguistic findings suggest that one of a few possible sources for progressives are basic movement verbs such as 'come' and 'go' (pp. 132-133), which is what the previous literature has proposed for the progressive

marker in O'odham, as briefly noted in §5.0 (also see §5.2.2). The progressive marker additionally co-occurs with main verbs that "describe activities that have some characteristic and overt location, rather than with predicates describing internal and non-observable states" (p. 133). In other words, if an activity is overt and observable, then it must also occur somewhere in space as opposed to non-observable or internal states. From an evolutionary perspective, Bybee et al. (1994) hypothesize that progressive constructions initially contain all of the following combined pieces of meaning:

(1) (a) An agent
(b) is located spatially
(c) in the midst of
(d) an activity
(e) at reference time (Bybee et al., 1994, p. 136).

Over time, the original meaning of the construction can gradually undergo erosion, and as the earlier semantic elements weaken and broaden, progressive constructions can come to be used in a wider range of contexts. For example, they can begin to occur with predicates where the subjects are experiencers of an event. This expansion likely occurs "at the same time as the notion of activity expands to include developing states" (p. 136) such as 'becoming' or 'getting tall/wise/angry, developing a headache/an understanding,' etc. The meaning that remains when this expansion takes place is a temporal one which expresses "an ongoing action or state of affairs" (p. 136) over a period of time. At this stage, the locative component has been bleached. A person who is *getting tall* or *wise,* for example, does so over time, but not at a specific location. It is this loss, Bybee et al. (1994) argue, that allows the construction to expand: "[If] a construction means 'to be in the midst of doing,' the only change that would give it more contexts of use would in fact be the erosion and loss of the

strict locative meaning, which would then allow the expression of temporal meaning in a wider range of contexts" (p. 137).

To summarize, progressive constructions are hypothesized in some cases to evolve from source verbs expressing movement such as 'come' and 'go,' and their original meaning likely depicts a subject that is actively carrying out or moving around doing some ongoing activity in some location and at some juncture in time. They begin with dynamic predicates that describe active, changing situations and entail agentive subjects with control over the event, but as the meaning of the construction evolves and generalizes, it can potentially expand to include predicates with non-agentive subjects and expressions of developing or ongoing states, and the locative meaning can eventually be lost.

5.2 Descriptions of the progressive in O'odham

This section is devoted to providing a review of previous O'odham scholars' observations regarding the suffix *-him.* It addresses what they note about a) its grammatical functions as well as their (sometimes) conflicting interpretations of its possible variations in form (§5.2.1), and b) the suffix's hypothesized historical source (§5.2.2). All examples presented in this section have been converted into the Alvarez-Hale orthography by me regardless of the original authors' orthographic conventions. Morphological glosses are also my own unless otherwise stated.

5.2.1 Grammatical functions of *-him* and some notes on form

The suffix *-him* is generally described as expressing the progressive aspect by O'odham language scholars such as Mason (1950), Hale (1959), and Saxton (1982). Mathiot (1973) and Zepeda (1984) similarly describe the suffix in manners that are in line with a progressive meaning. Regarding form, all of the authors note in some capacity that the suffix can be

preceded by the vowel /a/ in some cases; however, they differ in their interpretations of the presence of this vowel. Mathiot (1973) additionally points out that the suffix can also sometimes be preceded by /i/. The authors' discussions regarding meaning and form are summarized below.

Mason (1950) provides one of the earliest descriptions of O'odham which is based primarily, if not entirely, on data prepared by native O'odham speaker and linguist Juan Dolores (e.g. 1913, 1923), including O'odham texts. Based on Dolores's materials, Mason (1950) assigns the labels "progressive" as well as "continuative" (p. 52) to the suffix forms *-him* and *-im*, the former label of which is employed here. Mason describes the suffix as adding "a continuative participial sense to the main verb…" (p. 53), meaning that it signals that the action expressed by the verb is marked as ongoing. A few of Mason's examples are provided below in (2). The author also states that the suffix can sometimes take the allomorphic form *-ahim* shown in example (2)e.

(2) a. *bei-**him***
take-**PROG**
'taking up'

b. *kaij-**im***
say-PROG
'saying'

c. *ñe'i-**him***
sing-**PROG**
'singing'

d. *ñeoki-**him***
talk-**PROG**
'had been talking'

e. *cikpen-**ahim***
work-**PROG**
'had been working' (Mason, 1950, p. 53; modified examples)

Mason (1950) – and others, as it will be seen – also observes that the verbal suffix can give "the sense of motion simultaneous with the action of the main verb" (pp. 52-53), suggesting a possible meaning of the suffix that has since been described in cross-linguistic studies as expressing "associated motion" (Guillaume, 2016). I interpret this meaning to be a likely precursor to the progressive meaning of the suffix. I will revisit this topic in greater detail in §5.2.2.

Next, Hale (1959) simply refers to *-him* as well as *-ahim* (written by Hale as *-himɨ* and *-ahimɨ*) as marking the progressive. Hale describes the presence of /a/ on purely phonological grounds (see Hale, 1959, pp. 60-61). His description of the suffix does not extend beyond this brief explanation.

Saxton (1982) specifically labels the form *-ahim,* including the /a/, as the past imperfective (p. 125), and in Saxton, Saxton, and Enos (1983), *-ahim* is described as reflecting "an action in progress in the past" (p. 2). Separately, Saxton (1982) glosses the form *-him,* without the presence of /a/, as a marker which expresses the meaning "to go along V[erb]ing" (p. 161) and similarly in Saxton et al. (1983) *-him* is glossed as "go along (doing something)" (p. 22) where the suffix adds a motion meaning to the verb to which it attaches. A couple of examples with Saxton's original glosses illustrating this meaning are given below. In these cases, the author glosses the preceding /a/ as a separate gerundive suffix, *-a,* although it is not clear what criteria he uses to make this decision based on the examples he provides, and this interpretation of /a/ has not been addressed elsewhere.

(3) *'a:g-a-**him***
say-GER-**PROG**
'to go along saying'

(4) *ñeid-a-**him***
see-GER-**PROG**
'go along seeing' (Saxton, 1982, p. 216)

Next, Zepeda (1984) argues that events are "augmented by the presence of *-him*" (pp. 103-104) and the suffix extends the action expressed by the verb. She often translates the form *-him* as "going along V-ing," associating the meaning of the suffix with motion as Mason (1950), Saxton (1982) and Saxton et al. (1983) do. Examples from Zepeda are shown below.

(5) *waila-**him***
dance-**PROG**
'going along dancing'

(6) *ceggia-**him***
fight-**PROG**
'going along fighting'

(7) *ne'e-**him***
sing-**PROG**
'going along singing' (Zepeda, 1984, p. 101; modified examples)

Zepeda in some cases also records the form *-im*.

(8)	'Ab	'o	him	hegai	c	ñeok-**im**.
	LOC	3.AUX.IPFV	walk.IPFV	3SG	and	talk- **PROG**
	there	aux	walk	him	and	talking:along[15]

'He is walking along talking.' (Zepeda, 1984, p. 104; modified example)

The author also notes the possible presence of /a/ before the suffix (i.e. *-ahim*) arguing that it is purely epenthetic in contrast with Saxton (1982) and Saxton et al. (1983), for example, who sometimes assign different yet related meanings to the forms *-ahim* and -

[15] Third line and forth lines reflect Zepeda's original glosses and translations, respectively.

him. Zepeda (1984) points out that "the suffix -him … generally requires that it be preceded by a vowel" (p. (p. 101) and verbs that are consonant-final usually require "an epenthetic vowel between the word [i.e. the verb stem] and the suffix" (p. 101). Some examples from Zepeda (1984) illustrating the epenthetic /a/ are:

(9)	*huhu'id* [16]	'chase'	→	*huhu'id**a**him*	'going along chasing'[17]
	hukṣan	'scratch'	→	*huksan**a**him*	'going along scratching,'
	keihin	'dance'	→	*keihin**a**him*	'going along dancing'
	ṣamud	'herd'	→	*ṣa:mud**a**him*	'going along herding something'
	piast	'make music'	→	*piast**a**him*	'going along making music'
	(Zepeda, 1984, p. 102)				

Finally, Mathiot (1973) describes *-him* as follows: it "indicates that the action – or condition – can be resumed after an interruption…" (p. 57). Mathiot refers to this as the 'interruptive aspect.' During a personal correspondence with Mathiot, she additionally stated that the addition of *-him* means that the action expressed by the predicate "requires time to manifest" (personal communication, August 1, 2018), but the action could be *temporarily halted* or *interrupted* and then *resumed* once again. Mathiot's choice in terminology may be the result of a desire to make a direct contrast between the 'interruptive aspect' and what the author refers to as the 'completive aspect,' which "indicates that the action is *definitively terminated*" (p. 57; my emphasis). Two of Mathiot's contrasting examples of the interruptive and completive aspects point in this direction.

(10) Cikpan-a-**him** 'o.
work-EPEN-**PROG** 3.AUX.IPFV
'He has been working/he was working (i.e. he is not working now but he may be working again)' (Mathiot, 1973, p. 58; original translation & parenthetical information).

[16] All examples in this discussion are in their imperfective forms unless otherwise noted.
[17] Zepeda's (1982) original glosses.

(11) Nt o ku:p-**oho** g tianna k o m-'oi
1SG.AUX.PFV FUT close-**COMPL** ART store CONJ FUT 2SG.OBJ-accompany.PFV
'I'll close the store and then I'll go with you (I'll go with you after I have closed the store)' (p. 58; original translation & parenthetical information)

In (10), a person has been or was *cikpanahim* 'working' over the course of some period of time in the past, as reflected in the translation, but may potentially resume working. Mathiot contrasts this with the completive aspect example in (11) which describes a scenario in which a person will definitively *ku:poho* or 'close' the store. My interpretation is that Mathiot's 'interruptive aspect' does, in fact, also depict ongoing, progressive actions as is illustrated by example (10) above.

Mathiot (n.d., n.d.a) also includes numerous examples of the 'interruptive' marker like the following in her two-volume dictionary:

(12) 'In has ñ-hu'uid-a-**him.**
LOC INDEF.PART 1SG.OBJ-pursue-EPEN-**PROG**
'He is trailing me everywhere.' (Mathiot, n.d., p. 309)

(13) B 'o kui-**him** g cukuḍ.
LOC 3.AUX.IPFV hoot-PROG ART owl
'An owl is coming this way, hooting.' (p. 429)

Example (12) above depicts a person who is *hu'uidahim* 'pursuing' or 'trailing' somebody over time, and (13) describes an owl that is *kuihim* 'hooting' for some undefined period of time. While Mathiot labels these and numerous additional examples as being in the interruptive aspect, they do, in fact, also convey a progressive meaning based on the author's own translation of the examples, and as others have attested for this suffix elsewhere.

Regarding form, Mathiot, like Zepeda, interprets /a/ as epenthetic: "Whenever the verb theme ends on a consonant when it occurs without an inflectional suffix, an epenthetic vowel may be added" before *-him*, as well as a few other verbal suffixes (p. 66). Mathiot

(1973) is the sole person to additionally note the presence of /i/ that optionally appears preceding *-him*, describing it in this context as a "morphemic component" (p. 67). Some examples from various sources of where this vowel can potentially appear are:

(14)	*ṣoak*	'cry'	→	*ṣoak(i)him*	'crying'
	kaij	'say'	→	*kaij(i)him*	'saying,'
	'oimmeḍ	'wander'	→	*'oimmel(i)him*	'wandering around'
	huḍuñ	'go down'	→	*huḍuñ(i)him*	'going down'
	huḍḍawag	'pester'	→	*huḍḍawag(i)him*	'pestering'

Mathiot refers to this as a morphemic component rather than an epenthetic vowel because of its phonological effect on specific (but not all) verb-final consonants that precede the suffix. For example, <ḍ> becomes <l> (/d/ → /l̠/) after the suffix *-him* is added as illustrated above with *'oimmeḍ* 'wander,' which becomes *'oimmel(i)him* 'wandering around.' In some cases, the morphemic component /i/ "is elided, but the effect on the preceding consonant remains" (p. 67). The author's description of the phonological contexts in which either /a/ or /i/ appear is limited to the observation that they appear between consonant-final verb-stems and *-him* as well as some other suffixes; no other details are given. The appearance of these vowels with the progressive marker will be treated in more detail in §5.3. At a later date, Mathiot (personal communication, 2018) also suggested the possibility that /i/ might be a separate morpheme with its own meaning, although this has not been definitively corroborated.

In sum, different scholars describe the different yet related forms of the progressive marker in varying ways; however, what is clear is that they all associate (explicitly or implicitly) a progressive type meaning with the related forms of the suffix – one that extends the activity expressed by a particular predicate in time. In the case of Mason (1950), Saxton

(1982), Saxton et al. (1983), and Zepeda (1984), a motion meaning is sometimes also associated with the suffix, providing a clue regarding its original source.

5.2.2 The historical source for *-him*

This section addresses the hypothesized source for the progressive marker based on observations for O'odham specifically as well as Uto-Aztecan as a whole, as well as its hypothesized evolution into the progressive marker.

O'odham scholars suggest that the progressive marker *-him* evolved from – or at least recognize that it is related in some way to – the verb *him* 'go, come walk' (Mason, 1950; Saxton, 1982; Saxton et al., 1983; Zepeda, 1983; see also Fitzgerald, 2004). Historically, the verb *him* 'go, come, walk,' probably occurred within [Verb-Verb^{him}] formations in which *him* was the second member.

Table 24 summarizes the meaning labels assigned to specific forms of the progressive marker discussed in §5.2.1, those who note a motion meaning for the marker, and those who explicitly suggest its original source or, at a minimum, comment on the semantic relationship of the progressive marker with the independent verb, *him* 'go, come, walk.' Parentheses ( ) indicate that the author interprets the sound to be epenthetic or, in the case of /i/, a "morphemic component" (Mathiot, 1973, p. 67) that is sometimes elided.

	Progressive/ Continuative	**Past Imperfective /Progressive**	**Interruptive (*ongoing action* can be resumed after interruption)**	**Motion**	**Suggested Historical Source**
Mason (1950)	*-him, -im, -ahim*			*-him, -im, -ahim*	*him* 'go, come, walk'
Hale (1959)	*-him, -(a)him*				
Saxton (1982), Saxton et al. (1983)	*-him*	*-ahim*		*-him*	*him* 'go, come, walk'
Zepeda (1984)	*-(a)him, -im*			*-(a)him*	*him* 'go, come, walk'
Mathiot (1973)			*-(a)him, -(i)him*		

Table 24. Functions and historical source associated with the progressive marker.

The association of both motion and progressive type meanings with the suffix and the hypothesis that the verb *him* 'go, come, walk' is the source are in line with Bybee et al.'s (1994) cross-linguistic findings. The findings suggest that expressions of the progressive aspect can in some cases originate from motion verbs.

Beginning with an examination of dictionary entries and grammatical descriptions of O'odham, the verb *him* has been assigned various meanings, the most relevant of which convey motion and are provided below:

- "to go" (Mason, 1950, p. 52)
- "to walk, go on foot in the direction or the manner specified by context (such as to come, to go somewhere, to go away)" (Mathiot, n.d.a, p. 261)
- "to go on for a certain amount of time" (Mathiot, n.d.a, p. 262)
- "to go" (Saxton, 1982)
- "*V* move along; progress; walk…" (Saxton et al., 1983, p. 22)
- "walking" (Zepeda, 1983, 1984)

I include a few examples below in (15)-(17) from the data which illustrate the motion usages of the independent verb, *him,* glossed as 'go,' 'come,' and 'walk,' respectively. The third line in each example reflects Juan Dolores's original word-for-word glosses.

(15) …m-o 'ab kaij g 'O'oi Wonamim m-apt-ṣ
…SBDR-3. AUX.IPFV LOC say.IPFV ART Painted Hat SBDR-2SG.AUX.PFV-QUOT
…that there said the Painted Hat that you

'am o **him**-k 'am o ñei,
LOC FUT **go**-SUCC LOC FUT see.PFV
there will go there will see
'[and he arrived there and told him], "Painted Hat [person's name] said that you will go over there to see him.' (*Boy*: 135)

(16) C-ḍ a cum si ha-ñeid-a-him g hemajkam
1PL.AUX.IPFV-RMT although INTS 3PL.OBJ-look.at-EP-PROG ART people
we although were looking very carefully the people

m-at 'ab hu wo **him**-a-d g
SBDR-3.AUX.PFV LOC FUT **come**-EP-CONT ART
that from somewhere will be coming the

t-ga:g-i-him-ḍam tt o wo:p 'u:hum…
1PL.OBJ-look.for-EP-PROG-NMLZ 1PL.AUX.PFV FUT run.PL.PFV back
one looking for us we would run home
'We had one kid keep watch in case somebody came looking for us, so that we would run home…. (*JD*: 286, 311)

(17) Gmhu **him**-a-c 'am cegito-him m-an-s has
over.there **walk**-EP-CONC LOC think-PROG SBDR-1SG.AUX.IPFV-DUB what.ABST
over there walking there thinking asking what

ñ-ju:-k o ñ-do'ibia.
1SG.REFL-do.PFV-COMPL FUT 1SG.REFL-save.PFV
to do will save myself
'As I walked I was wondering how I could escape.' (*JD*: 254; 308)

The various motion meanings that are attributed to *him* suggest that, prior to becoming a grammatical marker for the progressive, the verb's meaning had already generalized. It conveys movement or motion along a path – in any direction. This basic

meaning makes it a prime candidate for grammaticalization. In fact, the directionality is often determined by the verb's co-occurrence with particular locatives rather than the verb, itself. For example, in (15), the locative *'am* adds a translocative motion meaning to the verb, while *'ab* conveys a cislocative motion meaning in (16).

Mason (1950) categorizes the progressive suffix *-him* as one of a small handful of possible "subordinate compounded verb stems" (p. 52) in O'odham, a category which refers to a small set of verbal or verb-like suffixes that can combine with another verb stem in "subordinate" or second position as shown in (18) below. In the case of *-him*, Mason (1950) posits that the suffix is "certainly related to the common verb *him* 'to go'" (pp. 52-53) and also, as noted earlier, that the verbal suffix can carry the meaning "motion simultaneous with the action of the main verb" (pp. 52-53). The clearest illustrative example that he provides is the following:

(18) 'i-ge-wipia-**him**[18]
DIR.IMP-IMP-hunt-**GO**
'go and hunt around here' (Mason, 1950, p. 53; modified example)

Similarly, Saxton (1982) implies that the verb *him*, which he translates as "to go," is the source for *-him*. He lists *-him* as one of a small handful of suffixes with clear verbal sources that can combine with another verb-stem to create another verb (see section "V+V=V" in Saxton, 1982, pp. 175-177). Saxton provides the general gloss 'to go along doing' for constructions in which the suffix appears and gives the following examples:

(19) 'a:g-a-**him**
tell-GER-**PROG**
'to go along telling' (Saxton, 1982, p. 175)

[18] Mason's original morpheme breaks.

(20) ceggia-**him**
fight-**PROG**
'to go along fighting' (p. 175)

(21) ma:k-**him**
give-**PROG**
'going along giving' (p. 62)

It is not entirely clear based on the morphological glosses, translations, and lack of context in the above examples, however, if Saxton is suggesting that some entity is physically going along a path in the constructions or if he uses the translation 'going along V-ing' to express an activity that is occurring over some period of time. At a minimum, Saxton's translations suggest a connection between the physical movement and the progressive meanings.

Zepeda (1984) also implicitly suggests the relationship between the verb *him* and the progressive suffix, *-him.* In her analysis, Zepeda considers the possibility that the suffix is, in fact, the independent verb *him* which she glosses as 'walking.' The author also states that *him* in combination with another verb stem perhaps "should be considered a compound" (p. 103). However, Zepeda argues, there are instances where both the independent verb, *him*, and the suffix *-him* occur in the same sentence, where the verb form, *him*, clearly carries the meaning of 'ambulatory movement' whereas the suffix, *-him*, instead is used to extend some action conveyed by a verb to which it attaches. A few of the examples Zepeda uses to illustrate this point are as follows:

(22)	Husi	'o	si	gikuc-**him**	c	'ab	**him**.
	Joe	3.AUX.IPFV	EMPH	whistle-**PROG**	CONJ	LOC	**walk**.IPFV
	Joe	aux	very	whistling:along	and	there	walking[19]

'Joe is walking along and is really whistling.' (Zepeda, 1984, p. 104).

[19] Zepeda's original glosses.

(23) 'Ab 'o **him** g kawyu kc si
LOC 3.AUX.IPFV **walk**.IPFV ART horse CONJ INTS
there aux walk g horse and very

'e-mo'o-gid-a-**him**
3.REFL-head-shake.DNMLZ-EP-**PROG**
himself:shaking head along
'The horse is walking along shaking its head.' (p. 104)

(24) 'Ali 'o 'ab **him**-a-c ṣoak-**him**.
baby 3.AUX.IPFV LOC **walk**-EP-CONC cry-**PROG**
baby aux there walk:and going:along:crying
'The baby is walking along crying.' (p. 103)

The assumption is that, as with Saxton (1982), Zepeda recognizes the semantic relationship between the verb *him* 'go, come, walk' and the progressive marker.

Finally, Langacker (1977) notes of Uto-Aztecan languages overall "that compounding is probably one stage in the evolution of independent verbs into verbal affixes" (p. 75) and in some cases "the boundary between compounding and affixation will not always be sharp, being essentially a matter of degree of grammaticization" (p. 75). The author states that particular verb-like affixes in Uto-Aztecan languages are characterized as such because "they often derive historically from verb roots, [and] convey notions that are often expressed periphrastically by verbs…" (p. 133). Langacker also points out that verbal suffixes which specifically express motion are common in Uto-Aztecan languages. Furthermore, "[m]any are transparently related to independent verbs or can be seen to bear such a relation diachronically" (p. 147). These suffixes, according to Langacker, commonly convey the meanings 'go,' 'come,' and 'go around.'

Assuming that [Verb-Verb^{him}] compounds are the source for the progressive marker in O'odham, this offers a likely link to why the previous literature often ascribes both a progressive and motion meaning to the suffix *-him* as noted in §5.2.1 as well as in Table 24.

Subsequent to [Verb-Verb^{him}] compounds, I hypothesize that the verb likely extended towards and was reinterpreted as an **associated motion** grammatical suffix. Such morphemes add some type of motion meaning to the verb to which it attaches. Previously, these grammatical markers have been described in languages of Amazonia and South America, as well as Aboriginal languages of central Australia (Guillaume 2019). More specifically, these morphemes expresses **translational motion** or "spatial displacement/change in location" (p. 81). This meaning can be illustrated for the O'odham suffix, *-him*, in a number of the previous examples (see (18)-(21) above; also see (38) in §5.5.3). In these cases, it is clearly the suffix, itself, that adds the motion-meaning component to the verbal event. It may be the case that this function of *-him* served as an intermediary stage between [Verb-Verb^{him}] compounds and the progressive function of the suffix. The development of *(-)him* can be schematically represented as follows:

(25) [Verb-Verb^{him}] → [Verb-*him*$^{\text{TRNSLOC}}$] → [Verb-*him*$^{\text{EARLY.PROG}}$]

The combined synchronic observations for O'odham in particular and Langacker's (1977) discussion for Uto-Aztecan as a whole suggest the likelihood that [Verb-Verb^{him}] compounds served as the source construction for the progressive marker *-him* in O'odham. Furthermore, specific components inherent in the meaning of *him* 'go, come, walk' that overlap (i.e. are polysemous) with the progressive marker *-him* point in this direction.

5.3 Determining the phonological shape of the progressive suffix

Before moving on to the primary study, it is necessary to establish the form(s) of the progressive suffix that should be included in the study of the grammaticalization of the progressive aspect given the various and sometimes conflicting accounts of the /a/ and /i/ that

can precede the progressive marker. As discussed in §5.2.1 and summarized in Table 24, different authors interpret the variable forms of the suffix in different manners. Mason (1950), Saxton (1982) and Saxton et al. (1983) sometimes associate *-ahim*, in other cases *-him* and *-im* with progressive meaning. In another case, Saxton (1982) interprets the /a/ preceding *-him* as a separate past gerundive suffix, *-a*. Zepeda (1984) argues that /a/ is purely epenthetic, as does Mathiot (1973). Mathiot additionally notes the possible presence of /i/ (i.e. *-ihim*), referring to it as a "morphemic component" and also contemplating whether or /i/ in this context should in fact be treated as a separate morpheme whose meaning is yet to be determined. Resolving this issue is crucial because, as discussed in §1.1, theories of grammaticalization include the hypothesis that unidirectional grammatical change can be paralleled by phonological reduction in form – but not the reverse. If *him* 'go, come, walk' is the source for the progressive marker, then its form is expected to either be maintained as *-him* or reduced as we see sometimes in the suffix's manifestation as *-im*, but it should not be augmented (e.g. *-ahim, -ihim*). Based on a brief study conducted here, the conclusion is that the presence of /a/ and /i/ can be explained on purely phonological grounds. The vowels do not, in the majority of cases, appear in the same phonological contexts and are in complementary distribution. As a result of these findings, I interpret the progressive suffix to be *-him* and in on occasion reduced to the form *-im*.

As Zepeda (1984) notes, an epenthetic vowel generally appears following all consonant-final verbs appearing with a suffix (see examples in (9) in §5.2.1 above). Based on this information, I identified examples of verb stems with all of the possible final consonants in combination with *-him* and the form *-im* to determine where /a/, /i/, and also /_/ tend to occur. The possibilities are: <p>, <b>, <t>, <d>, <ḍ>, <k>, <g>, <s> <ṣ>, <m>, <n>, <ñ>,

<l>, and <w>. To illustrate the process, I identified examples of /t/-final verb stems such as *ciñvot* 'get, grow silk (of corn)' occurring with *-him* or *-im* to determine if either /a/ or /i/ can precede the suffix. In all of the /t/-final verb stems co-occurring with *-him* or *-im* that were examined, /a/ always appeared between the two components (e.g. *ciñvot-a-him*; see list of examples in Table 25).

Tokens were collected in the following manner: First, verbs occurring with *-him* and/or *-im* were identified in the oral texts and recordings used for this study. However, this proved to be an incomplete list of possible consonantal verb stem endings occurring in combination with *-him*/*-im*. In order to establish a complete list of all of the possible combinations, I conducted a random search of several different verbs with the specific final-consonants lacking from the list in Mathiot's (1973) dictionary. Once I identified a sample of verbs with all of the possible final-consonants, I carried out an exhaustive search for all examples of the specific verb with *-him* and *-im* in both the texts and in Mathiot (n.d.a, n.d.b). The search included all of the following possible combinations for each of those verbs:

- V + *-ahim*
- V + *-him*
- V + *-ihim*
- V + *-im*

What became apparent, however, is that Mathiot (n.d.a, n.d.b) does not record the form *-im* in the dictionary, although there is an indication that Mathiot heard the reduced form. In a phonetic transcription of the *Coyote Stories* (Pancho & Mathiot, 1959-1960), Mathiot transcribes the progressive marker as *-him, -*^{h}*im, -im,* and in one case *-am,* in the latter three cases noting either a reduction of the initial /h/ or a lack thereof. Besides a few examples

from the *Coyote Stories*, the reduced form of the suffix otherwise occurs in only a few cases in the historical data.

Table 25 and Table 26 below list a) all of the verbs that I analyzed, b) the forms of the progressive marker, *-him* or *-im* that co-occurred with a given verb, as well as c) the presence (or lack thereof) of intermediary /a/ or /i/. In cases where only one verb is listed, this can be taken to mean that other examples occurred neither in the texts nor in Mathiot's dictionary. The notations used in the tables are as follows:

- *-(i)-* or *-(a)-* = The specific verb + *-him/-im* is found both with and without the interceding vowel (e.g. *ga:g-(i)-him* means that *ga:g* + *-him* occurs both as *ga:g-i-him* and *ga:g-_-him*).
- *-a-* or *-i-* = The specific verb + *-him/-im* always appears with the interceding vowel
- *-_-* = The specific verb + *-him/-im* always occurs without the interceding vowel in the examined examples.

Verb-final consonant	Verb[20]		With *-him*
<p>	*gewṣp*	'freeze (of the ground)'	*gewṣp-_-him*
	jukṣp	'rain on'	*jukṣp-(a)-him*
	kekiwup	'stand up, stop repeatedly'	*kekiwup-_-him*
	bidṣp	'to paint/ dirty an obj'	*bidṣp-a-him*
	'e-ciosp	'support oneself with a cane'	*'e'cioṣp-a-him*
<b>	*ku:b*	'be blowing around (of dust or smoke)'	*ku:b-_-him*
	'e:b	'stop crying'	*'e:b-_-him*
	to:b	'twist obj'	*to:b-a-him*
	va'ub	'get dewy, damp'	*va'ub-a-him*
<t>	*ciñvot*	'get, grow silk (of corn)'	*ciñvot-a-him*
	cu'it	'clear up obj'	*cuit-a-him*
	wagt	'dig a hole'	*wagt-a-him*
	wakut	'make a flat surface out of obj'	*wakut-a-him*
	kait	'make, get, seeds	*kait-a-him*
<d>	*'a:gid*	'tell'	*'a:gid-a-him*
	ñeid	'see, look at, watch'	*ñeid-a-him*
	namkid	'pay'	*namkid-a-him*
	'amicud	'recognize, understand'	*'amicud-a-him*
	ba:mud	'ask for a hand/help'	*ba:mud-a-him*
<ṣ>	*'e-ma'ikkoṣ*	'bounce several times (reflexive)'	*'e-ma'ikkoṣ-a-him*
	keikkaṣ	'kick obj reiteratively'	*keikkaṣ-a-him*
	celitṣ	'scratch'	*celitṣ-a-him*
	cekṣṣaṣ	'make, draw, several unbroken lines in one location (e.g. several furrows in field)'	*cekṣṣaṣ-a-him*
	ta:tṣ	'split obj in several strands (of beargrass or yucca for basket making)'	*tatṣ-a-him*
<m>	*bijim*	'go slowly past/ around'	*bijim-_-him, bijimm-_-im*
	maṣcam	'teach'	*maṣcam-a-him*
	ṣo:m	'sew a seam on'	*ṣo:m-a-him*
	tatam	'touch repeatedly'	*'amicud-a-him*
	hehem	'laugh'	*hehem-(i)-him*
<n>	*ta:pan*	'split, crack'	*ta:pan-a-him*
	ñe'owin	'quarrel with'	*ñe'owin-a-him*
	gewittan	'hit with a stick'	*gewittan-a-him*
	wansan-[21]	'laugh'	*wansan-a-him*
<w>	*meliw*	arrive somewhere running or driving	*mel[u]w-_-him*
	ṣaṣaw	echo'	*ṣaṣaw-_-him*
	tataw	'shine (said of the sun)'	*tataw-_-him*
	'uiw	'to fart'	*'uiw-a-him*

Table 25. Predominant patterns for verbs with final consonants <p>, <b>, <t>, <d>, <ṣ>, <m>, <n>, <w>.

[20] All verbs are in their imperfective forms.

[21] Form only occurs with a suffix.

Verb-final consonant	Verb		With *-him*
<k>	*ma:k*	'give'	*ma:k-_-im*
	ṣoak	'cry, rattle, moo, bleat'	*ṣoak-(i)-him*
	totpk	'boil (food)'	*totpk-(a)-him*
	ku:k	'make a roaring sound (of wind, ocean, etc.)'	*ku:k-_-him, ku:k-_-im*
	mu:k-[22]	'die'	*mu:k-_-him*
<g>	*ga:g*	'look for, hunt for'	*ga:g-(i)-him*
	bihag	'coil, surround, corner'	*bihag-_-him*
	huḍḍawag	'coax, pester'	*huḍḍawag-i-him*
	'a'amog	'make an announcement (PL.SUB)'	*'a'amog-_-him*
	huhug	'vanish, disappear, pass away; run out, exhaust	*huhug-_-him*
<s>	*vakois*	'be curved'	*vakois-_-him*
<c>	*je:kc*	'to find the tracks left by obj once	*je:kc-_-him*
	ba'iwc	'to get ahead of obj	*ba'iwc-_-him*
<j>	*gikuj*	'whistle'	*gikuj-_-him, gikuj-_-im*
	kaij	'say, make a sound'	*kaij-(i)-him, kaij-_-im*
<ñ>	*sispañ*	'break open obj'	*sispañ-_-him*
	kopñ	'make a popping noise, explode or blast once'	*kopñ-_-him*
	huduñ	'go down'	*huduñ-_-*^{h}*im*
	wawañ	'form a line, a ridge'	*wawañ-_-him*
	wu:ş[a]ñ	'rise, come out; get out; be born'	*wu:şñ-_him*
<l>	*'ial*	'to roll over'	*'ial_him*
<d>→<j>[23]	*ceṣad-*[24]	'climb, reach the top; rise (of heavenly bodies)	*ceṣaj-(i)-him*
<ḍ>→<l>	*ge'eḍ*	'become bigger, grow'	*ge'el-_-him*
	ceweḍ	'become tall (of a plant)'	*cewel-_-him, cewel-_-im*
	heweḍ	'blow (wind), be blowing in the wind'	*hewel-_-him, hewel-_-im*
	'oimmeḍ	'be/stay around a place, wander'	*'oimmel-i-him*
<ṣ>→<s>	*kamṣ*	'to put obj into one's mouth'	kams-_-him
	mauṣ	'put one's hand somewhere'	*maus*-_-him
	ge:ṣ	'fall, get in'	*ge:s-i-him*

Table 26. Predominant patterns for verbs with final consonants <k>, <g>, <s>, <c>, <j>, <ñ>, <l>, <d> → <j>, <ḍ> → <l>, <ṣ> → <s>.

[22] Form only occurs with a suffix.
[23] An arrow (→) indicates a sound change that occurs in the presence of /i/. (See discussion of Mathiot, 1973 in §5.2.1.)
[24] Special form only occurs with a suffix.

The patterns included in Table 25 show that in a majority of cases /a/ appears between the verb stems with the final consonants <p>, <b>, <t>, <d>, <ṣ>, <n>, <w> and *-him/-im*. Only <m>-final verb stems showed a departure from this pattern. The verb *bijim* 'go slowly past or around' occurred without any interceding vowel and the suffix appeared in both the forms *-him* and *-im* (*bijimm-him* and *bijim-im*). Examples with the verb *hehem* 'laugh' appeared as either *hehem-i-him* or *hehem-_-him.*

Table 26 reveals that the most pervasive pattern for verbs with final consonants <k>, <g>, <s>, <c>, <j>, <ñ>, <l>, <d> → <j>, <ḍ> → <l>, <ṣ> → <s>, was the lack of a vowel (notated as _), although according to Mathiot (n.d.a, n.d.b), many of the verbs listed in Table 26 can potentially occur with /i/. While this may be the case, I identified few concrete examples in the data. Other patterns in the data were verbs followed by /i/ or /_/ (notated as *-(i)-*), or only /i/. The only consonant that displayed a different pattern was <k>-final verb stems as in the example *totpk* 'boil (food)' which could occur either with or without /a/ as in *totpk-(a)-him.* All other <k>-final verb stems had the pattern of /i/ or no vowel preceding *-him* or *-im.*

In a few cases in this set, the progressive marker is in the form *-im* (or on occasion the phonetically transcribed reduced from -[^{h}*im*] in the *Coyote Stories* (Pancho & Mathiot, 1959-1960). When Table 25 and Table 26 are compared, this occurs predominantly with the verbs where /i/, according to Mathiot, can potentially appear as an interceding vowel; it never appears with a verb when the interceding vowel is otherwise /a/ in the data, suggesting that the reduced form of the suffix is context-specific.

The data largely suggest (though not exclusively) that the interceding vowels /a/ and /i/ occur in complementary distribution, providing evidence that their presence is a result of

the phonological environment in which they occur. Based on these findings, I view *-him* as the basic form of the suffix that is associated with the progressive meaning which on occasion reduces to *-im*. I interpret the vowels /a/ and /i/ as epenthetic and also assume that they are devoid of meaning (i.e. not treated as morphemes).

5.4 Data and token selection

In order to investigate the grammaticalization of progressive constructions involving the suffix *-him* in O'odham, I analyze data from the three different time periods. This includes a) the historical data from the early 1900s, b) the midcentury data which is based on materials collected during the late 1950s and early 1960s, and the modern data from the early 2000s. I describe the details concerning the data at length in §4.1. As stated earlier, the purpose of looking at data from three different time periods is to provide a hypothesis regarding, in this case, the grammaticalization of O'odham progressive constructions based on detected meaning changes in the data which spans over approximately 100 years and how those meaning changes correlate with the early-, middle-, and late-shift periods; while written records suggest that very few O'odham people spoke any English during the early-shift period and were primarily monolingual O'odham speakers, by the late-shift period of which we are in the midst today, over half of the population are monolingual English speakers (see timeline summary in Figure 5, §2.4.4). The last goal is to decipher what conclusions can be made based on these findings regarding language change and language shift.

Table 27 below includes the number of tokens of the progressive marker occurring in the historical, midcentury, and modern data collections and their combined total. Following the table, I explain how the tokens were identified.

	Total # of tokens
Historical	46
Midcentury	50
Modern	24
TOTAL:	120

Table 27. Token totals for progressive marker in each data period.

After entering the **historical data** into Fieldworks (FLEx) free software, I used the program's search capabilities to identify and extract forms most commonly associated with the progressive marker, namely *-him* and sometimes *-im* (see §§4.1.1 and 4.2). In a couple of rare cases, the form *-m* also occurred in the data as a possible example of the progressive marker based on its meaning. I identified and extracted approximately 52 tokens in total for coding. I omitted six tokens from the study after determining that they were not, in fact, examples of the progressive marker. For example, the desiderative also takes the form *-im*, and so I eliminated those tokens that clearly conveyed a desiderative rather than progressive meaning from the analysis. In a couple other cases, the tokens were in sections of the texts that were songs, a narrative style that is not being considered here. Forty-six tokens remained after these omissions.

The **midcentury data**, as noted in Chapter 4, are comprised of four different *Coyote Stories* (Pancho & Mathiot, 1959-1960) told by Jose Pancho as well as Mathiot's (n.d.a, n.d.b) two-volume dictionary which the author made publicly available online in the early 2000s (no precise date is provided). All of these materials are available online as downloadable PDFs (see link in References).

For the *Coyote Stories* in particular, Mathiot conveniently includes a phonetic transcription of the texts, as previously noted, in addition to versions in Mathiot's

orthography. After the *Coyote Stories* were downloaded, I conducted a search for each instance of the progressive suffix in the word-search bar. The forms that the progressive marker can take in the phonetically transcribed versions were as follows: *-him, -ʰim, -im,* and in one case *-am.*[25] I identified all of these cases as the progressive based on the fact that in the versions of the stories that are in Mathiot's orthography, they are given in their regularized form, *-him*, which Mathiot refers to as marking the "interruptive" aspect (see discussion of author's use of this term in §5.2.1).

For each lexical entry in the digitized version of Mathiot's (n.d.a, n.d.b) two-volume dictionary, as described in §4.1.2, the author includes several illustrative examples taken directly from various orally produced narratives to illustrate how the given lexical item is used in authentic discourse. Numerous examples of the progressive exist throughout the dictionary. An example is provided below (from Mathiot, n.d.a, p. 161).

gakşp Vintr Unit (for Repet see **gakşşap**; for Distr see **gagakşp-**) [Neutr: def gakş. Correl: gakş-k; immed: gakş-ka-'i. Interr: gaks[i]-him /gakşpa-him/ gakşp[i]-him]: to get dry, to dry out (such as plants, clothes on a line)
ex: Gakş 'atkï g gakidaj. The saguaro fruits are dry (have gotten dry).--Gakş 'atkï g nşoñiaka. The wood that I cut is dry. --Gakş g 'eñga. The clothes (on a line or on bushes) are dry.--Gakş g havihogdag k 'egañva g mu:ñ. The pods have gotten dry and the beans spilled out. --…mat hekid 'ab o 'i gakş hega'i ha:hagaj. …when its leaves (of tobacco plants) have dried out. --Hu:ñ 'o gakşpahim. The corn is drying out (as because of lack of water). --Gakşpahim g ha'icu m'e'es. Your plants are drying out.--Pi 'ipt(<'apt) kia vapga g 'e'e'es k abş gakşpahim. You have not watered your plants yet, and they are drying out.--Gakshim 'o g bahidaj. The saguaro fruits are getting dry (on the plants). --… k ha'i 'ab 'i gaksihi g bahidaj. …and some of the saguaro fruits up there (on top of the saguaro) were getting dry.

In order to select a sample of tokens of the progressive, I randomly selected approximately 20 pages in both dictionary volumes (totaling around 40 pages). As noted earlier, some

[25] Mathiot does not use the standard IPA representations for the phonetic transcriptions. For this reason, I do not utilize the standard brackets.

examples in Mathiot's dictionary are from the *Coyote Stories.* In the case that a dictionary entry included an example sentence that I had previously collected from one of the *Coyote Stories,* I selected a different example in order to avoid any redundancies. In all other cases, I identified and selected the first entry on a page with an example sentence of the progressive marker for inclusion in the sample. In an effort to have as wide a range as possible of verbs with which the progressive marker occurs, I selected one example per verb from the dictionary for analysis. Otherwise, I made no effort to find specific types of examples in Mathiot's dictionary entries in order to procure as random a sample of progressive constructions with *-him* as possible. During the identification process, I noted that Mathiot (n.d.a, n.d.b) generalizes the form of the progressive marker to *-him* in all cases. As a result, I did not identify any alternate forms of the progressive marker (such as *-im*) in the dictionary.

The progressive marker from all of the midcentury data combined totaled 50 tokens. It should be pointed out that while the examples in the *Coyote Stories* represent an exhaustive list, the examples from the dictionary do not. The goal instead was to roughly match the number of tokens that occur in the historical texts.

Finally, I entered the modern data, which are comprised primarily of YouTube videos, into FLEx (see §§4.1.3 and 4.2). I identified progressive marker forms in the same manner as I did for the historical data (using the FLEx search function). Twenty-four tokens occurred in the modern data.

5.5 Coding procedures for progressive constructions with illustrative examples

In this section, I detail how I code the data in order to analyze *-him* constructions (primarily) once they have achieved progressive status in O'odham. I also interweave a number of examples into this discussion based on my own descriptive analysis of the data. The

examples both illustrate and justify why the coding criteria I have selected are necessary as well as what the criteria highlight about different facets of progressive constructions.

I base much of the coding procedures on specific hypotheses presented in Bybee et al. (1994) which I describe in §5.1. To recap, they propose that progressive constructions in their earliest stages of grammaticalization occur with predicates that entail more agentive subjects or more subject control over the verbal event (see §5.5.1). They occur with dynamic predicates which describe situations that involve change over time (Comrie, 1976). In the case of progressive constructions that have motion verbs such as *go* and *come* as their source, they are hypothesized to originally depict a subject located somewhere in space "moving (around) doing something" (Bybee et al., 1994, p. 33), or in the case of O'odham, moving along a path doing something. As the constructions grammaticalize, they may expand, occurring with predicates that entail less subject control or agency and stative predicates that describe situations which involve very little observable change over time. The meaning of the construction may also generalize to express situations in which neither the expression of location nor a motion meaning are present as is the case with ongoing and developing states. The development of the different meaning components of earlier and more advanced progressive constructions as a whole are summarized in (26).

(26) **[Verb-*him*]$_{\text{EARLY.PROG}}$** → **[Verb-*him*]$_{\text{PROG}}$**
More subject control → Less subject control
Dynamic predicates → Stative predicates
Path-motion meanings → Non-motion meanings

Based on these premises, I code each construction in which a token of *-him* occurs for the following: **a)** subject control (i.e. agency) based on the meaning of the predicate, **b)** situations described by the predicate (i.e. dynamic vs. stative), and **c)** motion meaning types

(or lack thereof) expressed by the progressive construction as a whole. A detailed categorization of these parameters can potentially shed light on how the progressive aspect has evolved or changed in O'odham over the time period I am examining here. I describe each of these criteria in greater detail in this section.

Lastly, I conduct a separate analysis of whether the predicates to which the progressive marker attaches carry specific meanings (e.g. *walk, hear, eat,* etc.) versus generic meanings (e.g. *do, happen, make*). This is in order to determine if collocations between semantically specific versus generic verbs and the progressive marker show any differences over time. This is worth exploring given that in contexts of severe language shift, speakers may tend to use lexical items with less specific meanings more often (Mithun, 1990).

5.5.1 Subject control

When a subject is examined in relation to the predicate with which it co-occurs, it becomes clear that **control** must be taken into account in gauging subject agentivity. Schlesinger (1995) describes control as follows: "[A]n entity that is in CONTROL of the event … steers the activity in the event and may be able to terminate or obviate it" (p. 33; original emphasis). In their earliest stages, progressive constructions hypothetically occur with predicates that entail more subject control (i.e. highly agentive subjects) and gradually expand to include subjects with less control (i.e. less agentive or non-agentive entities such as patients). I elaborate on the concept of control in what follows and layout how I code for subject control in the data.

To begin with, Comrie (1981) states that the distinction between semantic categories such as agent, patient, and others do not represent "discrete semantic relations, but rather a continuum…" (p. 53) and this continuum "can be regarded as a continuum of control" (p. 53). Croft (1998a; 2012) also argues that on one end of a continuum, subjects that are

interpreted to carry the most agent-like qualities tend to be involved in what Croft refers to as Controlled Activities (see Croft's examples in (27) below). On the other end of the spectrum are subjects that lack control and tend to be categorized as patient-like. Subjects in these cases are involved in Uncontrolled Activities as well as Transitory States "which generally come about via an external cause or at least an internal cause not fully under the participants control" (Croft, 2012, p. 258). These findings are based on Croft's multidimensional scaling analysis[26] of the following set of semantic predicate classes. The author identifies these classes in the active-inactive systems described in Holisky (1987), Gregores & Suárez (1967), and Mithun (1991):

(27) Controlled Activities: agentive processes, e.g. *run, dance, go out, etc.*

Inactive Actions (Holisky's 'locative statives'): *sit, stand, lie, hang, live, stay, etc.*

Bodily Actions: normally uncontrolled, e.g. *cough, sneeze, shiver, sweat;* some are more controllable e.g. *laugh, spit, urinate*

Other Uncontrolled Activities: e.g. *die, slip, grow, trip, get lost,* etc.

Dispositions: properties of actions that are also attributed to inherent traits of individuals: *proud, wise, evil, courageous, jealous,* etc.

Inherent Properties (*states)*: construed as permanent unchanging properties of entities, e.g. *red, tall, round, etc.*

Inchoatives (of dispositions/properties): *become proud, wise, tall, etc.*

Transitory (Mithun's '+affect') States: stative properties that are temporary and thus have come about through some process, e.g. *sick, tired, old, cold, etc.* (Croft, 2012, pp. 257-258; also Croft, 1998a, p. 52).

Croft's ranking for [A]gent-like and [P]atient-like coding tendencies in the active-inactive systems is as follows:

[26] The reader is directed to Croft (2012, pp. 128-132) for a description of this statistical analysis technique.

(28) *most A-like*
Controlled Activities
Inactive Actions
Inherent Properties, Dispositions
Bodily Actions
Inchoatives
Uncontrolled Activities, Transitory States
most P-like (Croft, 2012, p. 258)

These findings highlight that how agentive a subject is interpreted to be or how much control a subject is interpreted to have is directly associated with the type of activity the subject is engaged in. In other words, subjects cannot be analyzed in isolation but rather "[n]otions like control and experiencer refer to a relation between the predicate and one of its arguments" (Comrie, 1981, p. 56), and the more control a subject has over the event expressed by the predicate, the more agency it is interpreted to have, indicating a continuum.

Given that grammaticalization is a process that is gradual and develops along a continuum, analyzing subject types through the lens of control in order to determine subject agentivity can provide nuanced insights into how progressive constructions gradually change over time. The following progressive constructions in the O'odham examples below illustrate how subject control (based on the predicate) and subject agentivity are directly linked. In (29), some boys (mentioned in the lines leading up to this example) are collecting and rolling snow into a snow ball, and in (30) an older, more grown up kid groans due to pain he is experiencing from being physically punished; in the third example, (31), someone is asking another person how long they have been locked up in jail. All three subjects are human (and therefore animate), but they all execute different degrees of control and agency based on the different activities they are engaged in. The subjects in (29), the boys, display more agency

than the kid in (30), and the kid in (30) in turn has more agency than the person in (31) who is locked up in jail presumably against their will. The different degrees of agency are directly related to the degree of control the subjects have in executing the activities expressed by the predicates. The event of taking up snow in (29) requires more control than the act of groaning in (30) over which the subject perhaps has marginal control due to pain he is experiencing from being physically punished and, in (31), the subject has very little control over being locked up and would traditionally be referred to as more patient-like.

(29) Ḍ 'an bei-**him** g gew k hema 'olat
COP there take-**PROG** ART snow CONJ one roll.into.a.ball
there they were taking up the snow and one ball made

ge'ge'e
become.bigger.PLS
a very big one
'They [some boys] took some snow and built a huge snowball.' (*JD:* 248, 305)

(30) Hema hegam ge'egeḍ ṣoṣṣañ-**him**,
one those grown.ups groan-**PROG**
one those grown ups was groaning
'One of the older kids was groaning[,]' (*JD*: 291, 312)

(31) He'es 'ap 'i taṣ ku:pĭ-ka-**him**?
how.much 2SG.AUX.IPFV PUNC day be.locked.up-STAT-**PROG**
How long have you been locked up (in jail)?' (Mathiot, n.d.a, p. 449)

In order to come up with the most accurate scale of subject agentivity, I take into account subject control over the event expressed by the predicate. For each token of a progressive construction, I code for Croft's (1998a, 2012) activity types given in Table 28 below (described in greater detail in (27) above). The categories should be viewed as existing on a continuum rather than as separate and discrete. For example, Inactive Actions involve less control than Controlled Activities but more control than Inherent Properties and Dispositions. I give each token a ranking of 1-6, arranged from *most* to *least A[gent]-like*. I

use the label *least A-like* in place of Croft's *most P[atient]-like* as a cover term in order to include additional semantic roles such as experiencer.

	Activity Scale	**Activity Types**	**Examples**
Most A-like	1	Controlled Activities (agentive processes)	*run, dance, go out, throw etc.*
	2	Inactive Actions	*sit, stand, live, stay, watch, etc.*
	3	Inherent Properties	*red, tall, round, etc.*
		Dispositions	*proud, wise, jealous,* etc.
	4	Bodily Actions	more controllable: *laugh, spit, urinate;* less controllable: *cough, sneeze, shiver, sweat*
	5	Inchoatives	*become proud, wise, tall, etc.*
Least A-like	6	Uncontrolled Activities	*die, slip, grow, trip, get lost,* etc.
		Transitory States	*sick, tired, old, cold, etc.*

Table 28. Activity scale (based on Croft, 1998a, 2012).

For examples (29)-(31) above, I ranked them based on the activity scale as follows: 1 (boys taking up snow), 4 (a kid groaning), and 6 (a person who is locked up against their will).

5.5.2 Situations described by the predicate

Next, the expectation is that the progressive aspect occurs in its earlier stages with dynamic predicates and later expand to include stative predicates, both of which represent and describe different situation types. I elaborate on each of these categories below.

Dynamic predicates express situations involving "change or movement" (Dahl, 1985, p. 28) or some "change over time" (Croft, 2012, p. 34), and different phases of an activity can be detected from one moment to the next (Comrie, 1976) as with actions like "*write, walk, sneeze, ripen, drop*" (Bybee et al., 1994, p. 55; original emphasis). Included

under a definition of dynamic verbs are process verbs, a sub-category of dynamic predicates that express contexts involving "a change of state, such as *grow, improve, ripen, turn red*" (Bybee et al., 1994, p. 55; original emphasis).

In contrast to dynamic verbs, stative verbs depict situations that "do not change over time" (Croft, 2012, p. 34), as illustrated by the examples *be short, be kind, know, believe,* etc. This also includes a class of verbs such as *sleep, stand, lie, think, see,* etc. which in some languages such as English and O'odham can occur in the progressive aspect even though they otherwise involve no motion or change over time. Croft (2012) refers to this verb type as expressing "Inactive Actions" (p. 39; see also Croft, 1991, p. 97; 1998, p. 72). These depict an otherwise stative activity which is construed as including a starting point where the activity begins, a duration of time where the activity holds or is interpreted as ongoing, and an eventual end point where the action ceases (see Croft's, 2012 discussion of aspectual construals).

To illustrate dynamic versus stative predicates in O'odham, I describe some examples below. In the first example, (32), the progressive verbal construction includes *'oiyopi* 'wander (plural),' a dynamic predicate that represents a situation involving movement and, in combination with the progressive suffix *-him,* depicts the subject as in the midst of the ongoing activity at reference time, and in example (33) the dynamic verb *huḍḍawag* 'pester' is marked by the progressive suffix to express the ongoing dynamic activity in which someone is pestering another (presumably animate) entity. Example (34) includes an instance of the verb *cewel*[27] 'become tall, grow' which conveys the process or change of state meaning of plants becoming tall.

[27] The <*ḍ*> becomes <*l*> with the progressive and other suffixes.

(32) 'Am o ṣa'i 'oiyopi-**him** c 'ep o 'i hihi.
LOC FUT a.little wander.PL-**PROG** CONJ again FUT here come.PL.PFV
there we would a little while wander around and again will come this way
'We would wander about there for a little while, and move back again.' (*Antonio*: 67)

(33) …'am a cum huḍḍawag-i-**him**. …
…LOC indeed try.in.vain.PART pester-EPEN-**PROG**
'...he kept pestering him in vain. (Mathiot, n.d.a, p. 290)

(34) Kuk 'oitag t-'am g hu:ni wu:ṣañ k cewel-**im**.
Owl Hooting Village field ABS-at ART corn rise.IPFV CONJ become.tall-**PROG**
Owl Hooting Village's garden upon the corn came out and grew taller and taller
'At Kuk [Owl Hooting] village field the corn grew, becoming taller.' (*War*: 336)

The next set of examples illustrate different progressive constructions that include stative predicates which portray situations where no change can be detected. More specifically, the stative predicates combine with *-him* to express ongoing states. In example (35), the verbal construction *s-taḍañ-ka-him* literally means 'being wide.' Three pieces of evidence indicate that *s-taḍañ* 'be wide' represents a stative context. First, the meaning of the verb itself describes the unchanging state of some object (unless some other entity were to alter it in some way), secondly, the verb co-occurs with the prefix *s-*, which is a stative verb marker in O'odham (although not all stative verbs occur with this prefix; see examples (36) and (37) below), and thirdly, once the progressive marker *-him* is suffixed to the verb, the additional stative marker *-ka* is present which is used when stative verbs conjoin with *-him* (see Mathiot, 1973; Saxton, 1982; Saxton et al., 1983). The addition of *-him* denotes the ongoing state, in this instance, of being wide. In (36), the stative, Inactive Action verb *wo'o* 'be lying down' combines with *-him* which co-occurs with stative suffix *-ka*, emphasizing the ongoing state of the cow having been lying dead for an extended period of time. Example (37) illustrates a case in which the progressive marker conjoins with the verb *ke:k* 'stand' to depict another Inactive Action. The verb on its own is semantically stative; no motion or

change occurs. Its occurrence with the progressive marker highlights the subjects' holding of their standing position for some duration of time.

(35) Cum 'o s-taḍañ-ka-**him**.
used.to.PART 3.AUX.IPFV STAT-be.wide-STAT-**PROG**
'It used to be wide [lit. being wide].' (Mathiot, n.d.b, p. 221)

(36) Hegai gi:g m-a-ṣ si s-toha ḍ 'aṣ
that tallow SBDR-3.AUX.IPFV-QUOT INTS STAT-white COP just

waptop m-at-p taṣ 'ama'i wo'o-ka-**him**
maggots SBDR-3.AUX.IPFV-ASSUM long.time LOC lying.down-STAT-**PROG**

hegai haiwañ mu:k.
that cow dead.
'That white tallow that he had been eating were maggots that had been there as a result of the meat lying there a long time. (*Buzzard*: 18)

(37) Tt gemhu 'i ke:k-**im** k si ha-hi:nad.
1PL.AUX.PFV over.there PUNC stand-**PROG** CONJ INTS 3PL.OBJ-yell.out
we over took our stand separately and very loud yelled at them
'We were standing about and shouted at them.' (*War*: 169)

In the data, each progressive construction is simply coded as occurring with either a dynamic or stative predicate, the subtypes of which are summarized in Table 29.

Situation types occurring w/ *-him*	**Sub-categories**	**Examples**
Dynamic	Activities	*run, write, sneeze, etc.*
	Processes (Change of State)	*ripen, grow, improve, etc.*
Stative	States	*be kind, tall, smart, etc.*
	Inactive Actions	*sleep, stand, think, see, etc.*

Table 29. Situation types and subtypes described by the predicate.

5.5.3 Motion types

Next, given that the progressive marker in O'odham likely evolved from the independent verb *him* 'go, come, walk,' I also code for the **motion type** (or lack thereof) expressed by the

construction as a whole. First, **motion** itself refers to any perceptible physical movement either of an entire entity (e.g. a person, animal, automobiles, objects of nature such as clouds, the sun, a river, etc.) or a part of an entity (e.g. a person or animal's legs, toes, mouth, the leaves of a plant moving in the wind, etc.). The specific motion types that I code for are as follows: a) **path-motion**, i.e. the subject changes location along a path, b) **random motion** where the subject 'moves around doing something,' c) **localized motion** in which the subject carries out some physical activity that an onlooker would be able to perceive, but the motion or movement occurs in a single location (i.e. no translational motion/spatial displacement), and d) **non-motion** which includes events that are lacking in all types of perceptible physical movement at the time of speech, e.g. inactive actions such as *thinking, sleeping, sitting*, process or change of state situations like *growing, decreasing, improving*, as well as stative contexts, e.g. *being nice, sick, big.* In these cases, the expression of location is not present (e.g. the subject cannot be said to be carrying out the stative activity in a specific location in space). Based on the premise that the progressive suffix evolves from *him* 'go, come, walk,' the progressive construction as a whole may be more likely in its earlier stages to convey some type of linear path motion meaning that is of the translational type (i.e. involves spatial displacement). In their more advanced stages, *-him* constructions are hypothesized to involve lesser degrees of motion. In what follows below, I describe each of the motion types in greater detail, and I justify the organization from path-motion to non-motion construction types based on polysemous meanings that exist in the data. I illustrate these meanings through a number of examples in O'odham.

Path-motion constructions

Assuming that the progressive marker *-him* evolved from the verb *him* 'go, come, walk' within [Verb-Verb^{him}] compounds, the intermediary stage hypothetically involved the gradual grammaticalization of *him* into a translational motion suffix (see §5.2.2), depicting spatial displacement along a linear path. This is illustrated in example (38) below which notably comes from the historical data from the early 1900s and may illustrate a less grammaticalized instance of, or a precursor to the progressive function of *-him/-im*. In fact, this is the only example of this nature appearing in the data from all three time periods. In this case, *-im* conjoins with another dynamic verb, *gikuj* 'whistle' to depict the agentive subject, the narrator, *Shawañ 'U:we* (Smells Like Crow), as in the midst of doing some activity, in this case whistling, while moving along a path in the location of some corn fields for some unstated period of time. In this instance, it is specifically *-im* that adds the spatial displacement meaning to the verbal event. Note especially Dolores's original gloss in the third line of *gikujim* as 'whistling, I am going.'

(38)	…nt	heg	ñeid-ok	juṣal	gikuj-**im**.
	…1SG.AUX.PFV	that	see.IPFV-COMPL	quietly	whistle-**go/PROG**
	…I	that	having seen	softly	**whistling, I am going**

'…having seen that [corn tassels blowing in the wind], I was going while whistling quietly.' (*War*: 338)

It is crucial to highlight that an additional meaning is inherent in the path-motion example in (38) above and one that may have pushed it along the grammaticalization pathway. Note that the act of whistling is extended not only in space as the agentive subject "goes" while whistling, but the action is also extended for an unstated period of time; that is, movement along a path in space also entails an undefined period of time during which the movement occurs. It may be the case that speakers subconsciously inferred and extracted this inherent

and more abstract meaning in earlier constructions (see discussion on 'inference' in §1.1), leading to the eventual grammaticalization of *him* 'go, come, walk' into a grammatical marker of the progressive – one that includes a progressive meaning where a subject is carrying out some activity that is "ongoing at reference time" (Bybee et al., 1994, p. 126).

Given the likelihood of the source for the progressive in O'odham, in the subsequent stage, early progressive constructions as a whole might be expected to more likely convey some type of path-motion meaning. There are a number of cases where the dynamic verb to which *-him* attaches is itself a path-motion predicate, and the complex construction *as a whole* conveys a translational path-motion meaning (i.e. the expression of spatial displacement cannot solely be attributed to the suffix, itself). In (39), Coyote is *bijimhim* 'going past' around a body of water in the middle of which Turkey is sitting. In (40), the sun is described as moving vertically along a path in the sky or *huḍuñʰim* 'descending' towards the horizon. Example (41) illustrates a case in which the progressive marker attaches to the source verb *him* 'walk' from which it likely evolved which indicates the ongoing activity of walking. In all cases, spatial displacement of the subjects along a path is evident.

(39) …'an has bijim-**him** ṣu:dagï we:gaj…
…LOC PART go.passed-**PROG** water around…
'…He circled the water ...' (*Coyote Turkey*: 33, 64, 75)

(40) Nia k 'oia woho 'am heba'i ju: g taṣ
see.DM AND thereupon truly LOC somewhere do.PFV ART sun

huḍuñ-**ʰim** 'am hi a'a…
descend-**PROG** LOC maybe think.PFV
'So, thereupon, as he had said, sometime when the sun was starting to go down, [Coyote climbed on top of his house and stood up there]. (*Coyote Skunk*: 31, 64, 80)

(41) K ab masma b wuḍ 'elida m-att
CONJ thus in.a.certain.way thus COP think.IPFV SBDR-1PL.AUX.PFV

'an haha o 'i si 'i hihim-**im** …
LOC then FUT PUNC INTS PUNC walk.PL-**PROG** …
'and it is now our thought that we in an intense way need to start canvassing [lit. walking] again [to inform people about a certain legislative matter]' (*#4*: 142)

There are also a few instances where a generic verb such as *juñ* 'do' or 'happen' conveys a path-motion meaning based on context. In the lines leading up to (42) below, Coyote is moving towards a group of turkeys. The verbal construction *'e-juñhim* 'doing (reflexive)' in (42) refers anaphorically to the prior motion activity. I determine which motion type to assign a generic verb on a case-by-case basis.

(42) B 'e-juñ-**him**-k gahu 'ai hega'i
thus 3.REFL-do-**PROG**-COMPL over.there reach 3SG

matp uḍ si ge'ecu towa,…
SBDR-3.AUX.PFV-ASSUM COP INTS oldest turkey…
'Having done that [moved towards the turkeys] he had approached the one who was the oldest turkey...' (*Coyote Turkey*: 7, 44, 70)

Finally, a few examples, like (17) repeated below as (43), also occur in the data:

(43) Gmhu him-a-c 'am cegito-**him** m-an-s has
over.there walk-EP-CONC LOC think-**PROG** SBDR-1SG.AUX.IPFV-DUB what.ABST
over there walking there thinking asking what

ñ-ju:k o ñ-do'ibia.
1SG.REFL-do.PFV-COMPL FUT 1SG.REFL-save.PFV
to do will save myself
'As I walked I was wondering how I could escape.' (*JD*: 254, 308)

Note that in the example above, the progressive verb *cegitohim* 'thinking' is preceded earlier in the sentence by the independent verb *himac*, which breaks down into *him* 'walk' and the suffix *-c*, which is a concursive marker indicating that two activities are occurring

contemporaneously. In other words, the concursive marker, in this case, signals that the "walking" and "thinking" are occurring at the same time. I also categorize cases like these as expressing path-motion.

Random motion

The next category of progressive constructions continues to convey translational motion but of a type that is more random in nature. If in its earlier stages the *-him* construction imparts the meaning of moving along a linear path from point A to point B while doing something over an extended period of time, then the hypothesized next step would be to express observable, physical motions that appear more random or distributed and involve a less clear path-type of motion. That is, these constructions continue to convey translational motion, but the notion of path is less evident. The meaning of the progressive construction gradually develops from a more concrete one towards a more abstract one.

In the first example of this set, (44), the meaning of *woihim* can be interpreted in a couple of different yet related ways. Dolores provides the gloss 'lie down one after another,' which suggests that a group of people are in the midst of moving into a prostrate position around a fire indicating motion. In Saxton et al. (1983), the entry for *wo'ihim,* specifically, can have the interpretation, 'move in a prone position' as well as 'roll over' (p. 65) indicating the subjects move around in some manner while lying down. The meaning of the verb stem, *wo'i-*,[28] otherwise carries the meaning 'in prone position' (Saxton et al., 1983, p. 65). In either interpretation, the addition of *-him* adds a motion meaning, either 'to move into a prostrate position' or 'to be moving around while in a prostrate position.' The example can also be read as expressing an ongoing physical action that occurs while the people are joined around the

[28]*Verb-* denotes a verb stem that requires a suffix of some type.

fire. The next example in (45) portrays a group of women and children who are in the process of moving around 'gathering themselves' during some period in the past. In example (46) the speaker talks about how a group of people were canvassing neighborhoods in an attempt at '*u'ahim* 'picking up' or collecting signatures in support of specific legislation. The motions in these cases are a further departure from the earliest meaning of *him* 'go, come, walk.'

(44) 'an we:big o 'i woi-**him** k gahu
LOC around FUT PUNC lie.down-**PROG** CONJ over.there
there around it would lie down one after the other and over there

wo t-nam.
FUT 1PL.REFL-meet.PFV
would meet ourselves
'…[we would be] lying around [the fire] and joined around it,' (*Antonio*: 54)

(45) T kiap 'e-hemapad-a-**him** g 'u'uwi c 'a'ali.
3.AUX.PFV yet 3.REFL-gather-EP-**PROG** ART women CONJ children
There had been gathering themselves the women and children
'They had been gathering themselves, the woman and children [during a time of war].' (*War*: 41)

(46) 'am masma 'ab hihi-him c 'an cem 'u'a-**him**
LOC in.a.certain.way LOC walk.PL-PROG CONJ LOC try.PART pick.up-**PROG**

g cecegig…
ART name.PL …
'…we were in that manner canvassing [lit. walking] around and trying to collect [lit. picking up or carrying] people's signatures.' (*#4*: 60)

If hearers subconsciously infer an abstract meaning that is included in the earlier path-motion meaning – one that connotes ongoing actions over a period of time at some location as opposed to movement along a path – then the contexts in which *-him* is utilized will also expand. The constructions as a whole develop to express actions that involve motion, the suffix continues to attach to dynamic predicates that describe an overt, ongoing activity at some location at reference time.

Localized motion

Localized motion describes contexts in which a subject clearly exerts energy in order to execute some type of visible, physical movement (e.g. the tapping of the fingers, the thrust of an arm, the movement of the jaw while yawning, singing…) while otherwise remaining in a single, restricted (self-imposed or otherwise), stationary location. In other words, while physical movement is observable, the subject in its entirety does not move *around* or undergo spatial displacement or translational motion while carrying out the physical activity.

In the first example, (47) below, the narrator describes a group of women who are perhaps standing in a single location facing each other and are in the midst of physically arguing with one another, an activity which entails the movement of the facial muscles, the lips, tongue, etc. The women may also be gesticulating; however, no indications suggest that the women are moving around while doing so but are likely in a single, stationary location. Next, example (48) describes the iterative action of a horse kicking someone. While it is clear that the horse must physically move its legs to execute a kicking action, there is again no suggestion that the horse, itself, undergoes any type of spatial displacement. The third example, (49), describes a scenario in which Hunter looks at what he is eating, his facial muscles and mouth executing the motions involved in mastication.

(47) 'E-'ebgio 'am 'e-ñe'owin-a-**him**…
3.REFL-frighten.PFV LOC RECIP-argue-EPEN-**PROG**…
they frighten themselves there they were talking back and forth…
'They frighten themselves, there they were arguing with one another,'
(*War*: 14)

(48) Kawyu 'o ñ-keikkaṣ-a-**him**.
horse 3.AUX.IPFV 1SG.OBJ-kick.ITER-EPEN-**PROG**
'The horse has been kicking me.' (Mathiot, n.d.a, p. 378)

(49) Ku-t-ṣ 'oya ce: g Mo'abḍam ma:s
CONN-3.AUX.PFV-QUOT then discover.PFV ART Hunter appear

g haiwañ jewalig 'am ko'i-**him**.
ART cow rotten.OBJ LOC eat-**PROG**
'and Hunter saw that he had been eating a rotten cow.' (*Buzzard*: 17)

The localized physical motion examples above illustrate that as the progressive construction generalizes and the meaning of *-him* becomes more abstract, it eventually extends to contexts including more subtle, non-translational motions, while at the same time, more emphasis is placed on the meaning component, 'ongoing action at some reference time.'

Non-motion

Finally, there are cases where no observable motion occurs. In these cases, no reading other than the abstract progressive meaning of ongoing action at reference time is possible for the suffix. This is illustrated first in (50) below which depicts a young boy who is repeatedly *koksim* 'sleeping' or napping while being carried on his older brother's back. In example (51), the narrator describes a period in time where the central figure of the story, a boy, *'e-'amicudahim* – internally concentrates or is literally 'knowing himself' until he turns into an eagle. The final example in (52) includes the process verb *'alicud* 'turn into a child' which combines with the progress marker to describe the development of an elderly man who is becoming childlike.

(50) Ṣ, koks-**im**, hekaj 'ab wulṣp
QUOT sleep.REP-**PROG** because LOC tie.an.OBJ.IPFV
It is said, he repeatedly slept for that reason he was tied

'ab 'e-si:s 'ab.
LOC 3.POSS-older.brother LOC
there his older brother at
'He was napping, because of him being tied to his older brother.' (*War*: 225-226)

(51) ku-ṣ 'am s-'e-'amicud-a-**him**,
CONN-QUOT LOC INTS-3.REFL-know-EP-**PROG**
'and so then he was concentrating intently [lit. knowing himself].' (*Boy*: 99)

(52) Hega'i keli 'e-'alicud-a-**him**.
that old.man 3.REFL-turn.into.a.child-EP-**PROG**
That old man is getting into his second childhood. (Mathiot, n.d.b, p. 406)

The examples above illustrate an expansion of *-him* to a meaning that is highly abstract – where motion or physical involvement in the action of any type is not present. Note also that in example (52) above, the process of becoming childlike cannot be said to occur at a specific location but instead refers to an internal, non-observable, process which illustrates a further extension of the contexts in which progressives can be used – one that is far from its original meaning of 'changing location along a path.' As the progressive marker continues to evolve from its hypothesized original concrete motion meaning of 'go, come, walk,' it broadens and generalizes, appearing in a wider range of contexts.

There are also a few instances where motion of some type, however, is metaphorically expressed, as in the construction in example (53). In this case, the narrator describes a time when someone observed that only the O'odham language was *wu:ṣñim* 'coming out' from some white person's mouth, metaphorically drawing the image of O'odham words streaming out in a path between someone's lips.

(53) …"g ḍ a'i O'odham-aj 'ab wu:ṣñ-**im** ciñ 'eḍ 'amjeḍ…
…3SG COP only O'odham-INSTR LOC come.out-**PROG** mouth in from…
'"…it was only O'odham coming from his mouth…' (#3: 144)

I categorize cases such as these as the non-motion type based on the definition of what "motion" is in this study – that the construction must describe some type of observable physical movement – either of the spatial displacement type which is related to the earlier

translational motion meanings of the suffix or of the non-translational type as in the localized motion cases described earlier.

I code each progressive construction in the data for the motion types described in this section and rank them from 1-4, with 1 being associated with the most motion, i.e. path-motion, which is also likely the earliest meaning based on the hypothesis that the verb *him* 'go, come, walk,' is the source for the O'odham progressive marker, and 4 marking examples that are completely devoid of any type of motion and are most abstract. The rankings are summarized in Table 30 below.

	Motion scale	**Motion Type Constructions**	**Description**	**Examples**
Most motion	1	Path-motion	The construction as a whole expresses a subject that is moving along a linear path while doing something; translational motion	*walking, running, going, coming… while talking, whistling, etc.*
	2	Random motion	The subject is randomly "moving (around) doing something" (Bybee et al., 1994, p. 133); translational motion	*moving around while picking up objects, gathering things, etc.*
	3	Localized motion	The subject carries out an observable physical activity while in a stationary location; non-translational motion	*typing, scratching, talking, eating, crying, coughing, breathing, etc.*
Least motion	4	Non-motion	The subject is in the midst of carrying out some non-motion activity.	*sitting, living, concentrating, improving, getting big etc.*

Table 30. Motion types and scale.

I rank the examples (38)-(41) above as 1 (whistling while going, going past/circling, walking, respectively), examples (44)-(46) as 2 (moving into lying position/lying down, gathering

selves together, picking up/collecting signatures), examples (47)-(49) as 3 (arguing, kicking, eating), and (50)-(52) as 4 (sleeping, knowing, turning into a child).

5.5.4 Specific versus generic predicate types

The possibility exists that as O'odham is used less often in the community, speakers may rely more readily on generic verbs with less specific meanings to express a wide range of progressive activities. As speakers have fewer chances to use a language, retrieving specific-meaning verbs may become increasingly more challenging, even if those verbs are stored in the speakers' mental lexicon. In more extreme cases, lexical items with specific meanings are also more likely to disappear first leaving behind words with comparatively more general meanings (e.g. Mithun, 1990, pp. 18-19). Based on this premise, the final variable that I code for is whether the progressive marker co-occurs with **specific** or **generic meaning predicates** in order to shed light on potential differences in collocation trends in the three different data periods. Examples of predicates with specific meanings are *run, see, scratch,* etc. I have already presented a number of cases in O'odham of specific meaning predicates throughout the present chapter (see (50)-(53), above, for example). Generic meaning predicates, on the other hand, carry meanings such as *do, make,* and *happen.* In O'odham, those verbs are *juñ* and *wua*.[29] These are used by a speaker to refer to or represent some other dynamic action or set of actions associated with an event. Example (42) from earlier is one instance of this. I present a few more below. First, in (54), the verb *wua* (or *wui-* when attached to a suffix) refers to doing some activity repeatedly or habitually and is used with

[29] Both forms usually co-occur with the particle *hab* which is frequently translated as 'thus' and denotes that some activity is carried out in a certain manner.

-him to reinforce the ongoing nature of the 'doing.' In this case, *wua* serves as a stand-in for the more specific, dynamic, physical motions Coyote executed in an attempt to expel the cholla cactus needles from his throat (e.g. coughing, hacking, etc.) that he had swallowed. In the second example (55), the narrator, Juan Dolores, uses the verb *juñ* which can mean 'happen,' 'do' or 'engage in an activity' to refer a set of the events that transpired after his parents left him at a boarding school. The dynamic actions that he refers to as 'happening' are the physical and verbal abuse students received during his years at the boarding school. In (56), *juñ* refers to activities or language use occurring in O'odham on the internet.

(54) Cum 'e-**wui-him**-c 'am a 'i huhug.
try.in.vain.PART 3.REFL-**do.HAB-PROG**-CONC there/then pass.away
'No matter what he tried to do [to get the cholla from his throat], that was the end [i.e. Coyote choked and died].' (*Coyote Quails*: 36, 65, 56-57)

(55) …ḍ mu'i ha'icu pi 'ap 'e-**juñ-him** 'am n-ta:gio
COP many thing(s) NEG good 3.REFL-**happen-PROG** LOC 1SG.OBJ-toward
there were many thing not good happening there in any way

ñ-ḍ 'aṣ pi has o ñ-dodda.
1SG.AUX.IPFV-RMT just nothing FUT 1SG.REFL-do.smth.to
I just could not help myself.
'…many bad things happened [corporal punishment, getting scolded, etc.] to me. I was completely helpless.' (*JD*: 261, 307)

(56) …'eḍa 'am hab 'aṣ 'i kia meḍ 'am hab 'e-**juñ-him**
yet LOC thus just PUNC still run.IPFV LOC thus 3.REFL-**happen-PROG**

heg 'eḍa World Wide Web.
3SG inside World Wide Web.
'…it [exposure and use of the O'odham language] is still progressing [lit. running] and happening inside the World Wide Web…' (*#3*: 352-353)

5.6 Results: Quantitative

In this section, I present the quantitative results for each of the categories that I describe and illustrate in §5.5. These are: a) subject control, b) situation described by the predicate (i.e.

dynamic or stative) to which the progressive suffix attaches, c) motion type expressed by entire progressive construction, and, lastly, d) specific versus generic verbs occurring with the progressive marker. As progressive constructions evolve, specific changes to the first three categories (a-c) are expected to occur. To recap, once the motion verb *him* 'go, come, walk' has grammaticalized into the earlier stages of the progressive marker, the progressive constructions as a whole often co-occur with agentive subjects, dynamic predicates, and meanings that convey translational motion/spatial displacement along a linear path. As progressive constructions generalize in meaning and expand, they hypothetically increasingly co-occur over time with predicates involving less subject control or agency, as well as stative predicates, and constructions that express lesser degrees of motion (see §§5.1 and 5.5) or, at a minimum, their distributions will remain stable if some of the features are already present in the historical data. If the patterns are not as expected, we can ask whether or not decreased language use is a possible factor. Furthermore, we must also consider whether or not any clues are present which suggest that language contact has promoted any unexpected types of changes. At this point, no clues point in this direction. Finally, the last category, (d), assesses whether or not generic verbs co-occur more frequently over time with progressive constructions than verbs with semantically specific meanings. If this is the case, this could be an indicator of changes in the lexical domain as O'odham is gradually used less often and the community experiences more advanced stages of language shift.

For the results, first, I present token frequencies in percentages for each category across the three data periods. The goal in doing so is to provide an overall picture of how distributions have or have not changed over time and if it is possible to detect any grammaticalization trends in the data. For example, if less-agentive subjects significantly

increase in frequency as time passes, this would be one possible indicator that the construction continues to extend towards a more grammatical status. Finally, I conduct Chi-Square tests to determine if a statistically significant relationship exists between the different progressive construction categories noted above and the three data periods. The null hypothesis of the Chi-Square test in this case is that no statistically significant relationship exists between the progressive aspect categories and the data periods during which the progressive constructions are produced (i.e. the variables are independent from one another).

5.6.1 Subject control

As progressive constructions continue to grammaticalize, they may gradually extend towards less agentive subjects. Figure 6 below shows the findings for subject agentivity based on activity type (see Table 28 in §5.5.1) for the historical, midcentury, and modern data. The percentages represent the number of progressive construction tokens that display the different activity types for each separate data period. These are organized based on the ranking scale presented in Table 28 (1=*most A[gent]-like* to 6=*least A-like*, i.e. patients and experiencers). The category of Inherent Properties, Dispositions (ranking of 3) does not occur in any of the three periods.

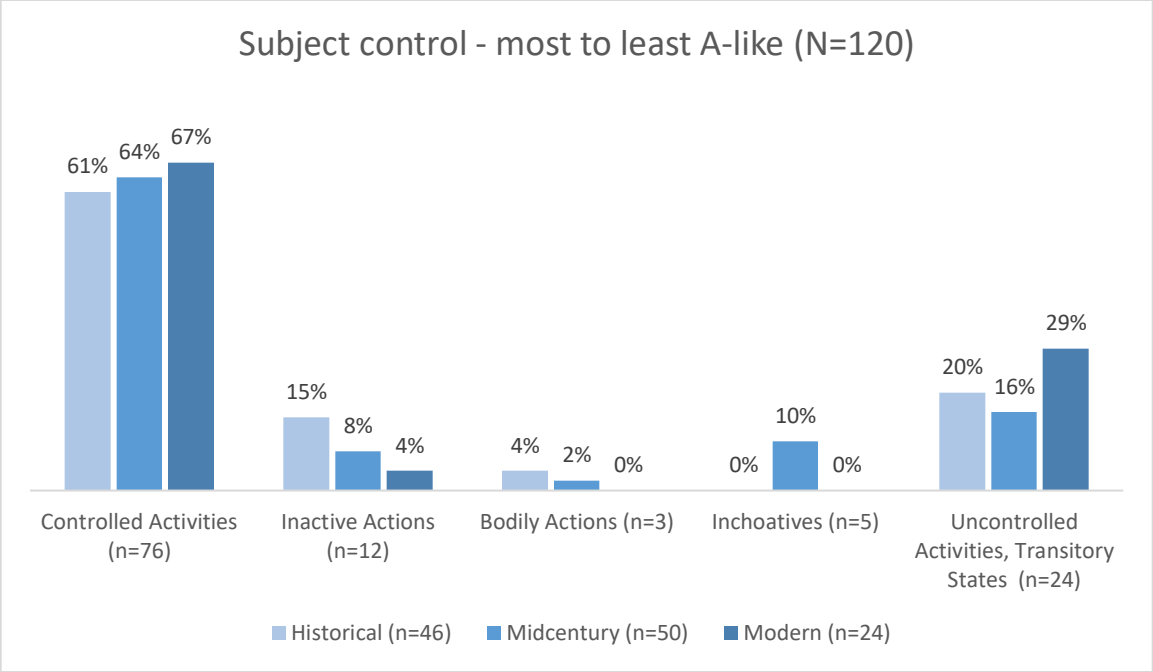


Figure 6. Percentages of activity types associated with degree of subject agentivity.

Subjects identified in progressive constructions are most frequently highly agentive across the historical (61%), midcentury (64%), and modern data (67%), occurring in the context of Controlled Activities, a ranking of 1, with slight increases of 3% between the historical and midcentury data, and the midcentury and modern data. The intermediary categories, Inactive Actions, Inherent Properties, Dispositions, Bodily Actions and Inchoatives, are the least frequent overall, with few tokens overall. The constructions falling into the category of Uncontrolled Activities, Transitory States with a ranking of 6 (=least agentive)) involving patients and experiencers are the second most frequent across the data periods, 20%, 16%, and 20% for the historical, midcentury and modern data, respectively. Between the historical and modern data, there is an increase of 9%, with the midcentury data having the lowest number of occurrences.

In sum, all three groups of data show similar trends with a majority of progressive constructions having subjects that are the most Agent-like, the least number of subjects

falling somewhere in the middle, and then a relatively larger number of subjects that fall towards the least Agent-like end of the spectrum.

Finally, I conduct a Chi-Square test for the categories of Controlled Activities, Inactive Actions, and Uncontrolled Activities (Agentivity rankings of 1, 2, and 6) – that is, excluding categories that have frequency counts of 0. The results of the test are as follows:

The Chi-Square is 3.3577. The *p*-value is .499849. The result is *not* significant at $p < .05$.

The *p*-value supports the null hypothesis; there is no significant relationship between degree of agentivity of the subjects and the three data periods. In other words, the differences between the three time periods are not significant. In this regard, it appears that progressive constructions are more-or less stable over time, changing very little.

5.6.2 Situations described by the predicate

During the earlier stages of grammaticalization, progressive constructions co-occur with dynamic predicates and in more advanced stages they may also co-occur with stative predicates. The percentage of dynamic and stative predicates occurring with the progressive marker *-him/-im* in the data for all three periods is represented in Figure 7.

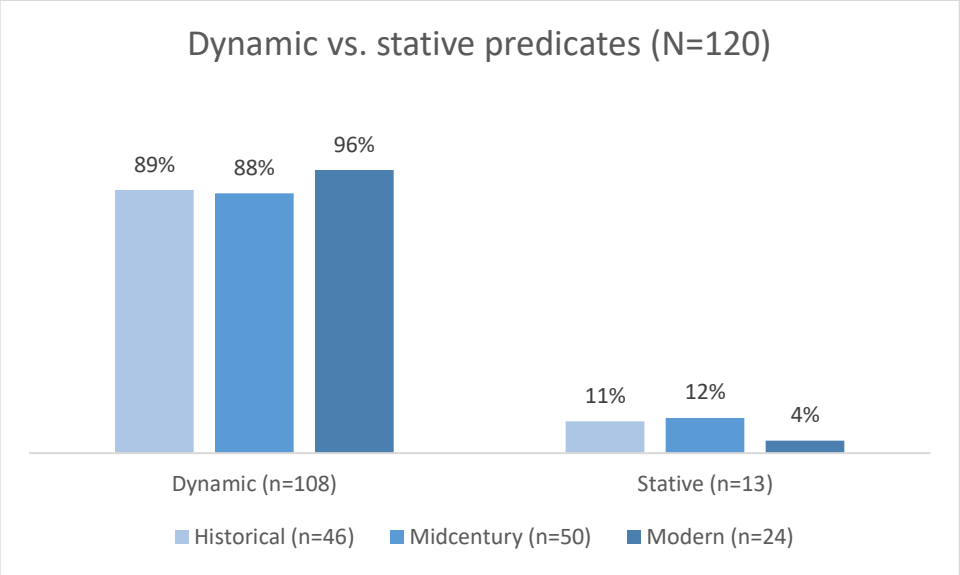


Figure 7. Percentages for dynamic vs. stative verbs occurring with *-him/-im*.

Dynamic verbs occur with *-him* overwhelmingly in the historical (89%), midcentury (88%), and modern data (96%). Stative verbs occur with low frequency (11%), (12%), and (4%), respectively, in all three periods. For both verb types, the differences between the three data pools indicate similar distributions.

Finally, a Chi-Square statistical test of the predicate types shows no significant difference between any of the three periods; the relationship between predicate type and data periods is statistically insignificant in favor of the null hypothesis as the results below show.

The Chi-Square statistic is 1.1683. The p-value is .557586. The result is not significant at $p < .05$.

Both the distributional patterns and Chi-Square test indicate that the predicate types that tend to co-occur with progressive constructions are stable with little significant change over time.

5.6.3 Motion types

I hypothesized earlier in §5.5.3 that progressive constructions as a whole likely start off expressing path-motion meanings and at the other end of the grammaticalization cline they

may eventually express less dynamic motions or even static, non-observable states. Based on this premise, the third set of criteria that I code for and analyze include the following motion types: Path-motion (translational), Random Motion (translational), Localize Motion (non-translational), and Non-motion. Figure 3 shows the frequency in percentages of the particular motion-type constructions in the data. These are ranked from 1 to 4 (1=most motion, 4=least motion).

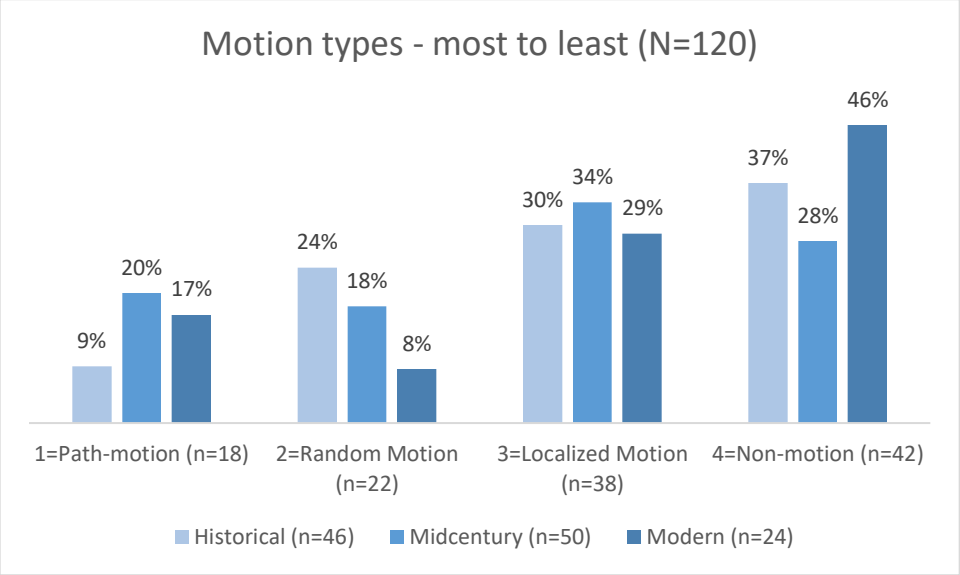


Figure 8. Percentages for motion types expressed by progressive constructions.

The breakdown for the four different motion types for each data period is as follows: progressive constructions expressing Path-motion (assigned a ranking of 1=most motion) on the motion scale – are the least frequent in the historical data (9% of the total tokens), followed by Random Motion (24%) and Localized Motion (30%). For the historical data, the Non-motion type (4=least motion) occurs with the highest frequency (37%). For both the midcentury and modern data, Random Motion is the least frequent (18% and 8%, respectively), followed by Path-motion constructions for both data periods (20% and 17%, respectively). Looked at another way, the higher motion constructions (Path-motion and

Random Motion) are the least frequent across the data periods, and the lower motion type constructions (Localized and Non-motion) are the most frequent.

The additional line graph in Figure 9 below highlights trends in a slightly different way, with each line representing the four different motion types and the points on each line indicating the data periods.

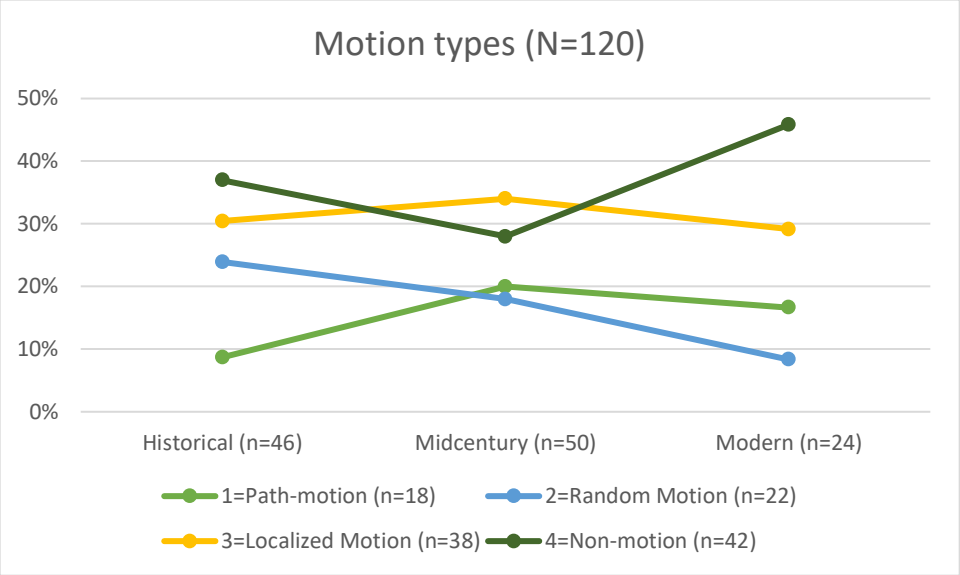


Figure 9. Trends for each of the three data periods for Motion Types.

The line graph shows that the highest percentage of tokens (46%) occurs with Non-motion (4= least motion) progressive constructions in the modern data, suggesting a general upward trend of constructions falling under the non-motion categories; as progressive constructions advance and generalize in meaning, this is the expected pattern. The lowest percentage of tokens occurs with Path-motion constructions in the historical data (9%). The line graph also reiterates the earlier point that, overall, Localized and Non-motion types (the lower motion categories) are the most frequent, and Path-motion and Random Motion (the highest motion categories) are the least frequent across the data periods.

The results of the Chi-Square test are as follows:

The Chi-Square statistic is 5.9003. The p-value is .434445. The result is *not* significant at p <.05.

Although there are some differences for motion types in each of the three data periods – particularly the higher percentage of Non-motion constructions in the modern data – a Chi-Square test shows no statistically significant difference between data periods.

5.6.4 Specific versus generic meaning predicates

Lastly, Figure 10 below summarizes the distributions for specific versus generic verbs co-occurring with the progressive marker.

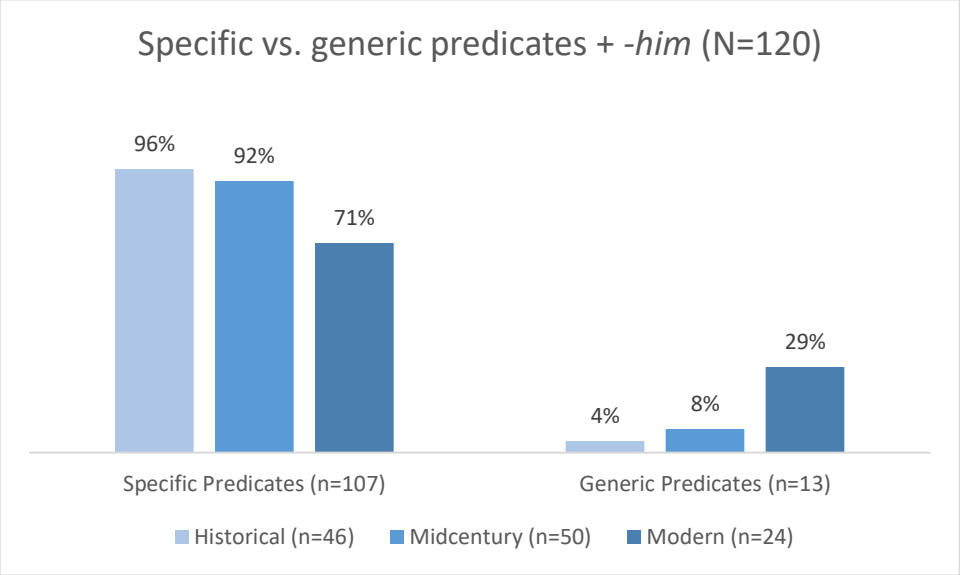


Figure 10. Percentages for specific and generic predicates occurring w/ *-him*.

When all of the data across all three periods are examined, specific-meaning predicates occur in the progressive aspect most frequently in the historical (96%), midcentury (92%), and modern data (71%) and predicates with generic meanings occur with the lowest frequencies, 4%, 8% and 29%, respectively. When generic predicates in particular are compared across the data periods, they clearly occur with the greatest frequency in the

modern data (29%) compared to the other two periods, at 4% and 8% for the historical and midcentury data, respectively.

In order to determine if the relationship between specific and generic predicate types and data period is statistically significant, I implement a Chi-Square test.

The Chi-Square statistic is 10.7694. The p-value is .004586. The result is significant at $p < .05$.

The results of the test indicate that the difference in distributions is statistically significant with a *p*-value less than .05, rejecting the null hypothesis. The combined results may suggest that speakers tend to rely more often on generic verbs as they use the language in fewer domains.

Further examination of specific versus generic-meaning predicates

The findings in the preceding section indicate that the increase in generic-meaning predicates is statistically significant in the context of progressive constructions. In terms of percentages, this change in distribution is most notable for the modern data when compared to the other two periods (see Figure 10 above). In order to make a further determination if this pattern exists in the language as a whole and not solely in relation to progressives, I carry out an additional sample study. I analyze a sample of verbs outside of the progressive aspect, specifically in the historical and modern data. If the findings reflect a parallel trend when the historical and modern data are compared, this would suggest that current-day speakers are, in fact, displaying an overall general preference for generic-meaning verbs (*juñ, wua* 'do, happen,' sometimes 'make'), but that the pattern is not specific to progressive constructions. In other words, it cannot be said that there is a change, specifically, in collocations between semantically specific versus generic verbs and the progressive marker.

In order to carry out this additional study, I selected a verb sample for analysis using the FLEx search feature (see discussion of Image 3 in §4.3.4), which allows one to isolate all lexical items that have been tagged with the label '(v)erb.' FLEx then provides a randomized list of all verbs occurring in the texts. Once the list was procured, I simply went down the verb list in the random order produced by FLEx and took note of whether the verb fell into the specific or generic category. Only copulas were omitted. For the historical and modern data, I coded 100 tokens for each period, totaling 200 tokens. All speakers are represented in the sample across the different texts (see Table 18 and Table 21 for the information on the historical speakers and texts; see Table 20 and Table 23 for the modern materials). This was done in order to eliminate, as much as possible, patterns that may be speaker specific, which could potentially skew the results in one direction or another. Figure 11 below includes the findings in percentages.

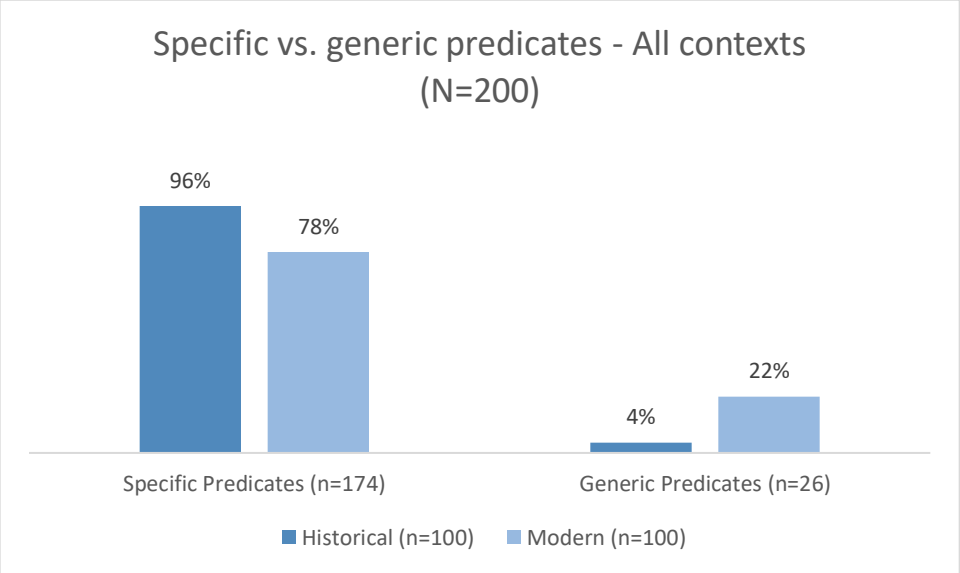


Figure 11. Percentages for specific and generic predicates *overall.*

The figure above shows that while specific-meaning predicates are used more often in both the historical and modern data at 96% and 78%, respectively, the percentage of generic-

meaning predicates increases over time, up from 4% in the historical data to 22% in the modern data.

The results of the Chi-Square test indicate that the difference between periods is statistically significant.

The Chi-Square statistic is 14.3236. The *p*-value is .000154. The result is significant at $p < .05$.

The findings provide evidence that modern speakers utilize generic verbs more often and more generally in O'odham – regardless of whether or not the verb is in the progressive aspect. The pattern is not likely due to a change that is limited to the types of verbs that collocate with the progressive marker. Instead, the change appears to be taking place in the language as a whole, as speakers both hear and use the language on a less frequent basis.

5.7 Discussion

In this final section, I discuss what the findings in §5.6 suggest regarding the impact language shift has had on the grammaticalization of progressives in O'odham over the approximately 100 years that the data spans. Overall, it appears that language shift has had little impact. The variables most often remain stable and in one case show a possible (albeit tentative) trend that one would predict based on previous cross-linguistic grammaticalization studies. The area of language that the shifting linguistic landscape may have impacted is the lexical domain – although this change does not appear to be limited to progressive constructions. I did not find any suggestive evidence that contact shapes any of the changes that I identified.

First, the findings for the first variable, subject agentivity, show little signs of change over time when I examine the historical, midcentury, and modern data. Subjects more often

than not tend to be highly agentive, but non-agentive subjects such as patient and experiencer are possible in all three data periods. The same pattern is true of predicate types, dynamic versus stative, occurring with the progressive marker, *-him*; the constructions most commonly occur with dynamic predicates, while stative predicates are also present but at lower percentage rates. When I consider motion types expressed by progressive constructions as a whole across the data, the higher motion types, Motion-path and Random Motion, are the least frequent and the lower motion types, Localized Motion and Non-motion, are the most frequent across the data.

In all three cases (i.e. subject agentivity, predicate type, and motion type) the results of the Chi-Square test confirm that the relationships between variables are not statistically significant.

Regarding the motion-type category, however, Non-motion constructions occur at the highest percentage rate in the modern data. This may suggest (tentatively at this point) a grammaticalization advancement trend from higher motion type constructions towards Non-motion type progressive constructions over time. More data may support or refute this tentative finding.

The combined findings for all of the categories might also suggest that motion type is a better predictor of the grammaticalization of the progressive aspect – especially if the source of the progressive marker is a motion verb as in O'odham – while subject agentivity and predicate type are weaker predicators of change. These various parameters are worth exploring further in future studies.

With regard to the severe language shift that has transpired among the O'odham community, as stated earlier, the data show little indication of change in terms of the fabric of

the constructions overall. The structures remain stable, with the possible exception of an increase in the usage of Non-motion type constructions. The latter detail suggests that, despite the scenario of language shift, the constructions continue to evolve in a manner that would be expected, extending towards more abstract meanings. That is, language shift has not disrupted the hypothesized grammaticalization trends with regard to motion type.

Separate from the variables above, one area where a possible change related to decreased domains of language use does emerge is in the frequency of predicates with specific meanings (e.g. *walk*, *whistle*, *groan*, etc.) versus those with generic meanings, in this case *juñ* and *wua*[30] 'do,' 'make,' or 'happen.' This change, however, is not restricted to progressive constructions; evidence points in this direction both within the context of progressive constructions as well as across verb tokens in the language more generally (i.e. those not in the progressive aspect). Regarding the former, comparatively, semantically specific verbs decrease in frequency in the modern data (down to 71% compared to 96% and 92% in the historical and midcentury data, respectively), while the frequency of generic verbs increases in the modern data (29% compared 4% in the historical and 8% in the midcentury data). Furthermore, the result of the Chi-Square test indicates a statistically significant relationship between the variables of predicate type and data period. When verbs are examined more generally (i.e. not just those in the progressive), similar statistically significant trends are evident across time when the historical and modern data are compared. In the historical data, 4% of the verbs are of the generic type, and in the modern data 22% are.

[30] *Wua* also carries an iterative or habitual sense.

These combined finding suggests that while severe language shift has not appeared to impact the components that are hypothesized to be integral to progressive constructions, speakers today may be showing an *overall* preference for lexical items with less specific meanings, as has been noted in other language communities experiencing language shift (e.g. Mithun, 1990).

Generic verbs arguably carry a lighter cognitive load in that they can be used to describe a wide range of activities versus verbs with more specific meanings which occur in a narrower range of contexts. This may be one possible explanation for why the fluent speakers of the late-shift period – a period during which O'odham is used infrequently and which the modern data represents – may show a preference for these types of verbs given the limited opportunities to use the language. In the most extreme interpretation, it may also suggest a decrease in vocabulary; it is more likely that verbs with specific meanings would fade away before verbs with generic meanings. Finally, even if speakers have the more specific-meaning verbs stored in their mental lexicon but have few opportunities to use the language, recalling generic verbs is an easier task.

One caveat is in order, however. The preference for generic verbs may also be attributed to the different discourse types included in the data. While the historical and midcentury data are based on narratives (traditional, personal, etc.), the YouTube recordings in a majority of cases involve long answer or interview formats (described in §4.1.3). These formats tend to lend themselves to different topics than a traditional or personal narrative would, at least in the data that I consider here. The long answer and interview formats usually center around current events at the time of recording – discussions about legislation, community gatherings, topics of controversary or concern in the community – whereas

traditional or personal stories tend to involve overt, physical activities associated with war, daily chores, etc. It may be the case that the discourse types themselves influence the kinds of predicates that speakers tend to use. A study involving more uniform discourse types might produce different or similar results.

Chapter Six: Demonstratives

6.0 Introduction

As with the previous study on progressive aspect in Chapter 5, the purpose of examining demonstratives is to determine how they have changed, if at all, during a period of language shift among speakers of O'odham over the last century. This study differs, however, in that I do not attempt to provide evidence for their hypothesized source and their development into grammatical devices. Instead, I assume that demonstratives start off as grammatical and then later expand to serve a multitude of different functions (see further discussion in §6.1.4). My primary concern in this chapter is to explore this expansion of demonstratives. Diessel (1999) proposes a number of hypotheses for the various trajectories of change that demonstratives may undergo and what those changes are expected to look like based on evidence in the author's typological study. Drawing on these hypotheses, we can also examine to what extent the changes, if any, in the historical, midcentury, and modern data align with those expectations. Furthermore, what do the findings reveal about the impact or lack of impact the shifting linguistic landscape has had on the evolution of demonstratives in the language, as speakers increasingly use English in a majority of domains in the US? For example, the data may reveal changes that have been promoted by contact, and/or they may show signs that grammaticalization has been disrupted in some manner. Overall, very little discussion about demonstratives and grammaticalization exists in the O'odham literature. In addition to examining language change in the context of a language shift scenario, the present study also provides a starting point for research on this topic. Furthermore, few studies of O'odham describe the range of possible functions that demonstratives may serve, particularly within

authentic O'odham discourse as this study does – an understanding of which is crucial for an in-depth analysis of patterns of change over time.

Diessel (1999) argues that demonstratives show various possible grammaticalization pathways or clines that they may follow, which can gradually lead to a demonstrative's development into third-person pronouns, definite articles, connectives, determinatives, among many other possibilities (see §6.1.4 for more details). Furthermore, Diessel suggests that the path down which a demonstrative follows largely depends on the syntactic context or construction in which it occurs. Demonstratives also hold a range of pragmatic functions, the most basic of which is to "orient the hearer outside of discourse in the surrounding situation" (p. 2), as well as to assist the hearer in organizing referents within the discourse text itself (see §6.1.3). How demonstratives are used in discourse also contributes to their potential grammatical development.

In O'odham, the primary demonstratives are *hegai/heg* 'that', *hegam* 'those', and *'i:da/'i:d* 'this', *'idam* 'these', illustrating a distinction between distal and proximal, respectively, as well as number. There is evidence that demonstratives have, in fact, expanded to carry varying grammatical functions in O'odham. For example, the distal demonstrative series (*hegai/heg* 'that', *hegam* 'those') often convey third-person pronoun meanings (Zepeda, 1983, p. 159). Both distal and proximal demonstratives also show evidence of having become what Diessel (1999) labels 'determinatives' (p. 135; based on Quirk, Greenbaum, Leech, & Svartvik, 1972, p. 217). These are demonstratives that obligatorily mark the nominal head of a relative clause or serve as the head of a relative clause itself and are also entirely devoid of any other discourse pragmatic meanings associated with earlier stages of development. The proximal demonstratives, *'i:da,/'i:d* 'this'

and *'idam* 'these', have been discussed very little in terms of their grammatical functions, a gap that the current study addresses.

I begin in §6.1 by providing an overview of how demonstratives have cross-linguistically been described in the literature from a functional, usage-based standpoint. I also discuss the relevant grammaticalization pathways that demonstratives are hypothesized to potentially take, as described primarily in Diessel (1999) and Himmelmann (1996). In §6.2, I briefly discuss how previous scholars have described demonstratives in O'odham in particular. Sections 5.4 and 6.4 address methodological concerns for the present study, namely token selection and coding procedures for identifying demonstrative types and functions in the data. Section 6.5 includes the results of my descriptive analysis of demonstratives in which I present and discuss numerous examples of demonstratives in O'odham based on the coding guidelines presented in §6.4. In §6.6, I present numerical distributions and statistical analyses of various demonstrative patterns in the historical, midcentury, and modern data. Finally, in §6.7, I discuss my interpretation of the findings with regard to how language shift has or has not impacted the evolution of demonstratives.

6.1 Background on demonstratives

In what follows, I present an overview of how demonstratives are addressed in the literature. This includes a definition of and an introduction to their most basic functions (§6.1.1) followed by a more in depth description of the syntax of demonstrative constructions (§6.1.2), as well as an overview of their pragmatic functions (§6.1.3). Finally, I discuss some previous findings on the different grammaticalization paths that demonstratives have been hypothesized to take over time (§6.1.4).

6.1.1 Defining demonstratives

Demonstratives such as *that* and *this* in English are, in their earliest stages, deictic expressions or reference devices used as a type of "verbal pointing" (Halliday & Hasan, 1976, p. 57). The speaker uses a demonstrative in order to situate the referent on a proximity scale in either space or time. According to Diessel (1999), nearly all languages *at a minimum* make the relative distinction between **distal** (far from the speaker) and **proximal** (near the speaker)[31] (p. 2). This is true of O'odham, as well. Demonstratives are additionally used as tools for organizing discourse (see §6.1.3).

6.1.2 Basic syntactic formations

Cross-linguistically, demonstrative constructions typically take on two basic syntactic patterns, either pronominal or adnominal (see Diessel, 1999; Himmelmann, 1996). **Pronominal demonstratives** occur in isolation as independent pronouns or verbal arguments in their own right and, as a result, are also frequently referred to as **demonstrative pronouns** in the literature (e.g. Diessel, 1999; Dixon, 2003; Himmelmann, 1996; Halliday & Hasan, 1976; Lyons, 1979). This is illustrated in example (1) from English.

(1) **That** was really scary [referring to a movie the speaker just saw].

In the example above, *that* occurs independently and serves as a pronoun (or NP) which represents the movie the speaker is referring to.

Demonstratives can also be **adnominal** where the demonstrative is in juxtaposition to a noun, together comprising a NP. Demonstratives in these instances are also sometimes

[31]Diessel (1999) also includes demonstrative adverbs and demonstrative identifiers in his discussion. These are not addressed in this study. The reader is referred to Diessel, CH 4 for a description of other demonstrative types.

referred to as **demonstrative adjectives** (e.g. Dixon, 2003; Himmelmann, 1996), as well as **demonstrative determiners** (e.g. Diessel, 1999; Halliday & Hasan, 1976). Example (2) illustrates an adnominal construction.

(2) **This dog** is so soft!

The demonstrative *this* in (2) is followed by its referent, *dog*, which in this case together comprise the subject NP of the sentence.

6.1.3 Discourse functions

Demonstratives also serve a number of important discourse pragmatic functions. The **exophoric function** of demonstratives draws attention to an entity in the physical surroundings. Demonstratives also play various **endophoric functions** within discourse text itself. **Discourse deictic demonstratives** highlight entire segments of surrounding discourse, **tracking demonstratives**, as the name implies, help the hearer keep track of previously mentioned referents, and **recognitional demonstratives** activate shared knowledge between the interlocutors. The latter three basic endophoric categories figure more prominently in discussions concerning the development of more grammaticalized uses of demonstratives given that they shape, to some extent, the possible grammatical outcomes (see §6.1.4) while exophoric demonstratives do not; however, I also address exophoric demonstratives in some compacity given that no extensive discourse-based study of demonstratives in O'odham currently exists.

Exophoric demonstratives

Halliday and Hasan (1976) distinguish between exophoric and endophoric demonstrative functions. Exophoric demonstratives are "situational" in that they refer "to a thing as

identified in the speech situation" (p. 32). The function of the exophoric type is to "focus the hearer's attention on persons, objects or locations in the outside world" (Diessel, 1999, p. 6). I will only discuss two types of exophoric demonstratives, namely **gestural** and ***Deixis am Phantasma*** in order to illustrate a couple of the different ways that these may manifest (for other types, the reader is directed to Diessel, 1999, pp. 94-95).

Features of gestural demonstratives

Exophoric demonstratives of the gestural type (Fillmore, 1997, p. 63; Levinson, 1983, pp. 65-66) are set apart from endophoric demonstratives by the fact that a) the speaker is at the deictic center, b) the demonstratives indicate relative physical distance[32] (e.g. proximal versus distal in English), and c) they can though are not required to be accompanied by a pointing gesture (Diessel, 1999, p. 94; also see Himmelmann, 1996, pp. 224-225). The set of examples in (3) below illustrates the exophoric use with both pronominal and adnominal demonstratives.

(3) a. **Those** look beautiful! [speaker pointing to a bouquet of flowers across the room]
b. **Those flowers** look beautiful! [speaker pointing]

From the vantage point of the speaker, the flowers physically exist in the person's physical surroundings, the speaker points to the referent, the flowers, in order to draw the hearer's attention to them; hence, it is 'gestural,' although the speaker could have also refrained from pointing.

[32] There are arguably some exceptions; see discussion in Diessel, 1999, pp. 38-39.

Features of 'Deixis am Phantasma'

There are also cases which Bühler (1934) first referred to as *Deixis am Phantasma* (or 'Deixis in the Imagination'). In these instances, the perspective is shifted from the viewpoint of the speaker to the perspective of a person or other animate entity within a narrative or other type of discourse. Himmelmann (1996) provides the following example in which the speaker describes a scene playing on a video from Chafe's (1980) *Pear Stories*:

(4) And he's…you see a scene where he's… coming on his bicycle **this** way.
(Himmelmann, 1996, p. 222)

In this case, the speaker treats the physical space in front of them as if the bicycle were, in reality, moving "this way," possibly even gesticulating to show the imaginary path down which the bicyclist is moving. This strategy may be employed in any discourse setting where the speaker treats the imagined space and the referents within it as if they truly existed in the speaker's physical surroundings.

Endophoric demonstratives

Any demonstrative that does not indicate a referent in the (real or imagined) surrounding, physical situation is classified as endophoric. Endophoric demonstratives allude to referents that are mentioned within the verbal utterance itself. They refer to something that is "textual" rather than "situational" (Halliday & Hasan, 1976, p. 33) and carry a range of pragmatic functions within discourse. The most basic categories are discourse deixis, tracking (anaphora), and recognitional. These functions may later expand to more grammaticalized, endophoric uses (see §6.1.4).

Features of the discourse deictic function

The first subcategory of endophoric demonstratives are the discourse deictic type. In these cases, the demonstrative is employed to refer to "propositions or events" (Himmelmann, 1996, p. 224). They "focus the hearer's attention on aspects of meaning expressed by a clause, a sentence, a paragraph, or an entire story" (Diessel, 1999, p. 101). Furthermore, they create "an overt link between two discourse units" (p. 102). The proposition occurs in an immediately adjacent discourse segment (Himmelmann, 1996, p. 224). The demonstrative may refer to the proposition either anaphorically or cataphorically, pointing backwards or forwards in the discourse. Anaphoric reference can be conveyed with either distal or proximal demonstratives while proximal demonstratives are used for cataphoric reference; this appears to be true of English at least on which Diessel basis his discussion (Diessel, 1999, p. 103; based on Fillmore, 1997; Halliday & Hasan, 1976). The function of the discourse deictic demonstrative is to focus the hearer's attention on a particular discourse segment, whether it be a sentence, paragraph, entire story, etc. Diessel illustrates this with the following example:

(5) a. A: I've heard you will move to Hawaii?
B: Who told you **that**?

b. A: Listen to **this**: John will move to Hawaii. (Diessel, 1999, p. 102; modified)

In (5)a Speaker B uses the distal demonstrative *that* to refer anaphorically to the entirety of Speaker A's proposition, not to a specific nominal referent. In (5)b the proximal demonstrative *this* is used to cataphorically anticipate and place focus on the information in the clause that follows it. In both cases, note that the demonstratives are used pronominally. Based on findings in Himmelmann (1996), the discourse deictic use occurs predominantly

with pronominal demonstratives. The hypothesis is "that discourse deixis is *the* typical use for demonstrative *pronouns…*" (p. 225; original emphasis), although one can imagine cases where discourse deixis with adnominal demonstratives is also possible as in a modified version of example (5) from above shown in (6) below.

(6) a. A: I've heard you will move to Hawaii?
B: Where did you hear **that nonsense**?

b. A: Listen to **this nonsense**: John's leaving us and moving away to Hawaii.

In (6)a above Speaker B uses the adnominal phrase *that nonsense* to refer anaphorically to Speaker A's entire proposition. In (6)b the speaker might be upset that their good friend John is moving far away and uses *this nonsense* to cataphorically highlight the bad news expressed in the subsequent sentence.

Once the referent – in this case an entire proposition – is highlighted, it typically does not continue as salient in the discourse. The referent is usually mentioned once and then the discourse continues without the referent being mentioned again. Diessel argues that demonstratives in these cases "function to provide a thematic link between two propositions (or speech acts) at one particular point in the progressing discourse" (Diessel, 1999, p. 102). They are "similar to a sentence connective in that it functions to combine two chunks of discourse" (p. 102).

Features of the tracking function

Tracking demonstratives (often referred to as 'anaphoric' demonstratives) point back to and are co-referential with a non-topical referent represented by a noun or noun phrase mentioned at some earlier point in the discourse. They indicate a discourse participant that is "somewhat unexpected and not currently in the focus of attention" (Diessel, 1999, p. 96). They often

refer to '**semi-active' referents** that are "not at the center of the hearer's consciousness at the current point in the discourse" (Croft, 2019, p. 68 based on Chafe, 1987; see more on referent status in §6.1.4). Tracking demonstratives may also anaphorically refer to a non-prominent referent in order to mark a topic shift in the discourse. Demonstratives that do so often occur as a second mention (Diessel, 1999; based on the studies of Cyr, 1993a, 1993b, 1996; Himmelmann, 1996; Lichtenberk, 1996), although they may occur at other points, as well, based on my own findings (see §6.5.1).

To begin with, the example below in (7) from the *Pear Stories* illustrates the use of tracking demonstratives to refer to a semi-active referent.

(7) Something that I noticed about the /movie/ particularly unique was that
the colors . . were (.35)
just (.5)
very strange. (.6 (.2)
Like (.3))
the green was a (2.2)
inordinately bright green, (.55)
for the pears, (.4 . . and (.25))
these colors just seemed a little (.5)
kind of bold, almost to the point of (1.15)
being artificial. (Himmelmann, 1996, p. 227; modified)

In the example above, the speaker talks about 'the colors' and then continues on to describe specific colors (e.g. green) and then briefly brings up 'the pears.' At this point, 'the colors' is probably no longer the central referent in the hearer's consciousness and has become semi-active as a result of the introduction of the other referents (see §6.4.2 for more on how referent status is operationalized). Once the speaker returns to the semi-active referent, it is tracked with the adnominal construction, 'these colors' (Croft, 2019, p. 73).

The next constructed example illustrates a case in which a tracking demonstrative indicates a topic shift:

(8) The other day we were with Phillip, and he was telling us a story about **a man** who was a shapeshifter. **This man** could turn himself into a bird. In order to do so, **he** would concentrate with great focus…

In the above example, the speaker first introduces *Phillip* to the discourse, and then we are told what Phillip is doing – telling a story. The speaker also introduces a new participant into the discourse at this juncture with an indefinite noun phrase, *a man*. At this point, the hearer does not know for certain whether or not the newly introduced referent will become the main participant in the discourse segment. For example, it could be the case that the speaker continues to talk about Phillip. However, the speaker mentions the referent, *man,* for a second time, but in this case the speaker employs an adnominal construction, *this man.* The demonstrative tracks the previously introduced referent, *a man,* and it also signals a topic shift to the referent that has not, up to this point, been established as central to the discourse. It is the 'unexpected' referent in this sense. It isn't until the third mention of the referent that the speaker employs the third-person pronoun, *he,* fully establishing the referent, *man,* as prominent in the discourse.

Finally, there is additional evidence that tracking demonstratives may *reactivate* a non-topical referent that is introduced "at some distance in the preceding discourse" (Diessel, 1999, p. 98; referencing Lichtenberk, 1988). The referent may have appeared in the discourse recently enough that it maintains semi-active status in the speaker's consciousness but exists somewhat on the periphery perhaps due to the presence of other participants in the intervening clauses. However, evidence points to more extreme cases, as well, in which the tracking demonstrative reactivates a non-topical discourse participant that occurred far enough back that the referent has likely become 'inactive' in the hearer's consciousness (see §6.1.4 for more on referent accessibility). At some point, the previously stated referent may

become relevant to the discourse once more, and the speaker employs a demonstrative to retrigger the referent (see examples (67) and (68) in O'odham in §6.5.1).

What is true in all cases – whether marking a topic shift or reactivating a referent from earlier in the discourse – is that the referents are not what "the hearer would expect in this context (i.e. the most topical NP)" (Diessel, 1999, p. 99).

Finally, in terms of demonstrative type, Himmelmann (1996) tentatively suggests that proximal demonstratives tend to be employed as tracking devices, although the author also points out that there appears to be "significant cross-linguistic variation with respect to these observations…" (p. 226). Diessel (1999) simply states that languages tend to prefer one demonstrative type over another for the tracking function (p. 99). Regarding preferences for pronominal versus adnominal constructions, both have been found to be used for tracking, although the results of Himmelmann's (1996) study suggest that there may be a preference for adnominal constructions (see pp. 252-254).

Features of the recognitional function

Finally, recognitional demonstratives were first described in significant detail by Himmelmann (1996), although prior to this, the function had been noted tangentially by others (e.g. Auer, 1984, p. 636; Goddard, 1983, p. 54; Wilkins, 1989, p. 121). According to Himmelmann, recognitional demonstratives tend to mark a referent that is new to the discourse, but at the same time is assumed to be known and identifiable to both the speaker and hearer based on shared knowledge or experience. In Croft's (2019) terms, the referent carries 'inactive' status given that it is new to discourse and the hearer does not yet hold the referent in their consciousness (see §6.1.4). In other words, the referent is "*discourse new*"

and assumed by the speaker to be "*hearer old*" (Diessel, 1999, p. 106 citing Prince, 1992). Himmelmann (1996) also adds,

> A central feature of this use is that the speaker anticipates problems with respect to the information used in referring to a given referent. That is, the speaker is uncertain whether or not the kind of information he or she is giving is shared by the hearer or whether or not this information will be sufficient in allowing the hearer to identify the intended referent (Himmelman, 1996, p. 230).

The following example illustrates the recognitional use:

(9) …it was filmed in California, **those dusty kind of hills** <u>that they have out here by Stockton and all</u>, … so … (Himmelmann, 1996, p. 230; modified)

In this example, *those* is used in an adnominal construction by the speaker to signal presumed shared knowledge between both the speaker and hearer. In order to maximize the chances that the hearer will be able to recall the referent, the speaker provides additional details about the hills – that they are the ones in Stockton which the speaker believes that the hearer has also seen. According to Himmelmann (1996), this is a common tendency for the recognitional use in order "to make the intended referent more accessible" (p. 230). This anchoring information frequently occurs in the form of a relative clause as in the underlined portion of (9) above. Furthermore, additional indicators such as false starts, hesitations, and pauses are often present which reflect the speaker's uncertainty that the hearer is able to accurately identify the particular referent.

A few additional features are commonly associated with the recognitional use. First, according to Diessel (1999), the recognitional use tends to be limited to the adnominal

construction as shown in (9) above. Secondly, Himmelmann (1996) observes that distal demonstratives are strongly associated with the recognitional use. Thirdly, the referents of recognitional demonstratives are generally "of only peripheral importance (low topicality)" (p. 230). Once the referent is introduced to the discourse, they typically do not continue as salient participants in the discourse; they are introduced, and the speaker moves on.

Table 31 below summarizes the exophoric and basic endophoric demonstrative functions.

	Subtype	Description	Possible DEM types: *Distal, Proximal, Both*	Possible syntactic formations: *Adnominal, Pronominal, Both*
Exophoric (situational)	Gestural	• Refers to entity in the physical surroundings • *May* be accompanied by pointing gesture to referent	Both	Both
	Deixis am phantasma	• Deictic center is shifted from speaker to the perspective of some entity in a story, narrative, imagined scenario, etc. • *May* be accompanied by pointing gesture to referent	Both	Both
Basic endophoric (textual)	Discourse deictic	• Anaphoric or cataphoric • Refers to proposition or event expressed in an immediately adjacent clause, sentence, story, etc. • Connects discourse segments	(English) Anaphora: proximal or distal Cataphora: proximal	Both; strong tendency for pronominal
	Tracking	• Anaphoric • Refers to and is co-referential with prior NP in the discourse • Tracks unexpected, non-topical referents (those with either 'semi-active' or 'inactive' cognitive status) • May indicate a topic shift • May reactivate a previously introduced referent	Preference for one or the other, depending on the language	Both; possible preference for adnominal
	Recognitional	• Indicates an 'inactive' referent that is new to the discourse but known to both speaker and hearer • Referent does not continue as salient once introduced • Can occur with a relative clause or other clarifying descriptive information • Pauses, hesitations, false starts, etc. commonly present	Tendency for distal	Adnominal

Table 31. Summary of basic pragmatic functions of demonstratives (based primarily on Diessel, 1999; Himmelmann, 1996; also Croft, 2019).

Exophoric demonstratives refer to an entity in the physical surroundings of the interlocutors. The speaker is at the deictic center, the demonstratives make distance contrasts such as proximal and distal in a majority of languages, and they frequently (but not necessarily) co-occur with a pointing gesture. It is also possible for the deictic center to shift from the speaker to a person or other entity in some type of imagined scenario (i.e. *Deixis am Phantasma*). Endophoric demonstratives are textual in that they refer to an entity or proposition uttered within the discourse itself. Endophoric demonstratives can further be broken down into the discourse deictic, tracking, and recognitional uses. Discourse deictic demonstratives refer either anaphorically or cataphorically to a proposition or event in an adjacent discourse segment. Tracking demonstratives carry the function of keeping track of a specific, non-topical referent mentioned at some point earlier in the discourse in the form of a noun phrase. Tracking demonstratives may also reactivate a non-topical referent that has gone unmentioned for some time. Finally, recognitional demonstratives indicate a shared referent between speaker and hearer based on assumed shared knowledge – one that is new to the discourse or inactive in the hearer's consciousness.

6.1.4 Demonstratives and grammaticalization

This section provides background on how demonstratives cross-linguistically tend to evolve based on previous observations. These details guide my own analyses of the different ways that demonstratives have developed in O'odham and whether or not those developments align more-or-less with previous findings. While the primary concerns in this study are the more grammaticalized demonstrative types and their endophoric precursors (i.e. discourse deictic, tracking, and recognitional demonstratives), I also address the likely starting point for all

demonstratives – the exophoric category – in order to provide a complete picture of the hypothesized evolution of demonstratives in the world's languages.

Regarding the hypothesized lexical source for demonstratives, Diessel (1999), in fact, makes the case that demonstratives may be one of the few grammatical items that do not have prior lexical sources, even though proponents of grammaticalization show that grammatical items generally evolve from an earlier source (e.g. Bybee et al., 1994; Hopper & Traugott, 1993; Lehmann, 1985, 2002). It may be that demonstratives are a part of the basic vocabulary of the world's languages and that they start off as grammatical, especially given their hypothesized earlier exophoric function. At this stage,

> Demonstratives are used to orient the hearer in the speech situation, focusing his or her attention on objects of interest. This is one of the most basic functions of human communication for which there might be a particular class of linguistic expressions that emerged very early in the evolution of language (Diessel, 1999, p. 152).

Exophoric demonstratives carry "the prototypical use from which all other uses derive" (Diessel, 1999, p. 110; for an alternative view, see Himmelmann, 1996, pp. 223-224; also Hanks 1990; Fuchs 1993; Laury 1997). One of the primary pieces of evidence in support of this view is that more grammaticalized demonstrative uses develop from the tracking, discourse deictic, and recognitional uses described earlier in §6.1.3; they never develop directly from exophoric demonstratives, suggesting that the latter represents the original, historical source from which all other types of demonstratives evolve (refer to Diessel, 1999, pp. 110-114; 150-153 for a complete presentation of the author's arguments). In other words, endophoric demonstratives have already somewhat grammaticalized from their original exophoric sources, which can then in turn further grammaticalize and expand into other more

grammatical types of morphemes. Furthermore, Diessel (2013) points out, in the author's cross-linguistic examination of over 300 languages, not a single case exists in which a prior lexical source can clearly be traced (for possible counter examples and arguments, see Heine, Kuteva, Long, Narrog, and Wu, 2020). The same appears to be true of O'odham or, at least, reconstructions of Uto-Aztecan (e.g. Langacker, 1977) offer no hypothesized lexical source from which demonstratives grammaticalized in the family. Based on the lack of evidence of a prior source, I adopt Diessel's view here.

Diessel provides a set of criteria in order to detect the degree to which demonstratives have grammaticalized. These pertain to functional, syntactic, morphological, and phonological changes, all of which may or may not be present. The criteria are not categorical but exist on a continuum. The author summarizes these changes as follows:

(10) **Functional changes**
a. Grammatical items that developed from demonstratives are no longer used to focus the hearer's attention on entities in the outside world.
b. They are deictically non-contrastive.

Syntactic changes
c. Their occurrence is often restricted to a particular syntactic context.
d. They are often obligatory to form a certain grammatical construction.

Morphological changes
e. They are usually restricted to the distal or, less frequently, the proximal form.
f. They may have lost their ability to inflect.

Phonological changes
g. They may have undergone a process of phonological reduction.
h. They may have coalesced with other free forms. (Diessel, 1999, p. 118)

The functional types of changes in (a) and (b) simply indicate that demonstratives have evolved beyond orienting the hearer in the external surroundings and expressing deictic distinctions in physical space such as proximal and distal. They have evolved to refer to entities stated within discourse, itself.

The other developments listed above in (c)-(h) are self-explanatory. These types of changes, if present, suggest more advanced stages of grammaticalization for demonstratives.

Diessel (1999) additionally observes that the evolutionary paths that a demonstrative may take is determined by the syntactic context or construction type in which it occurs (p. 115). For example, Diessel observes that while pronominal demonstratives commonly evolve into third-person pronouns, adnominal demonstratives "give rise to grammatical items that function as operators of nominal constituents" (p. 115).

These observations can be illustrated by Croft's (2019) **Accessibility Scale** which sheds light on a few different possible paths that pronominal and adnominal demonstratives may follow (see also Ariel, 1988, 1990; Givón, 1983). To begin with, the scale illustrates that different referring phrases (e.g. pronominal and adnominal demonstratives, pronouns, zero expression, etc.) tend to express different degrees of accessibility. **Accessibility** refers to how easily a referent "can be accessed by the hearer" (p. 67) once a **discourse file** for the referent has been opened – in other words, after a referent has been introduced into the discourse and is in the hearer's consciousness. Croft borrows the following terminology from Chafe (1987) to describe the different referent types: First, **active referents** represent highly salient referents "at the center of the hearer's consciousness" (Croft, 2019, p. 67) which tend to be expressed by pronouns that anaphorically indicate previously introduced referents in the discourse, as well as zero anaphora, and indexation on the verb. Next, **semi-active referents**, which I briefly introduced in §6.1.3, are in the hearer's short-term memory but not at the center of the hearer's awareness perhaps because other referents have been introduced into the discourse, a topic shift in the discourse has developed, etc. These tend to be expressed by tracking (anaphoric) adnominal demonstratives as well as definite noun phrases. Finally,

inactive referents are those of which the speaker and hearer share knowledge but have "not been activated in the discourse, at least not recently" (p. 68), for example, with referents that the hearer can infer from context or with first mentions in the recognitional use. It may also be the case that the referent has not appeared in the discourse for some time and presumably is no longer held in the hearer's consciousness. These tend to be expressed with non-anaphoric definite noun phrases, according to Croft, although evidence in the O'odham data suggests that anaphoric adnominal (i.e. tracking) demonstratives are also possible (see §6.4.2). These gradient degrees of accessibility and their referring phrases can in turn be ranked from high to low accessibility on the Accessibility Scale. The scale represents an "information status continuum" (Croft, 2019, p. 69) rather than discrete, non-overlapping categories. I provide a modified version of the scale below (also see Givón, 1983; Ariel 1988, 1990).

(11) ↑ **Higher accessibility**

Referring phrase	***Morphosyntactic expression***
zero anaphora	zero
predicate indexation	affix
(anaphoric) third-person pronouns	pronoun
first/second-person pronouns	pronoun
pronominal demonstratives	pronoun
adnominal demonstratives	determiner + noun
definite noun phrases	determiner + noun

↓ **Lower accessibility** (based on Croft, 2019, p. 69)

The crucial point to note is that the pronoun morphosyntactic expressions are ranked higher on the scale while the determiner + noun expressions are ranked lower. This split into higher and lower, Croft argues, begins with the two basic demonstrative constructions – pronominal

and adnominal demonstratives – which grammaticalize in different directions. This is illustrated in Figure 12 below.

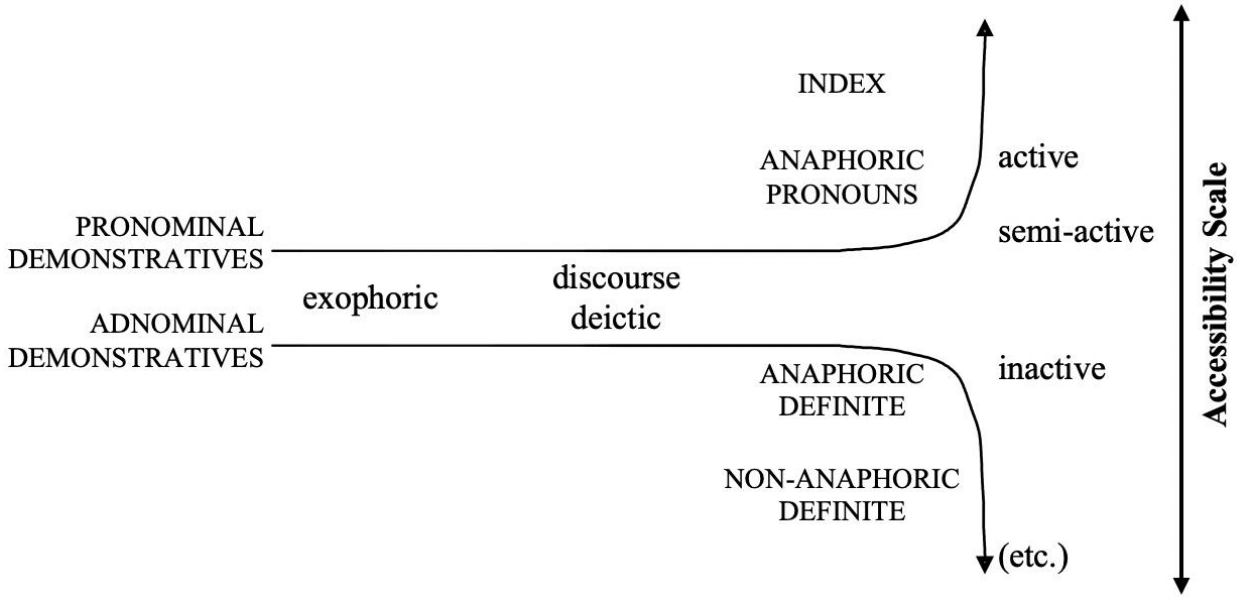


Figure 12. The grammaticalization pathways for basic demonstrative constructions (based on Croft, 2019, p. 76).

Both constructions start off as exophoric demonstratives which indicate referents in the surrounding physical situation. Later they extend to uses within discourse itself at which point their original deictic meanings have undergone some degree of semantic reduction. Discourse deictic demonstratives further extend towards tracking uses which often indicate semi-active referents while recognitional uses have a "somewhat lower accessibility information status than the tracking use, i.e. closer to inactive status" (p. 73). Eventually, pronominal tracking constructions may further extend to (anaphoric) third-person pronoun meanings which are higher up on the accessibility scale. Adnominal tracking constructions on the other hand move down the accessibility scale to express less accessible referents.

In his own study, Diessel accounts for eighteen different grammaticalization channels (see Diessel, 1999, Chapter 6 for a comprehensive description); the channels which are

relevant to the discussion of demonstratives in O'odham are: third-person pronouns, specific-indefinite articles, determinatives, and connectives. In addition to these, I also discuss what I have termed 'referring postpositions' (see §§6.4.3 and 6.5.1), and finally, demonstrative placeholders. These are described in greater detail below.

Third-person pronouns

It is common cross-linguistically for third-person pronouns to evolve from pronominal demonstratives (Croft, 2019; Diessel, 1999; Givón, 2001; Heine & Kuteva, 2002; Heine & Song, 2011; Himmelmann, 1996; Lehmann, 2002). This process is so pervasive, Himmelmann (1996) observes, that in many languages it is not always clear where the line should be drawn between 'demonstratives' and 'third-person pronouns,' being an issue of how advanced a demonstrative is on the grammaticalization path. Languages which utilize the same form for both functions as is the case in O'odham (§6.2) "have presumably extended the demonstrative pronouns, or at least one of the demonstrative pronouns … without a split in morphological form" (Croft, 2019, pp. 74-75). Diessel's (1999) crosslinguistic findings suggest that pronominal demonstratives that specifically serve the tracking function tend to develop into third-person pronouns. Diessel explains,

> At the beginning of this cline we find anaphoric [tracking] pronominal demonstratives tracking emphatic, contrastive and unexpected discourse topics. Anaphoric demonstratives that develop into third-person pronouns become … gradually extended to all persisting topics (Diessel, 1999, p. 120).

In terms of the Accessibility Scale (see Figure 12), pronominal demonstratives may move up the scale from indicating semi-active referents to those with active status in the hearer's

consciousness. Once a pronominal demonstrative is extended to active referents, then it can be said pronominal demonstratives have grammaticalized into third-person pronouns (see Croft's full discussion on pp. 72-76).

Specific indefinite articles

Diessel (1999) observes that some languages have a special indefinite article to refer to a participant that is *both* new to the discourse as well as to the hearer (pp. 109; 138-139). The most typical, crosslinguistic source for this article is the numeral one (p. 138; also Givón, 1995; Wright and Givón, 1987). English in many cases employs the article *a/an* to mark both specific and non-specific indefinites; however, it has been observed in colloquial English in particular that adnominal proximal demonstratives *this* and *these* have been extended to indicate a specific yet indefinite referent. This particular use of the proximal demonstrative is sometimes referred to as the "new-*this"* (e.g. Himmelmann, 1996; Wald, 1983). Here is an example from Givón (1990):

(12) …So next he passes **this bum** and[,] boy, the guy was real ragged, run down and all, was not even begging, just sitting there; so he stops and gives him a dollar and the next thing you know the guy is screaming… (Givón, 1990, p. 921, cited in Diessel, 1999, p. 138)

In this instance, the proximal demonstrative *this* introduces a specific and new (i.e. inactive) referent that has not been previously mentioned in the discourse. The example also highlights another common feature of the specific indefinite article; once the speaker introduces the referent, it often continues in the discourse as a major participant (Wright & Givón, 1987, pp. 15-16). This feature, however, is not always present (Wald, 1983, p. 103).

One other observation regarding specific indefinite articles is the possibility for it to occur with a relative clause which provides additional information about the referent. In

English, Perlman (1969) observes that this use of *this* "appears to serve as a signal for the additional information" (p. 78) about the new, but specific referent. This is exemplified below.

(13) **This man** who dropped by yesterday sold me some brushes. (Perlman, 1969, p. 76)

The referent *this man*, is new and inactive at this point in the discourse, and is further described to the hearer with the accompanying relative clause.

Determinatives

Diessel (1999) makes the case that adnominal, recognitional demonstratives can evolve into determinatives which function "to mark the nominal head of a relative clause" (p. 135; following Quirk et al., 1972, p. 217). At this stage, the demonstrative continues to indicate referents with inactive status, but the demonstrative does not signal a shared or known referent between the speaker and hearer. I have additionally observed that the demonstrative appears to indicate a category or type of referent which is restricted or defined by the relative clause with which it occurs. This point will be illustrated in the comparative recognitional and determinative examples that follow.

During the hypothesized earlier recognitional stage, adnominal demonstratives commonly co-occur with relative clauses which "make the intended referent more accessible" (Himmelmann, 1996, p. 230). Example (14) below illustrates the recognitional use in which the adnominal demonstrative, *that*, and the noun it refers to, *wine*, are further elaborated upon in the underlined relative clause.

(14) (Recognitional)

[Written in a text message] I'm finishing off **that wine** <u>you brought to our house the other night, and it's divine</u>!

In this case, *that* indicates to the hearer that the speaker has a specific bottle of wine in mind which is known to the both of them. The demonstrative does more than simply mark the nominal head, *wine*, of the relative clause in this instance. The speaker utilizes the relative clause to increase the certainty that the hearer (or reader of the text message in this case) will remember and accurately identify the specific and known referent that otherwise is inactive in the recipient's consciousness. Without the relative clause, the hearer may be less certain about which bottle of wine the speaker is referring to, especially in the case that several different bottles of wine had been brought over by various guests that evening.

In contrast to the recognitional use in (14) above, example (15) illustrates the related but more grammaticalized determinative function of a demonstrative which is also modified by a relative clause.

(15) …provision was made for payment for unemployment relief by nationwide taxation rather than by a levy only on **those states** <u>afflicted with manpower surplus</u>. (Himmelmann, 1997, p. 78 cited in Diessel, 1999, p. 135; modified)

In the case above, it cannot be said that the demonstrative *those* is being used to signal a shared referent, a particular group of states in this case, between the speaker and hearer. It also does not refer to a previously introduced discourse participant, and therefore is also not associated with the tracking use. Instead, the demonstrative in the adnominal construction serves the sole grammatical function of marking *states* as the nominal head of the relative clause that follows it. Furthermore, I would argue, at this stage the speaker does not actually have a specific referent in mind but rather a subtype or category of states that are *afflicted with a manpower surplus*. Whether or not the speaker accurately identifies which specific states are afflicted with the surplus is not important. Instead, the speaker's goal is to conjure up a referent subtype that belongs to the *manpower surplus* category. This is a detail that has

not, to my knowledge, been previously discussed in the literature. Example (16) further illustrates this point. In the example the speaker compares *true artists* to a type of scientist.

(16) The true artist is like one of **those scientists** who, from a single bone, can reconstruct an animal's entire body. (Himmelmann, 1997, p. 78 cited in Diessel, 1999, p. 108; modified)

Whether or not the hearer has ever actually known or heard of a scientist with the skills being described doesn't matter. What does matter is that the speaker has established a category of scientist in order to make a point – what defines a *true artist*. By contrast, in the earlier example in (14), the speaker clearly has a specific referent in mind – the particular bottle of wine someone brought over the other day – and the speaker wants to make certain that the hearer correctly recalls it. In (16) above, the demonstrative serves the sole grammatical purpose of marking *scientists* as the nominal head of the relative clause that follows it and is devoid of any recognitional (or other) function.

The same effect is true when the referent is singular as in the found example below, whereas examples (15) and (16) above evoke plural referents.

(17) I feel that as a YouTuber, everyone gets their moment in the spotlight before someone else comes along and takes your thunder, kinda like **that artist** who has a one-hit wonder and then no one ever hears from them again (Gutelle, 2014, Nov 26).

In (17) above, *that artist* does not refer to a specific artist that is known or expected to be recognized by the audience. This is apparent given the fact that this segment is from an online entertainment news source where the speaker has no knowledge of who the specific, individual readers are. Instead the effect of using the adnominal demonstrative refers to a common category of artist or label that probably most people have at least heard of, the "one-hit-wonder," but it does not necessarily refer to one specific artist.

At the determinative stage, according to Diessel (1999), either an adnominal or pronominal demonstrative may be employed for this grammatical function. In the case of the latter, the independent, pronominal demonstrative, itself, serves as the head of the relative clause. This contrasts to the recognitional function which is strongly associated with adnominal demonstratives. A couple of examples below illustrate pronominal demonstratives with the determinative function. First, in (18), the pronominal demonstrative does not trigger a known referent. Furthermore, which specific people backed up the plan is not important. Instead, the speaker's goal is to establish a category of people who "hailed the message" which is described in the relative clause. The demonstrative also does not align with other uses such as the tracking or discourse deictic functions, recalling neither a previously mentioned participant nor referring to another surrounding segment of discourse.

(18) **Those** who backed a similar plan last year hailed the message.
(Himmelmann, 1997, p. 77 cited in Diessel, 1999, p. 135; modified)

Similarly, in example (19) below, the pronominal demonstrative is used with the goal of establishing the type of people who will be impacted by the tax boost rather than bringing specific members to the hearer's mind.

(19) Similar payroll tax boosts would be imposed on **those** under the railroad retirement system. (Himmelmann, 1997, p. 78 cited in Diessel, 1999, p. 108; modified)

Diessel (1999) does not offer a hypothesis for why the recognitional use has a strong tendency to co-occur with adnominal demonstratives or why it is that once the construction has extended to the determinative stage, either adnominal or pronominal demonstratives are possible; however, in my view, if it is important for the hearer to accurately recall a known and specific referent with the recognitional function, explicitly stating the referent in the

demonstrative construction increases the likelihood of achieving this goal. In the case of the determinative function, if it is not important for the hearer to mentally conjure up a specific referent but instead a referent category that is defined in the relative clause that follows it, then either an adnominal or pronominal demonstrative would aptly serve this purpose.

Table 32 below highlights both the features that recognitional demonstratives and determinatives share as well as how they differ. This summary is based on both Diessel's (1999) and Himmelmann's (1996) observations, as well as my own.

Recognitional	**Determinative**
• Adnominal DEM (most typically)	• Adnominal or pronominal DEM
• May be followed by a relative clause	• Always followed by a relative clause
• First mentions (inactive referent status)	• First mentions (inactive referent status)
• Referents remain non-salient in the discourse	• Referents remain non-salient in the discourse
• DEM indicates a *specific* referent	• DEM indicates a referent *type/category*
• Possible indications that speaker seeks confirmation of known referent (e.g. hesitations, tag Qs, pauses, etc.)	--

Table 32. Summary of recognitional vs determinative demonstratives (based in part on Diessel, 1999 and Himmelmann, 1996).

Connectives

Connectives, or what Diessel (1999) refers to as "sentence connectives," commonly evolve from pronominal, discourse deictic demonstratives that co-occur with some other element such as a postposition or adverb (p. 125). Together the pronominal demonstrative and other element develop to indicate a "semantic relationship between the conjoined propositions" (Diessel, 1999, p. 125) which may be in the form of two separate sentences. Diessel's example below illustrates this pattern in (20) from Hixkaryana (Cariban), spoken in Brazil. The propositions may also consist of a single complex sentence in which one independent

clause is conjoined with a subordinate clause via a connective as in the Khasi example in (21), also in Diessel.[33]

(20) nomokyaknano tuna heno. **ɨre** **ke** romararin
it.was.coming rain QNT **DEM** **because.of** my.field

hokohra wehxaknano
NEG I.was
'It was raining heavily. Therefore I did not work on my field.'
(Diessel, 1999, p. 125 citing Derbyshire, 1985, p. 157)

(21) u khla u la baòm **naŋ-ta** u la thyú
ART tiger ART PAST ate **P-DEM** ART PAST slept
'The tiger ate then he slept.' (Diessel, 1999, p. 126 citing Nagaraja, 1985, p. 100)

Regarding (20), Diessel states that in Derbyshire's data on Hixkaryana, *ɨre ke,* which consists of the demonstrative *ɨre* and the causal postposition *ke* together function to link the two separate clauses. In (21), a complex sentence made up of "two clauses linked by *naŋ-ta* 'then', which is formed from the adpositional marker *naŋ* - and the demonstrative root -*ta*" (see Diessel, 1999, pp. 125-126 for a number of additional examples).

In the course of these developments, one might see demonstratives occur with increased frequency in a construction in which a demonstrative is juxtaposed with an adverb or a particular postposition as in the example above. There may also be indications that the component parts function together as a unit to convey a specified meaning.

[33] For this reason, I prefer the term 'connective' instead of Diessel's "sentence connective."

Other categories: Referring postpositions and placeholders

Diessel (1999) only very briefly alludes to or mentions the two remaining grammatical categories. The first of these I refer to as **referring postpositions** and the second category are often classified in the literature as **placeholders** (e.g. Ganenkov, Lander, & Maisak, 2010; Hayashi & Yoon, 2010, p. 36; Podlesskaya, 2010). I include these because there is evidence of similar uses of demonstratives in O'odham (see §6.5.2). I briefly describe both of these in what follows.

In Diessel's discussion of connectives, the author briefly notes a series of prepositions in German which consist of older forms of the oblique pronominal demonstrative (*dara, dar*) that are no longer used in modern German and a postposition (p. 126). A few of the examples are shown below.

(22)

Modern form	**Source**	**Gloss**
darauf	*da:r-u:f*	DEM.OBL-on.top.of
dazu	*da:ra-zuo*	DEM.OBL-to
dafür	*dara-fure*	DEM.OBL-for
dagegen	*dara-gegen*	DEM.OBL-against (Diessel, 1999, p. 126; modified)

According to Diessel, the demonstrative and postposition combinations can be used "to substitute for a prepositional phrase" (p. 126). A similar development appears in O'odham (see §6.5.2) in which an adposition plus a demonstrative turns into a new or alternative postposition (see §6.4.3). It is often if not always the case that the postpositions refer to specific entities mentioned previously in the discourse and in this manner they behave as 'referring' postpositions.

Finally, Diessel also briefly notes that demonstratives are a common source for different types of hesitation markers (p. 154), one of which is the placeholder function. This

development has also been discussed elsewhere and most often involves a pronominal demonstrative (e.g. Fox, Hayashi, & Jasperson, 1996; Ganenkov et al., 2010: Hayashi & Yoon, 2006; Hayashi & Yoon, 2010; Podlesskaya, 2010; Podlesskaya & Kibrik, 2006, 2009). The function arises when a speaker experiences difficulty in retrieving a specific lexical item. The speaker then utilizes a pronominal demonstrative "to fill the syntactic slot in which the missing word would be produced in an ongoing utterance while s/he engages in a word search" (Hayashi & Yoon, 2010, p. 36). The example below from Korean illustrates this function.

(23) kulikwu **ku:**-ke iss -cyanh-a.
and **MED.DEM**-thing exist-ASSERT-IE
'And there is *ku:-ke* [=that thing], you know.'

mwe-ya (.) ku (0.2) noph-un-ke **achi**.
what-be that high-RL-thing **arch**
'What is it.(.) That (0.2) The tall thing. The arch.' (Hayashi & Yoon, 2010, pp. 38-39; modified)

In (23) above, the pronominal medial demonstrative, *ku:* serves as a substitution for the lexical item that the speaker has trouble recalling. That is, it functions as a placeholder for where the sought-after lexical item would be situated if the speaker were able to recall it. Eventually, the speaker is able to retrieve the lexical item, *achi* 'arch.' Note also that the placeholder substitutes specifically for a noun, which, according to Podlesskaya (2010) is the most common finding (p. 13).

One more example from Russian additionally shows that the placeholder demonstrative often will mirror morphological features of the sought-after noun (Podlesskaya, 2010).

(24) On kupi-l vsjakie **et-i** … pirožny-e.
he buy-PAST.SG.M various **PH-ACC.PL** cakes-ACC.PL
'He bought various PH [whatchamacallit] … cakes.'
Lit. 'He bought various them … cakes.' (Podlesskaya, 2010, p. 12)

The pronominal proximal demonstrative *eti,* serves as a placeholder for the lexical item that the speaker is searching for. The placeholder also takes on the plural accusative form which matches the plural accusative marker used on the target noun *pirožny-e* 'cakes-ACC.PL'.

The placeholder function has added significance in the context of endangered languages in that word-retrieval issues may arise due to the language being used less often.

6.2 Previous descriptions of demonstratives in O'odham and Uto-Aztecan (UA)

In this section I present previous observations about demonstratives in the O'odham literature and, in a few cases, in UA as a whole. Not all of the functions that I discuss in my own study are included in the literature, but these prior observations shed light on a few of the grammaticalization paths that O'odham demonstratives may take.

Beginning with the most basic feature of demonstratives, O'odham scholars (Hale, 1959; Mason, 1950; Mathiot, 1973; Saxton, 1982; Zepeda, 1983) note a basic two-way deictic contrast in the O'odham demonstrative system, namely proximal and distal. This is true of a majority of UA languages (Langacker, 1977, p. 99). Demonstratives are also distinguished for number – singular and plural and they may occur in either adnominal or pronominal constructions. The most basic forms of O'odham demonstratives are *hegai* 'that', *hegam* 'those', *'i:da* 'this', and *'idam* 'these'. I include a few examples of adnominal constructions from Zepeda (1983) in (25)-(28) below. (All morphological glosses throughout this section are my own. Translations are the original authors'.)

(25) **Hegai** cehia 'o pi cicwi.
DIST girl 3.AUX.IPV NEG play
'That girl is/was not playing. (Zepeda, 1983, p. 9)

(26) **Hegam** gogogs c mimstol 'o wo:po'ŏ
DIST.PL dog.PL CONJ cat.PL 3.AUX.IPFV run.PL.IPFV
'those dogs and cats are/were running.' (p. 24).

(27) **'I:da** O'odham 'o ñeok.
PROX person 3.AUX.IPFV talk.IPFV
'This person is/was speaking.' (p. 8)

(28) **'Idam** cecoj 'o ñeñok.
PROX.PL boy.PL 3.AUX.IPFV speak.PL.IPFV
'These boys are/were speaking.' (p. 10)

Examples (29)-(31) show pronominal constructions. They also illustrate that demonstratives may serve as either subjects or objects. In (29) the plural form functions as the subject, and in (30) and (31) the singular distal demonstrative *hegai* functions as a direct object. Furthermore, the examples show that demonstratives may appear in different positions of the clause (i.e. initially, internally, or in clause-final position).

(29) **Hegam** 'o gḍhu daḍhă.
DIST.PL 3.AUX.IPFV over.there sit.PL.IPFV
'They are/were sitting over there.' (Zepeda, 1983, p. 52)

(30) Gogs 'o **hegai** huhu'id.
dog 3.AUX.IPFV **DIST** chase.IPFV
'The dog is/was chasing it (that, her, him).' (p. 34)

(31) Gogs 'o huhu'id **hegai.**
dog 3.AUX.IPFV chase.IPFV **DIST**
'The dog is/was chasing it (that, her, him).' (p. 34)

Example (32) illustrates a proximal, pronominal demonstrative.

(32) N-o **'i:da** ḍ m-si:ṣpa-kuḍ
INTER-3.AUX.IPFV **PROX** COP 2SG.POSS-pin-NMLZ.INST
'Is this your pin (safety pin)?' ' (Zepeda, 1983, p. 102; modified)

The singular forms can also be reduced to *'id* and *heg* in some instances as in (33) and (34) below. To my knowledge, the contexts in which reduction occurs has not been systematically described in the literature. I will address their distributions in §6.6.

(33) **'Heg** 'añ hab 'a:g ceoj.
DIST 1SG.AUX.IPFV thus say man
'That (is the) man I mean.' (Saxton, 1982, p. 191)

(34) **'Id** 'añ hab 'a:g ceoj.
PROX 1SG.AUX.IPFV thus say man
'This (is the) man I mean.' (p. 191)

Mathiot (1993) additionally records examples in which *hegai* 'that' is produced as *g*, particularly when it co-occurs with a postposition. I have only observed this in Mathiot's materials, which may suggest that this reduced form is specific to speakers of the Totoguañ subdialect of Tohono O'odham with whom Mathiot worked. An example from Mathiot (n.d.b) is included in the following example:

(35) 'am haha mawa **g'eḍa** 'e-huaṣomĭ
…LOC then put.hand **DIST.inside** 3.REFL-pouch
'…then he put his hand inside [that] his pouch.' (Mathiot, n.d.b, p. 27)

In (35) above, the distal demonstrative *hegai* 'that' is reduced to *g* which is written together with the postposition *'eda* 'inside'. The author does not discuss the reason for doing so, but it may be that Mathiot interprets these cases as unanalyzable chunks suggesting fusion of the construction at least in some cases.

O'odham demonstratives and their most common forms included in the literature are summarized in Table 33. I have also listed the basic demonstrative meanings attributed to each form as well as a few other meanings sometimes assigned to demonstratives.

	Full forms	Reduced forms	Basic DEM meanings	Other meanings
Distal	*hegai, hega'i*	*heg, g*	'that'	'he', 'it', 'him', 'that one' 'that only'
	hegam	--	'those'	'they', 'them'
Proximal	*'i:da*	*'id*	'this'	'it', that', 'this one'
	'idam	--	'these'	'they', 'them'

Table 33. O'odham demonstrative forms and meanings.

The literature also recognizes that demonstratives in O'odham have direct relations to articles, third-person pronouns, and determinatives (i.e. relative-clause markers). These relationships point to a few of the different grammaticalization paths that demonstratives have taken. In the case of articles, the process appears to be complete, whereas with third-person pronouns and determinatives, the processes appear to be ongoing.

First, I will address the article *g* 'the, a', which the literature states very likely evolved from the distal demonstrative *hegai* 'that'. Given the relative certainty of its source and evolution, this topic will not be treated beyond the discussion I present here; that is, I do not include it in my own study. Its evolution is worth noting, however, given its place in the story of possible grammaticalization pathways for demonstratives in the language.

The article *g* does not mark definiteness and can be glossed as either 'a' or 'the'. Mathiot (1973) simply refers to *g* as a generic "noun marker" (p. 107) which serves either a definite or indefinite article (p. 35), while Zepeda (1983) assigns the generic label "determiner" (p. 13). In Tohono O'odham, the article is sometimes produced as *eg* (/ɨg/), although orthographically it is always written as *g*. In Akimel O'odham, a dialect to which Tohono O'odham is closely related (see §2.1), the article is commonly recorded

orthographically as *heg*. These phonological details and the common cross-linguistic tendency for articles (usually of the definite type) to evolve from demonstratives point to the distal demonstrative *hegai* 'that' as a likely source. A couple of examples are shown in (36)-(38) below. The translations are verbatim from the original sources although *g* could be translated as either 'a' or 'the', as noted earlier.

(36) Him 'o **g** haiwañ.
walk.IPFV 3.AUX.IPFV **ART** cow.
'The cow is/was walking.' (Zepeda, 1983, p. 13)

(37) **g** ki:-j hega'i ceoj
ART house-POSS that man
'the house of that man' (Saxton, 1982, p. 193)

(38) **g** ki:-j **g** Huan
ART house-POSS **ART** Juan
'the house of Juana' (Saxton, 1982, p. 193)

Mason (1950) explicitly discusses the link between the distal demonstrative *hegai* and the article *g* 'the, a', noting that they are "obviously related" (p. 60).

Regarding UA, Langacker (1977) states that many languages have articles that can be traced back to demonstratives (p. 100). Langacker notes of their grammaticalization that the articles "can be regarded as demonstrative elements bleached of much of their semantic content and adapted for various grammatical uses" (p. 100). Of O'odham in particular, the author states that the article *g* 'the, a' is derived from *hegai* 'that'.

Next, the distal demonstratives *hegai/heg* 'that' *and hegam* 'those' are frequently cited as serving as third-person pronouns in O'odham (Hale, 1959, pp. 160-161; Mason, 1950, pp. 59-60; Mathiot, 1973; Saxton, 1982, p. 190; Zepeda, 1983, pp. 18, 35, 159). Langacker (1977) states that this is a common strategy in UA languages (p. 102). Although Mason (1950) recognizes the relationship between the distal demonstratives and their use as third-

person pronouns, from the author's point of view, O'odham has "no true third-person pronouns" (pp. 59). In fact, it is often difficult to distinguish between pronominal demonstratives proper and third-person pronouns. For example, there are cases in which *hegai/heg* can be glossed as either 'that' or 's/he, it', suggesting that pronominal demonstratives continue to evolve.

Both Mathiot (1973) and Saxton (1982) suggest that not only the distal but also the proximal demonstratives can carry properties of independent, third-person pronouns. Mathiot (1973) states that in either case they are used for emphasis as either subjects or objects (p. 93). Person marking otherwise is reflected by the auxiliary system for subjects and in a series of object markers on verbs (see §§3.3). Some examples from Zepeda (1983) illustrating third-person pronouns are included in (39)-(42). Examples (39) and (40) illustrate the use of *hegai* and the plural counterpart *hegam* as third-person, subject pronouns and in (41) and (42) as object pronouns. Note that in (42) the third-person, plural object marker *ha-* agrees in number with the plural object pronoun *hegam*, whereas the verb in (41) is zero-marked, reflecting the singular status of the pronoun *hegai.* All of the examples below are from Zepeda (1983).

(39) **Hegai** 'o tako ṣoak.
3SG 3.AUX.IPFV yesterday cry.IPFV
'He/she was crying yesterday.' (Zepeda, 1983, p. 20)

(40) S-hottam 'o hihim **hegam**.
STAT-quick 3.AUX.IPFV walk.PL.IPFV **3PL**
'They are/were walking quickly.' (p. 20)

(41) Gogs 'o **hegai** huhu'id
dog 3.AUX.IPFV **3SG** chase.PL.IPFV
'The dog is/was chasing it (that, her, him).' (p. 34)

(42) Gogs 'o ha-huhu'id **hegam.**
dog 3.AUX.IPFV 3PL.OBJ-chase.PL.IPFV **3PL**
'The dog is/was chasing them (those).' (p. 34)

Few clear examples of the third-person pronoun use of *'i:da* 'this' or *'idam* 'these' are included in grammatical descriptions of O'odham despite, for example, Saxton (1982) and Mathiot's (1973) claims that both the distal and proximal demonstratives may serve this function. The lack of examples may indicate that the distal series, *hegai/heg* and *hegam* are the primary third-person pronouns as Zepeda's (1983) O'odham grammar suggests. It may also simply reflect the overall asymmetric distribution of distal and proximal demonstratives in the language (see Figure 14).

Finally, demonstratives may also mark the nominal head of a relative clause or serve as the nominal head itself. Demonstratives are associated with relative clauses across UA languages (Langacker, 1977, p. 178). Both Saxton (1982) and Zepeda (1983), in fact, state that their presence is obligatory with relative clauses. Saxton provides the example shown in (43) below. Note the presence of the subordinator (SBDR) *m-* on the third-person, perfective auxiliary *at* (*'at* in isolation) following the adnominal demonstrative *hega'i ce'oj* 'that man'. The function of this prefix is to mark subordinate clauses of which the relative clause is one type (enclosed in brackets). Schematically, I represent this construction type as DEM N [*m-* AUX CLAUSE]$_{REL}$.

(43) Nt wo namkid **hega'i ce'oj** [m-at wo cikp]
1SG.AUX.PFV FUT pay.PFV **DIST** **man** SBDR-3.AUX.PFV FUT work.PFV
'I'll pay the man that works.' (Saxton, 1982, p. 194)

Saxton comments that *hega'i* in this case is "generic rather than demonstrative" (p. 194). The adnominal demonstrative *hega'i ce'oj* does not indicate a specific referent but any man who will work. This suggests that the demonstrative may be purely grammatical in this case, invoking neither the recognitional, nor tracking, nor discourse deictic meanings. Instead, it serves the purpose of marking the head, *ce'oj* 'man', of the relative clause. This example

specifically illustrates the grammaticalized function described earlier in §6.1.4 which Diessel (1999) refers to as 'determinatives.' I analyze and discuss this feature in greater depth in §6.5.2.

There are other cases, however, where the demonstrative of the relative clause construction does invoke a specific referent. Example (44) could be exophoric, referring to a specific child whose crying is within earshot, while example (45) could be either a tracking or recognitional use. In either case, the speaker appears to have a specific referent in mind (i.e. it is not strictly determinative in function; see discussion in §6.1.4). Without any context, it is difficult to state with certainty. Regardless, what remains true is that the distal, adnominal demonstratives mark the heads, *'ali* 'child' and *huawĭ* 'deer', respectively, of the relative clauses, but likely with additional pragmatic functions.

(44) **Hegai** **'ali** [m-o ṣoak] 'o wuḍ ñ-maḍ.
DIST **child** SBDR-3.AUX.IPFV cry.IPFV 3.AUX.IPFV COP 1SG.POSS-child
'That child that is crying is my child.' (Zepeda, 1983, p. 105)

(45) N-at mu: **hegai** **huawĭ** [m-apt gatwĭ]?
INTER-3.AUX.PFV die.PFV **DIST** **deer** SBDR-3.AUX.PFV shoot.PFV
'Did that deer you shot die?' (Zepeda, 1983, p. 106)

Pronominal constructions may also occur in these types of relative clause constructions as shown in (46) and (47) below. Example (47) includes one instance in which the proximal demonstrative *'i:da* is used as the head of a relative clause, although Saxton cites *hegai* most frequently, and Zepeda only includes examples with the distal *hegai.* Again, this may be due to the overall distributions of the two demonstrative types in the language (see §6.6.2).

(46) **Hegam** [m-ac 'am ha-ñeid] 'o wuḍ mamakai.
DET.PL SBDR-1PL.AUX.IPFV LOC 3PL.OBJ-see.IPFV 3.AUX.IPFV COP doctor.PL
'Those (people) that are/were watching there are/were doctors.' (Zepeda, 1983, p. 106)

(47) Nt wo ma: **'i:da** [m-apt ñ-ma:] …
1SG.AUX.PFV FUT give.PFV **PROX** SBDR-2SG.AUX.PFV 1SG.OBJ-give.PFV
'I'll give this that you gave me …' (Saxton, 1982, p. 253)

In sum, previous studies have noted (either implicitly or explicitly) a few different grammaticalization pathways in relation to demonstratives in O'odham. While the grammaticalization of singular distal demonstratives into an article or 'noun marker' appears to be more-or-less complete, in the case of pronominal demonstratives and their relationship to third-person pronouns, some ambiguity in meaning remains in the language suggesting that change continues to be an ongoing process. Previous cross-linguistic examinations note the common evolutionary relationship between pronominal demonstratives and third-person pronouns (see §6.1.4). According to Croft (2019), once pronominal demonstratives indicate active referents, they have transitioned into third-person pronoun status. Given the ambiguity in meaning that others have noted regarding O'odham pronominal demonstratives, the expectation is that, over time, pronominal forms would increasingly indicate active referents in the data. If this is not the case, then we can consider different possible explanations for this, one possibility being that the process has to some extent been disrupted due to language shift. Similarly, the evolution of demonstratives into determinatives also appears to be ongoing given that their source meanings remain detectable. According to Diessel (1999), demonstratives that co-occur with a relative clause can over time extend to a purely grammatical, determinative function in which the demonstrative either marks the head of the relative clause or functions as the head itself. The demonstrative at this stage is devoid of any other meaning. I also note in §6.1.4 that at this stage the demonstrative indicates a *type* of referent which the accompanying relative clause defines rather than a *specific* referent that both the speaker and hearer have in mind. The expectation in O'odham then is the

distribution of this function might increase over time as it advances towards more grammatical status and is used more frequently for the determinative function.

Similar grammaticalization trends have also been observed across the UA language family. Later in §§6.4.3 and 6.5.2, I include additional grammaticalized functions of demonstratives not previously discussed in the O'odham literature.

6.3 Data and token selection

As with the previous study in Chapter 5, I examine grammaticalization patterns of demonstratives across the historical data from the early 1900s, the midcentury data from the late 1950s and early 1960s, and the modern data from the early 2000s. The reader is directed to §4.1 for an in-depth description of these data collections. This timespan includes a period during which the community was predominantly monolingual in O'odham through the early 2000s by which time English is the primary means of communication.

The number of tokens for each demonstrative type (i.e. distal and proximal) that I examine are recorded in Table 34 below for the data collections (refer to §6.6 for a more detailed breakdown and analysis). The **combined total is 634 tokens** (450 + 184 = 634).

	Distal	**Proximal**
Historical	110	62
Midcentury	139	22
Modern	201	100
TOTALS:	450	184

Table 34. Token totals for demonstratives in each data period.

I identified tokens of demonstratives in the historical data by conducting a search in FLEx (see §§4.1.1and 4.2) for the full and reduced forms – namely *hegai, hega'i,*[34] *heg* 'that', *hegam* 'those' and *'i:da,'id,* 'this', *'idam* 'these' as listed earlier in Table 33 (see §6.2). The exception to this was Juan Dolores's autobiography (Dolores & Mathiot, 1991) which is in PDF format. In this case, I used the search bar to locate all of the demonstrative tokens, and then I entered the isolated examples into FLEx for morphological glossing. For the former distal series, I identified 110 tokens in the historical data. For the latter proximal series, I identified 62 tokens.

For the midcentury data, I employed only the *Coyote Stories* narrated by Jose Pancho (Pancho & Mathiot, 1959-1960) for this particular analysis; while Mathiot's (n.d.a, n.d.b) two- volume dictionary includes many examples which are extracted directly from various oral narratives, the examples are given in isolation and are completely decontextualized. Given that the meaning of demonstratives is largely discourse based, it is imperative to have access to all of the surrounding discourse in order to accurately interpret the function of a given occurrence of a demonstrative. For this reason, I exclude Mathiot's dictionary as a data source for this particular study. The uniformity of the texts will need to be taken into account particularly when looking at distributions in the quantitative portion of this study (§6.6) as it may emphasize patterns that may be due to the text type as well as the single speaker.

I identified each individual demonstrative token in the midcentury texts by using the search field in the PDF versions of the stories. I searched for all of the possible full and reduced forms noted above. Next, the phonetic version of each token was also noted in Mathiot's phonetically transcribed versions of the *Coyote Stories* primarily to check for

[34] Often glossed as 'that' and in other cases as 'that only'.

consistency on whether a token is produced in its full or phonologically reduced form (see additional coding notes in §6.4 and Table 36). I then copied these tokens into FLEx and morphologically glossed them. The distal demonstratives in the midcentury data totaled 139 tokens, and the proximal demonstratives totaled 22 tokens.

In the modern data, I identified tokens of both proximal and distal demonstratives with the search function in FLEx as I did with the historical data. The identified forms were *hegai, hega'i, heg, eg,* and *hegam* for the distal series totaling 201 tokens, and *'i:da, 'id, idam* for the proximal series with a total of 100 tokens.

Finally, I eliminated a small number of tokens from this study: 19 distal and 6 proximal tokens (25 tokens in total). These tokens are excluded from the totals listed in Table 34. In a majority of cases, I made the decision to omit a token because the referent of the demonstrative was not clear, even after discussing these cases with Mr. Robert Cruz, the primary O'odham language consultant. In other cases, I encountered some type of inconsistency in the data which led me to discard the token from the study. For example, in one case, Mathiot records *heg* in the orthographic version of one of the *Coyote Stories* but in the phonological representation of the same text, the demonstrative is not recorded. In a few other cases, the speaker produced a demonstrative construction at the point of a false start, making it impossible to assess either the referent the speaker had in mind or the function of the demonstrative.

6.4 Coding procedures for demonstrative constructions

According to Diessel (1999), there are a number of possible signs (e.g. phonological, syntactic, etc.) that a demonstrative has undergone some degree of grammaticalization. Through a detailed examination and comparative analysis of these possible indicators of

change in the historical, midcentury, and modern data for O'odham, it may be possible to observe the impact or lack thereof that language shift has had on the evolution of demonstratives. This section presents the coding criteria that I employ for analyzing the data. Later in §6.5, I present a detailed descriptive analysis of numerous examples in the O'odham data and in §6.6 a quantitative analysis.

Based on his typological study of demonstratives, Diessel (1999) observes that the demonstrative constructions, either pronominal or adnominal, largely determines the paths that they will take, which can be illustrated by the Accessibility Scale (see §6.1.4). As the constructions grammaticalize in certain directions, they may show other indicators of change over time. First, one **demonstrative type** (e.g. proximal or distal) may come to be more strongly associated with a particular function than another. For example, Himmelmann (1996) finds a strong association between distal demonstratives and the recognitional function (p. 235). Diessel refers to this as a type of "morphological" change (see pp. 118-119). Secondly, **phonological changes** such as reduction may also be observable over time as a demonstrative is increasingly used for more grammaticalized functions. Thirdly, there may be **syntactic changes**. A demonstrative may show signs of being "restricted to a particular syntactic context" (Diessel, p. 118) or an obligatory part of a larger construction, suggesting more advanced grammaticalization. Finally, a series of **functional changes** take place as demonstratives develop more grammaticalized meanings over time. Diessel posits that all demonstratives start off as exophoric, referring to entities in the surrounding real or imagined physical space, and later they point to and organize referents within discourse itself (see §6.1.4). I assume this is true of O'odham, as well. While I distinguish exophoric cases in the data from the endophoric cases out of necessity, of primary interest here are the latter;

namely, the basic endophoric functions (i.e. discourse deictic, tracking, and recognitional) as well as the more advanced grammatical developments that endophoric demonstratives may undergo over time.

Based on these premises, I code for and analyze the different components summarized in Table 35 below in the O'odham data.

Demonstrative type	**Phonology**	**Syntax**	**Function**
Distal • *hegai* 'that' • *hegam* 'those' Proximal • *'i:da* 'this' • *'idam* 'these'	Full forms • *hegai, hega'i* 'that' • *hegam* 'those' • *'i:da* 'this', • *'idam* 'these' Reduced forms • *heg, g* 'that' • *'id* 'this'	DEM (pronominal) • + [*m*-AUX CLAUSE]$_{REL}$ • + POSTP • + POSTP N • + Other DEM N (adnominal) • + [*m*-AUX CLAUSE]$_{REL}$ • + Other	(Exophoric) Basic endophoric • Discourse deictic • Tracking • Recognitional More grammatical • Third-person pronouns • Specific indefinite articles • Determinatives • Connectives • Referring postpositions • Placeholders

Table 35. Summary of coding for demonstratives.

After I code for each of the details listed in Table 35 separately, I then examine if any clear relationships exist between particular demonstrative types, phonological form, syntactic constructions, and the endophoric/more grammatical functions listed above in §§6.5 and/or 6.6. I also analyze if any trends can be more strongly associated with the early-, middle-, or late-shift periods (see §2.4) of which the historical, midcentury, and modern data represent. Before doing so, I will describe in greater detail how I code for each of the variables listed in Table 35 above.

Beginning with demonstrative type, each token of a demonstrative in the data is coded for whether it is proximal or distal. As reflected in Table 35 above and as described in §6.2, O'odham employs distal and proximal demonstratives – *hegai* 'that', *hegam* 'those', and *'i:da* 'this', *'idam* these' respectively, which are also marked for number (singular and plural).

I also code each token for phonological shape – whether the singular demonstratives occur in either their full or reduced forms – *hegai* versus *heg* 'that' and *'i:da* versus *'id* 'this' (also described in §6.2). The form that a demonstrative takes in association with the construction in which it occurs and its function can add to the picture of where it falls on a grammaticalization continuum. It has been observed that phonological reduction tends to be linked with more grammaticalized uses (Bybee et al., 1994; Hopper & Traugott, 1993; Lehmann, 1985, 2002).

Plural demonstratives in O'odham only appear in their full forms as shown in Table 35 above. As I briefly mentioned in §5.4, Mathiot also includes phonetically transcribed versions of the *Coyote Stories* (Pancho & Mathiot, 1959-1960). In the case of the distal demonstratives, Mathiot records the forms *hegai*, *heg*, *ʰeg*, *eg*, *egə*, *g*, *hegam*, *ʰegam*, and in one case, *hig.* For the proximal demonstratives, the author records *'i:da, 'i:də 'id,* and *'idam.* Mathiot generalizes these variants as the forms shown in the 'Orthographic representations' column in Table 36 below.

		Phonetic transcriptions[35]	Orthographic representations	Basic gloss
Proximal	Full	*'i:da*	*'i:da*	'this'
		'idam	*'idam*	'these'
	Reduced	*'id, 'idə*	*'id*	'this'
Distal	Full	*hegai, ʰegai*	*hega'i*	'that'
		hegam, ʰegam	*hegam*	'those'
	Reduced	*heg, ʰeg, hig*, eg, *egə*	*heg*	'that'
		g	*g*	'that'

Table 36. Summary of Mathiot's phonetic and orthographic demonstrative representations.

Finally, I code for several different syntactic patterns and several possible functions which are also listed in the previous table (Table 35) above. I elaborate on and provide further justification for the coding of these syntactic patterns and functions in greater detail in the subsequent sections.

6.4.1 Coding for syntax

The syntactic construction in which a demonstrative occurs plays a vital role in the possible grammaticalization pathways a demonstrative might take (Diessel, 1999, p. 116; also Croft, 2019). Furthermore, changes to syntax such as a demonstrative becoming fixed or obligatory within a larger construction may be indicative of more advanced grammaticalization (Diessel, 1999, p. 118). One facet of this study is to examine whether there are any syntactic indications that a demonstrative has extended to more grammaticalized uses. The goal of this section is to briefly introduce and summarize the different syntactic constructions that I code for in the data. These constructions will be relevant at various points in the descriptive

[35] In Mathiot's phonetic system. Refer to Appendix A for IPA equivalences.

analysis portion of this study in §6.5. Later in §6.6, I also analyze specific syntactic patterns in conjunction with various functions where relevant.

Basic demonstrative constructions

In order to carry out this examination, each token of a demonstrative in O'odham is first coded for whether it is **pronominal (DEM)** or **adnominal (DEM NOUN)**. Table 37 below provides a brief summary and examples of these basic demonstrative constructions in O'odham (see also sentential examples (25)-(34) in §6.2).

	Syntactic configuration in O'odham	**Examples**
Pronominal	DEM	Distal: • *hegai, heg* 'that' • *hegam* 'those' Proximal: • *'i:da, 'id* 'this' • *'idam* 'these'
Adnominal	DEM NOUN	Distal: • *hegai mistol* 'that cat' • *hegam mimstol* 'those cats' Proximal: • *'i:da goks* 'this dog' • *'idam gogoks* 'these dogs'

Table 37. Pronominal and adnominal demonstratives.

Pronominal demonstratives function as an independent pronoun while adnominal demonstratives consist of a demonstrative followed by the noun that it modifies, together comprising a noun phrase. Both constructions may further become components of specific larger constructions which I describe next.

Other syntactic patterns

Pronominal and/or adnominal demonstratives may gradually become obligatory components of larger, more complex syntactic constructions over time. While it is clear that pronominal and adnominal constructions determine the courses that demonstratives will take, it is also possible for these two basic demonstrative formations, themselves, to become components of larger constructions. This is likely the result of the basic types occurring with increased frequency over time in juxtaposition with other types of material (e.g. lexical, phrasal, clausal, etc.). Based on this premise, I extend this discussion on syntax beyond the basic pronominal and adnominal demonstratives to additional areas of grammar, taking note of the material that immediately follows both pronominal and adnominal demonstratives. As a result of this exploration, two additional frequent syntactic patterns emerge in the O'odham data.

The first pattern is noted by Saxton (1982) as well as Zepeda (1983) and was briefly discussed in §6.2. It involves a basic demonstrative (i.e. pronominal or adnominal) followed by the O'odham subordinator *m*-, which prefixes to the auxiliary system in the language, indicating person, number, and imperfective versus perfective aspects (see §3.6.1). The subordinator, as the name implies, marks different types of subordinate clauses. The relative clause is the primary concern in the discussion at hand. That is, I identify constructions that take the form of **DEM** or **DEM N + [*m*-AUX CLAUSE]**$_{\text{REL}}$.

The second frequent pattern is a pronominal demonstrative followed by a postposition or a postposition and noun (**[DEM POSTP]** or **[DEM POSTP N]**). All other patterns are coded as OTHER. These patterns are summarized in Table 38 below.

Basic DEM	**Additional syntactic patterns**	**Examples**
DEM (pronominal)	+ [*m*-AUX CLAUSE]$_{\text{REL}}$	*hegai* [*m-o ge'ej*] 'that [that is big]'
	+ POSTP	*'id wui* 'toward this' [lit. this toward]'
	+ POSTP N	*'id wui mi:stol* 'toward this cat' '[lit. this toward cat]'
	+ (OTHER)	--
DEM N (adnominal)	+ [*m*-AUX CLAUSE]$_{\text{REL}}$	*hegai mi:stol* [*m-o ge'ej*] 'that cat [that is big]'
	+ (OTHER)	--

Table 38. Demonstrative syntactic patterns and examples.

I single out these particular constructions based on their pervasiveness in the data and/or previous observations of their co-occurrence with demonstratives in O'odham (e.g. Saxton, 1982; Zepeda, 1983). This decision was also informed by Diessel's (1999) typological study of demonstratives which highlights similar cross-linguistic patterns to those described above in Table 38 for O'odham. The author's study suggests that demonstratives that occur in these constructions lend themselves to particular expanded grammatical functions.

First, Diessel (1999) notes the cross-linguistic tendency for adnominal demonstratives that co-occur with relative clauses to evolve into determinatives or grammatical markers of the relative clause head (pp. 135-137). At this more advanced stage, the demonstrative may be of either the adnominal or pronominal type, and the demonstrative is an obligatory part of the construction while at the same time being devoid of other pragmatic functions (see §6.1.4). Previous observations for O'odham (e.g. Saxton, 1983; Zepeda, 1983) similarly suggest that demonstratives are a fixed part of relative clause constructions in the language.

Diessel (1999) also observes that demonstratives that co-occur with adpositions may grammaticalize into connectives (pp. 125-127; also see §6.1.4). There is additional evidence

in O'odham of similar constructions also evolving into specialized postpositions which I assign the label 'referring postpositions.' In the O'odham data, demonstratives often occur in juxtaposition with varying postpositions. In these cases, the postposition always immediately follows the demonstrative (i.e. [DEM POSTP]). This is true whether or not the noun to which the demonstrative refers is included in the clause (see a more detailed discussion and O'odham examples in §6.5.1). In cases where a demonstrative and postposition co-occur with the referent, a special construction is employed in which the noun follows the postposition ([DEM POSTP N]). That is, an adnominal demonstrative followed by a postposition is not observed in the data (*[DEM N POSTP]).

Finally, a brief note on the morpho-syntactic variable of subject and object distinctions; Languages which use the same form for both pronominal demonstratives proper and more grammaticalized categories, as O'odham does (see §§6.2), can possibly develop a morphological distinction in form for subject and object pronouns (and potentially further into subject and object markers on the verb or noun; e.g. Givón, 2001). One function may become more phonologically reduced over time so that different forms are eventually employed for the two different grammatical functions (see Bybee, 2015, p. 150 for an example of such a phonological distinction in French, Spanish, and Portuguese). This does not, however, appear to be the case in O'odham. It is possible for either full or reduced phonological forms to function as either subjects or objects (compare, for example, (29)-(31) in §6.2 with (72)-(74) in §6.5.2). I will revisit this issue and offer more concrete findings which support this view in §6.6.4.

6.4.2 Coding for discourse functions

While the construction in which a demonstrative occurs plays a central role in the grammaticalization pathway it may follow, structure is also attached to function, and demonstratives play a number of important functional roles in discourse. Exophoric demonstratives focus the hearer's attention on entities in the surrounding physical situation. Endophoric demonstratives indicate referents mentioned within the text. They serve tracking, discourse deictic, and recognitional functions (described in §6.1.3). Endophoric demonstratives are the slightly more grammaticalized versions of their exophoric predecessors and may extend to more grammatical meanings over time (see §6.1.4). In this regard, the endophoric functions are more relevant to discussions concerning the evolution of demonstratives. That said, I include exophoric demonstratives in my coding and in the descriptive analysis in §6.5 (and, to a lesser extent, in the subsequent quantitative analysis in §6.6). Their inclusion allows me to be transparent about how I distinguish exophoric from endophoric demonstratives in the data. Furthermore, as stated in the introduction in §6.0, discourse-based studies of demonstratives in O'odham are lacking in the literature. Including exophoric demonstratives contributes to our understanding of their appearance as well as where they fit in with the overall makeup of demonstratives in the language.

In what follows, I describe the criteria I employ in order to identify each of the discourse functions noted above. (For examples of each function in O'odham, the reader is directed to §6.5.1). The first step in all cases is to accurately identify the referent. I carried out this initial step in close consultation with native O'odham speaker Mr. Robert Cruz. These discussions were imperative before moving on to all of the other remaining analysis procedures discussed in the subsequent sections.

Exophoric demonstratives

Exophoric demonstratives are the most basic type out of which all other demonstrative functions are hypothesized to evolve (Diessel, 1999, p. 110). To briefly recap, the demonstrative in these cases directs the hearer's attention to an entity that is physically present in the surrounding discourse situation, and the speaker is at the deictic center. The demonstrative may be of the gestural type where the speaker optionally produces a pointing gesture along with the demonstrative which, together, indicate the referent(s). In cases of *Deixis am Phantasma* the speaker may shift their vantage point to an entity within an imagined scenario, for example a character within a story. The speaker treats the imagined referent as if it is located in the physical space surrounding the speaker. Based on these variables, I use the following coding criteria to identify exophoric demonstratives in O'odham:

(48) **Identifying exophoric demonstratives**

- The referent exists in the surrounding situation (literally or imagined).
- The speaker could potentially point to the referent (literally or symbolically).
- Involves a deictic center in space from the vantage point of the speaker (literally or imagined).

Endophoric demonstratives

Any demonstrative that is not employed for the exophoric function is categorized as endophoric. The basic endophoric demonstratives, as described in §6.1.3, are used by the speaker to indicate referents that are stated within the text or verbal utterance itself. They are used to organize discourse, either to refer to larger segments of surrounding discourse (discourse deictic), to track previously mentioned referents (tracking function), or to indicate a referent that the speaker believes is also known to the hearer(s) based on prior, shared experiences or presumed shared knowledge (recognitional).

For each endophoric demonstrative category, I first establish what criteria I employ to identify and isolate the tokens that belong to each (tracking, discourse deictic, and recognitional). After I identify each of the categories based on my coding, I later discuss patterns with respect to function and their co-occurrence with distal versus proximal demonstratives, and pronominal versus adnominal demonstratives and how the patterns compare with previous cross-linguistic tendencies found in Diessel (1999) and Himmelmann (1996) (see Table 39). According to Diessel (1999), endophoric demonstratives are somewhat grammaticalized versions of their exophoric predecessors. As a function grammaticalizes, it may become more strongly linked with a particular demonstrative type or construction of which there are cross-linguistic tendencies.

Discourse deictic function

The discourse deictic use of a demonstrative refers to "propositions" or "aspects of meaning, expressed by a clause, a sentence, a paragraph, or an entire story" (Diessel, 1999, p. 101) rather than to specific nouns or noun phrases. The demonstrative may indicate the referent either anaphorically or cataphorically, and the referent usually does not persist as topical in the discourse; instead, it is referenced once, and then the discourse moves on. Himmelmann (1996) also observes that the demonstrative tends to refer not just to surrounding discourse, but immediately adjacent segments of discourse. Based on these previous observations, I identify discourse deictic demonstratives based on the following criteria:

(49) **Identifying discourse deictic demonstratives**

- The demonstrative refers to an entire proposition or event expressed in a clause, sentence, paragraph, story, etc. either anaphorically or cataphorically.
- The demonstrative refers to a chunk of discourse that is immediately adjacent to it.
- The referent is mentioned only once (i.e. it does not continue as salient).

After I have identified the discourse deictic demonstratives in the data, I examine whether they display a preference for pronominal or adnominal constructions and whether they tend to occur with distal or proximal demonstratives (see Table 39). According to Himmelmann (1996), pronominal demonstratives appear to be the primary device for discourse deixis in the author's language sample (p. 225). Finally, Diessel (1999) finds that, in English at least, anaphoric reference can be expressed with either a distal or proximal demonstrative, while cataphoric reference is expressed solely with a proximal demonstrative (p. 103; see also Fillmore, 1997; Halliday & Hasan, 1976).

Tracking function

With the tracking function, speakers can employ a demonstrative (either adnominal or pronominal) to refer anaphorically to a previously introduced entity in the discourse. Tracking demonstratives are co-referential with a discourse participant that is unexpected and non-topical. In terms of the Accessibility Scale (see §6.1.4), this unexpected referent often holds peripheral or semi-active status. Tracking demonstratives can also mark a topic shift, transitioning a non-topical referent to topical status. Often this occurs with second mentions of the referent. In other cases, the unexpected referent may be one that the speaker mentioned "at some distance in the preceding discourse" (Diessel, 1999, p. 98). It may even be the case that the referent has become inactive in the hearer's consciousness if it has not been mentioned for some time. The tracking device reactivates the referent when it becomes relevant to the discourse once again. In order to identify tracking demonstratives, I follow the criteria in (50) below.

(50) **Identifying tracking demonstratives**

- The demonstrative is co-referential with a previously introduced referent in the discourse text.
- The referent is initially introduced as a noun or noun phrase.
- The demonstrative indicates non-topical – either semi-active or inactive – referents.

In order to indirectly operationalize cognitive statuses of a referent, first, I apply Givón's (1983) referential distance measurement (explained below) to the isolated tokens (i.e. those that express co-reference with a noun or noun phrase initially introduced at some point earlier in the discourse), and secondly, I apply a more fine-grained analysis in certain cases in order to isolate tokens that mark a topic shift.

The measurement of **referential distance** gauges the number of intervening clauses between the demonstrative construction being analyzed and the referent to which it refers in the preceding discourse material (Givón, 1983, pp. 13-14). For example, a previous mention of a referent that occurs in the same clause is assigned a value of 0, a previous mention in the immediately preceding clause is assigned a value of 1, a referent two clauses before is assigned a value of 2, and so forth. While it is not possible to directly assess how accessible a referent is in the hearer's consciousness, this measurement provides an indirect estimation. Referents that are mentioned more recently are assumed to be more active in the hearer's consciousness and those that are mentioned more remotely in time are assumed to be less active, thus indicating an activation status continuum.

In order to detect demonstratives that indicate participants that are non-topical, I count the number of clauses between the demonstrative token being analyzed and the last representation or mention of the referent that the demonstrative refers to. A 'previous mention' refers not only to the original introduction of the participant in the form of a noun or noun phrase, but also to subsequent representations of that referent as pronouns,

indefinite/definite nouns, subject marking on the auxiliary, object indexation on the verb (see §3.3 and 3.6.1 for descriptions of the latter two patterns), and, finally, zero anaphora. I assign the label **'non-topical'** to referents that occur in three or more clauses (**3+ clauses**) in the preceding discourse. **Semi-active referents** appear in the **preceding 3-20 clauses**, and **inactive referents** occur in the **preceding 20+ clauses**. I adopt this value as the upper limit based on Givón (1983). At this point, it is safe to assume that the referent has become inactive in the hearer's consciousness. Clauses include independent, subordinating, and coordinating types (see descriptions for O'odham in Hale, 1983; Payne, 1987; Zepeda, 1983).

I make one minor adjustment to these coding procedures. As I discussed in §6.1.3, one possible function of demonstratives is to mark a topic shift. This often occurs with the second mention of the referent (Diessel, 1999; Himmelmann, 1996). In some instances, this 'second mention' may even occur in the second clause after a referent has been introduced to the discourse. Given this possibility, the referential distance measurement would exclude such cases from the count of tracking demonstratives in the findings. In order to avoid this exclusion, I separately examine second mentions occurring in the second clause and determine if the demonstrative indicates a topic shift (see example (8) in §6.1.3). If they do indicate a topic shift, these cases also fall under the tracking demonstrative classification.

Finally, after this subset of tracking demonstratives is established, I consider whether the tokens tend to more strongly associate with pronominal or adnominal, distal or proximal demonstratives (see Table 39). Regarding the former, previous findings suggest a preference for adnominal constructions (Himmelmann, 1996) and in relation to the latter, languages tend to show a preference for one demonstrative type – distal or proximal in the case of O'odham – over another (Diessel, 1999).

Recognitional function

As discussed in §6.1.3, recognitional demonstratives signal to the hearers that they are expected to have prior knowledge of the referent, but the referent is inactive at the moment it is introduced into the discourse. The recognitional demonstrative is also a tool with which the speaker can assess whether or not the hearer recognizes and accurately recalls a referent. The speaker may use additional comments in the form of relative clauses to help the hearer accurately recall the referent. Other signals may also suggest that the speaker is uncertain if the hearer remembers the referent and is attempting to ascertain if they do (e.g. hesitations, tag questions, etc.). Based on these observations, the following criteria are used to identify the recognitional use of demonstratives in O'odham:

(51) **Identifying recognitional demonstratives**

- The demonstrative introduces a referent that is new to the discourse but known to the interlocutors.
- The speaker adds remarks which suggest that the speaker is reminding hearer(s) of the referent (e.g. added explanations or descriptions of referent, relative clauses, etc.)
- Possible indications that the speaker is trying to gauge the hearer's accurate identification of the referent (e.g. tag questions, pauses, hesitations, etc.)
- Once the identity of the referent has been successfully established, the discourse moves on (i.e. referents do not continue as major discourse participants).

Finally, as I do with the other categories, I examine which constructions and demonstrative types the recognitional function tends to co-occur (see Table 39). This will tell us how similar the patterns in O'odham are to those found cross-linguistically (see Table 39). Diessel (1999) states that adnominal demonstratives tend to be used for the recognitional use, and Himmelmann (1996) finds evidence for distal demonstratives.

6.4.3 Identifying more grammaticalized demonstratives

While tracking, discourse deictic, and recognitional demonstratives are considered grammaticalized to some extent, they can advance further, expanding towards more grammaticalized uses (see §6.1). Based on previous discussions (Hale, 1959; Langacker, 1977; Mason, 1950; Mathiot, 1973; Saxton, 1982; Zepeda, 1983) and/or evidence in the O'odham data, the primary grammatical functions of demonstratives beyond the more basic endophoric types described in the previous section and which I further investigate are **third-person pronouns**, **specific indefinite articles, determinatives**, **connectives, referring adpositions,** and **placeholders**.[36] The goal of this section is to present the criteria I use to identify each subset of these demonstrative uses in the data.

Third-person pronouns

O'odham employs the same form for both pronominal demonstratives proper and third-person pronouns, as described in §6.2. Third-person pronouns are hypothesized to evolve from pronominal tracking demonstratives which mark referents that are non-topical; that is, they frequently have semi-active or, in some cases, inactive status. Over time, pronominal demonstratives may move up the Accessibility Scale to indicate active referents that are at the center of the hearer's awareness in the discourse (see (11) and Figure 12 in §6.1.4). At this stage, it can be said that pronominal demonstratives have extended to third-person pronouns. Third-person pronouns have the identifying features listed in (52) below.

[36] Articles (i.e. noun-markers) are excluded (see §6.2).

(52) **Identifying third-person pronoun uses of demonstratives**

- The demonstrative is pronominal.
- The demonstrative is co-referential with a noun or noun phrase appearing at some prior point in the discourse.
- The referent carries active status.

I indirectly determine activation status of the referent as I did with tracking demonstratives, applying Givón's (1983) measurement for referential distance. First, I identify all pronominal demonstrative tokens which indicate a referent that occurs within **0-2 preceding clauses**. Next, I exclude all cases which indicate a topic shift (see discussion for tracking demonstratives in §6.4.2). What remains I interpret as referents with active status and as third-person pronouns. Again, 'previous mentions' may take the form of noun or noun phrases, pronouns, subject marking on the auxiliary, object indexation on the verb, and zero anaphora.

After establishing this subset of tokens, I note whether both distal and proximal demonstratives may play this function in the descriptive portion of the study (§6.5.2), as Mathiot (1973) and Saxton (1983) suggest for O'odham (see §6.2).

Next, I closely examine correlations between the identified tokens and phonological form (i.e. full or reduced) in the quantitative portion of this study (§6.6.4). The expectation is that singular demonstratives will occur more frequently in their reduced form (plural forms, on the other hand, are nearly always in their full forms). In other words, phonologically reduced forms will be more strongly associated with a third-person pronoun meaning. This is based on the cross-linguistic tendency for phonologically reduced forms to correlate with a narrowing in meaning (Bybee et al., 1994, p. 19; Diessel, 1999, p. 118; see §1.1). This finding would additionally suggest a developing morphological split in form. The other assumption is that this pattern would strengthen over time in the data based on previously

observed crosslinguistic grammaticalization trends. These details may shed light on the extent to which pronominal demonstratives in O'odham have grammaticalized and extended to third-person pronoun meanings. If the trends show patterns that differ from the expectation, some possible explanations can be explored that may or may not be attributable to language shift in O'odham.

Specific indefinite articles

Some languages have an article that specifically marks specific yet indefinite referents – that is, referents that are inactive and new to both the discourse and hearer (Diessel, 1999; Wright & Givón, 1987). Once the new, specific referent has been introduced to the discourse, it may continue as salient. While "standard" English does not have a specialized, morphologically distinct lexical item for this function, in colloquial English the adnominal, proximal demonstratives, *this* and *these* show evidence of having grammaticalized to play this role in this type of discourse (see English examples (12)-(13) and discussion in §6.1.4). Similarly, O'odham does not have a separate, specialized mechanism for marking specific, indefinite referents; however, some preliminary evidence suggests that like English, O'odham may also on occasion employ the proximal demonstratives *'i:da* 'this' and *'idam* 'these' in an adnominal construction for this function. The criteria I employ in order to isolate a possible subset of specific indefinite articles in the O'odham data are included in (53) below.

(53) **Identifying specific indefinite articles**

- The referent is new to the discourse (i.e. inactive) and unknown to the hearer.
- The demonstrative is accompanied by additional modifying information, such as a relative clause.
- Once the referent is introduced, it continues to be "thematically important" (Wright & Givón, 1987, p.16).

After a subset is established with the above criteria, I examine possible correlations with demonstrative type based on the fact that English shows a preference for the proximal demonstrative. If O'odham shows the same pattern, this could indicate a contact effect. This will be explored in §6.5.2. Since articles always occur in adnominal constructions, I assume that the demonstrative will appear in its phonologically full form as is nearly always the case with this construction.

The significance of this potential development in O'odham raises some interesting issues given that languages that make this distinction usually utilize a *separate* article for this function – one that develops from the numeral one (Diessel, 1999, p. 138 based on Givón, 1995; Wright & Givón, 1987). In colloquial English, however, the source is clearly the proximal, adnominal demonstrative 'this'. Some preliminary evidence suggests that this may also be the case in O'odham. If a number of similar features between O'odham and English appear in the data – a preference for adnominal, proximal demonstratives, particularly in either the midcentury and/or modern data, this could indicate that language shift and contact may have played a role in this development.

Determinatives

To recap, adnominal demonstratives that are employed for the recognitional function may grammaticalize into an obligatory determinative or a relative marker. At this stage, the semantic content associated with the recognitional or other discourse pragmatic functions has been bleached (Diessel, 1999, p. 108). The determinative demonstrative in an adnominal construction serves the sole purpose of marking the nominal head of the relative clause, and in instances where the demonstrative is independent (i.e. pronominal), it functions itself as the head of the relative clause (p. 135). In most cases, the referent is usually a first mention;

it carries inactive status, just as its recognitional demonstrative predecessors do. Furthermore, as I argued n §6.1.4, at the determinative stage, the speaker does not appear to have a specific referent in mind but rather a referent subtype or category that is defined by the relative clause.

In O'odham, relative clauses take the form [*m*-AUX CLAUSE]$_{REL}$, which is always preceded by a demonstrative (see §6.4.1). This pattern, which can be employed for the recognitional function, may later expand to the more grammaticalized determinative use, at which stage the demonstrative is simply a marker of the relative clause – either modifying the head or functioning as the head of the relative clause itself.

In order to identify determinative demonstratives, I code for the following criteria listed in (54).

(54) **Identifying determinative demonstratives**
- The adnominal or pronominal demonstrative is followed by [*m*-AUX CLAUSE]$_{\mathbf{REL}}$.
- The referent is a first-mention (i.e. inactive).
- The referent continues as non-salient in the discourse after first mention.
- The demonstrative indicates a referent *subtype* that is established in the adjoining relative clause.

Finally, I consider whether determinatives tend to occur with distal versus proximal demonstratives, pronominal versus adnominal demonstratives and whether they tend to occur with the full versus phonologically reduced demonstrative forms.

Connectives

Pronominal demonstratives that co-occur with an adposition or adverb can together evolve into connectives. These probably evolve from pronominal demonstratives that are used for discourse deixis (Diessel, 1999, p. 125). Similar to discourse deictic demonstratives, connectives conjoin two adjacent chunks of discourse and they indicate "the semantic

relationship between the conjoined propositions" (p. 125). Over time, the pronominal demonstrative, together with the postposition or adverb, may show signs of fusion and, in turn, constructional change. Some of the meanings that Diessel has found these types of constructions to convey are *therefore, and/so then, and that, after that, then,* and *about that/it* (see Diessel, 1999, pp. 125-127).

Notably, a pervasive pattern exists in O'odham where a reduced form of a proximal or distal demonstrative is immediately followed by a postposition (either [DEM POSTP] or [DEM POSTP N]). These syntactic patterns were described in §6.4.1. In some cases, these constructions clearly convey a connective meaning.

In order to isolate the contexts in which a demonstrative functions as a connective, I identify demonstratives that are immediately followed by a postposition and also have discourse deictic-like functions where the construction appears to link one proposition to another one that either precedes (anaphoric) or follows (cataphoric) it. Finally, I consider whether the word gloss and/or free translation given for a construction expresses a connective meaning. These identifying criteria are summarized below.

(55) **Identifying connective DEMs**
- The demonstrative is followed by a postposition.
- The demonstrative and postposition connect two chunks of discourse.
- The glossing/translation indicates a connective meaning (e.g. therefore, because of, etc.) between two propositions.

I also note possible relationships between connective demonstratives, demonstrative type, and phonological form in order to assess if one type or form tends to be more strongly associated with this function over time.

Referring postpositions

In O'odham, another pattern that emerges with demonstratives that are followed by various postpositions is the possibility for them to convey specialized, postpositional meanings. Like connectives mentioned above, a postposition follows a demonstrative (either [DEM POSTP] or [DEM POSTP N]), but in these cases the demonstrative indicates a specific referent in the form of a noun, frequently in much the same way that tracking demonstratives do, based on examples in the O'odham data (see §6.5.2). For this reason, I call these cases "referring postpositions." This is a development that Diessel (1999) only briefly alludes to in passing (p. 126). More specifically, Diessel mentions the development of pronominal adverbs in German, which are composed of older demonstrative forms and prepositions. The author otherwise does not discuss these construction types as a possible source for the development of special adpositional meanings. Their existence in O'odham could suggest that this is a wider-spread phenomenon than previously assumed. The criteria I use to identify referring postpositions in O'odham is outlined below.

(56) **Identifying referring postpositions**

- The demonstrative is followed by a postposition.
- The demonstrative indicates a specific referent in the form of a noun of noun phrase.
- The glosses/translations suggest postpositional meanings.

As with the previous grammatical categories, I also consider their relationship with demonstrative types and also note their phonological form.

Placeholders

Finally, pronominal demonstratives may function as placeholders at a moment when a speaker experiences "word-formulation trouble" (Hayashi & Yoon, 2010, p. 33) or difficulty

retrieving a specific lexical item. Often the lexical item that the speaker is in search of is a noun. With the placeholder use, the demonstrative serves as a stand-in until the speaker is able to recall the lexical item. The placeholder function is particularly salient in the context of O'odham given that speakers of the language today have fewer opportunities to use it which could hypothetically result in lexical retrieval challenges. The criteria I employ for identifying placeholders are summarized in (57) below.

(57) **Identifying placeholders (based on Hayashi & Yoon, 2010; Podlesskaya, 2010)**

- The demonstrative serves as a substitute for a specific lexical item that the speaker has difficulty retrieving.
- The speaker is eventually able to recall the lexical item.
- The demonstrative occupies the same syntactic slot where the specific lexical item would have been uttered.
- The demonstrative mirrors the morphology (e.g. number) of the intended referent.

Finally, I look at whether or not placeholder demonstratives show any clear relationships with a type of demonstrative (proximal or distal) and the phonological form of the demonstrative.

6.5 Results: Descriptive analysis

In §6.4, I laid out the coding criteria I use in order to identify the different possible constructions associated with demonstratives, the possible discourse pragmatic functions, and the more grammaticalized uses of demonstratives in the O'odham data. In what follows, I provide a detailed descriptive analysis of the major patterns uncovered in this study based on the coding criteria laid out in §6.4. The aims in doing so are twofold. First, a detailed discussion of numerous examples provides additional transparency regarding how I identified different demonstrative functions in the data based on my coding. Secondly, this analysis reveals the range of functions demonstratives have in O'odham based entirely on their use in

authentic discourse, an area of the language that has up until this point been largely under-described. I discuss relevant quantitative distributions of patterns with regard to the historical, midcentury, and modern data in §6.6.

6.5.1 Description of exophoric and basic endophoric discourse functions

In this section, I provide a detailed account of the primary discourse functions of demonstratives based on the coding criteria stated in §6.4.2. More specifically, the examples illustrate the exophoric and basic endophoric functions (described in §6.1.3) that occur in the O'odham data used for this study. More grammaticalized uses of demonstratives will be treated separately in §6.5.2.

Exophoric demonstratives

The primary purpose for including exophoric demonstratives in the descriptive portion of the analysis is to illustrate how I distinguish them from endophoric demonstrative types. This analysis also contributes to our overall understanding of how demonstratives are used in O'odham. They figure less prominently in my analysis of grammaticalization trends which I address in §6.6. Exophoric demonstratives are the likely source out of which all other demonstrative uses evolve (Diessel, 1999, pp. 110-114). I assume that O'odham is no exception. They are employed to focus the hearer's attention on entities that exist in the surrounding, physical situation – either real or imagined (see §6.1.3). They may be of the gestural type where the speaker is physically at the deictic center and may possibly use a pointing gesture to indicate some referent, or the demonstrative may be used to express *Deixis am Phantasma* where the speaker shifts their viewpoint as if they are located in an imagined space, for example, in a story that is being told. (See (48) in §6.4.2 for coding guidelines).

Gestural examples

Several examples below from the O'odham data illustrate the exophoric use of demonstratives in the language. The first example, which is from a YouTube video, is gestural. In it, the speaker is in attendance at an event taking place at the time of the recording. The event celebrates the elders of the O'odham community who are also present. The speaker is in the midst of describing how and why the event came to fruition. At one point in the discourse segment shown in (58) below, the speaker refers to *'idam 'al elders* 'these little elders'. In the video, the speaker is seen pointing in the general direction of the event and where the elders are assumed to be located.

(58) … k ab cu'ig on there m-at 'an hig o hema
CONJ be.located on there SBDR-3.AUX.PFV LOC should FUT one

jegel-s g taṣ m-at 'am ab o
make.room.for-DUB(?) ART day SBDR-3.AUX.PFV LOC thus FUT

'e-ju: m-at 'am o ha-he:kegcul
REFL-make.PFV SBDR-3.AUX.PFV LOC FUT 3PL.OBJ-honor.PFV

'idam 'al elders.
PROX.PL little elders.
'…it is in there [a written proclamation] that they should set aside a day that will be made to honor these elders. (*#1*: 14)

It is clear in this case that the speaker employs the adnominal demonstrative construction to exophorically refer to the elders based on the fact that both the speaker and the elders are present at the event as well as the fact that the speaker is using a pointing gesture.

Deixis am Phantasma examples

The next two examples illustrate *Deixis am Phantasma,* both of which involve narrators who use the demonstrative as if they are present at the different scenes that they are describing,

when in fact they are not. In example (59) the narrator, Jose Pancho, is re-telling a story about a cat who is watching over his owner's bar where whiskey is sold. Some other animal friends, including Coyote, visit the cat while the owner is away. They are all sitting in the bar when Coyote requests that the cat pour them some *hegai wiñu* 'that whiskey', using a distal, adnominal demonstrative to indicate the referent. The narrator tells this portion of the story from the viewpoint of Coyote who is at the imagined deictic center. In the imagined space, the whiskey is situated in front of Coyote. One can even imagine the possibility of the narrator pointing to the space before him for dramatic effect when he quotes Coyote, pretending as if the speaker, himself, were Coyote referring to the whiskey. In the second example, (60), the speaker is talking about a time during his childhood when his uncle asks him to bring him an object. He does not know the word that his uncle uses, however, and is unsure of what he wants. The speaker describes what he does in response. He picks up different objects and asks his uncle, "*No ḍ 'i:da*" "'Is it this?"', using the proximal, pronominal demonstrative.

(59) … c 'at-p 'am hab haha 'i 'a:g hegai mi:tol
CONJ 3.AUX.PFV-ASSUMP LOC thus then PUNC say.IPFV DIST cat

"Pt 'ab hig o 'al hi a t-was
2SG.AUX.PFV LOC should FUT little maybe part 1PL.OBJ-DIP.OUT.PFV

hega'i wi:ñu tt o 'i:."
DIST **whiskey** 1PL.PFV FUT drink.PFV
'… Then Coyote coaxed the cat and told him: "Please pour us a little of that whisky for us to drink."' (*Coyote Cat*: 30-31, 50, 59)

(60) Nt pi ma:c hascu, ñ 'aṣ nanko ha'icu
1SG.AUX.PFV NEG know.IPFV what, 1SG.AUX.IPFV just various things

'an 'aṣ si bebbhe, "N-o ḍ **'i:da**?"
LOC just INTS get.REP INTER-3.AUX.IPFV COP **this**
'I didn't know what it was [that he was asking for] and so I would just grab various objects and ask, "Is it this?"' (*#3*: 276-278)

Both cases are classic examples of *Deixis am Phantasma* where the speakers are narrating a story as if they are physically located in the scenes that they describe. In example (59) the speaker uses a distal adnominal construction and in (60) the speaker employs a proximal pronominal demonstrative.

Endophoric demonstratives

The following series of examples demonstrate the basic endophoric functions of demonstratives in O'odham, which are the discourse deictic, tracking, and recognitional functions. At this stage, the demonstratives do not refer to the surrounding physical situation (literally or in the imagination) but instead to a referent that is stated within the text itself, and the demonstratives function to organize discourse in different ways. All endophoric demonstratives likely evolve from the exophoric type (see §§6.1.3 and 6.1.4).

Discourse deixis examples

In the case of discourse deixis, the demonstrative refers either anaphorically or cataphorically to an entire proposition or event that is expressed in an immediately adjacent segment of the text. The discourse deictic usage conjoins the ideas conveyed in one section of text to another (see §6.1.3; also coding procedures in (49), §6.4.2). Either adnominal or pronominal demonstratives may be employed, but there is a possible crosslinguistic tendency for the latter (Himmelmann, 1996, p. 225). Additionally, it has been observed in English that either distal or proximal demonstratives can be used to convey anaphoric reference while proximal demonstratives are preferred for cataphoric reference (Diessel, 1999, p. 103; based on Fillmore, 1997; Halliday & Hasan, 1976).

The first two examples illustrating the discourse deictic usage are **anaphoric**. In (61), the pronominal, distal demonstrative *heg* 'that' is employed to refer anaphorically to the text

in the previous line; that, as a boy, the narrator of the story had to ride behind someone else on the horse's back since he did not have his own horse to ride. The narrator reflects back on this and concludes that he probably did not like *heg* 'that', suggesting that it was unpleasant to ride on a horse behind someone else. The narrator then explains that he would get off of the horse and run and then the story moves on to another topic. The demonstrative, in other words, is not used to mark a referent as topical or central to the story but instead to indicate an additional small detail of the narrative before moving on.

(61) Ñḍ pi'a ṣoiga c 'an ' aṣ
1SG.AUX.IPFV-RMT NEG domesticated.horse CONJ LOC just
I did not have a horse and there just

ha-'oidc o daicug-ad. Ñ ṣ-p
3PL.POSS-behind FUT ride-FUT.IPFV 1SG.AUX.IPFV QUOT-ASSUMP
somebody's behind will be riding. I perhaps

heg s-'o:hod,
DIST STAT-get.tired.of
that disliked.
'I didn't own a horse, and I just rode behind someone else. I probably disliked that.'
(*Antonio*: 48-49)

The second example below in (62) utilizes the pronominal, proximal demonstrative *'i:da* 'this', indicating that both demonstratives types may express anaphoric discourse deixis. The demonstrative in this example points back, in this instance, to an entire story that the narrator was just telling. The demonstrative refers to the immediately adjacent text which includes all of the events leading up to the excerpt in (62). The example is followed only by one other line that essentially states, "The end." The referent is clearly not topical and is only mentioned one time before the story ends.

(62) Ñia hab masma hab 'e-ju: **'i:da**
see.DM in.this.certain.manner thus 3.REFL-happen.PFV **PROX**

kc 'i a 'i hugk 'i:ya'a.
CONJ here really PUNC be.at.the.end here.
'So, this is what happened [and] this is the end of the story. (*Coyote Cat*: 33, 52, 60)

In both of these cases the demonstrative is pronominal, which, according to Himmelmann (1996) is strongly associated with the discourse deictic function (p. 225). This is true for the most part in the O'odham data, as well, although on occasion, an adnominal demonstrative is possible, as in (63) below. In this example, the speaker is explaining that there are many O'odham who are protective or cautious of their language. In the next line, the speaker refers anaphorically to this statement as *hegai attitude* 'that attitude' and then again fully in English as *that language attitude.* The speaker then goes on to state that O'odham people have a right to their secret language, and then slightly shifts the topic and describes the types of language use that should not be put on the internet in order to protect the information from outsiders; that is, the speaker introduces the "language attitude" concept and moves on.

(63) Ha'u, 'am a weho m-o-- m-o s-da'ag
yes LOC indeed true SBDR-3.AUX.IPFV SBDR-3.AUX.IPFV STAT-attached.to

O'odham g 'e-ñe'ok, kc 'aṣhaba, **hegai attitude** uh tha-- h--,
O'odham ART 3.REFL-language, CONJ but, **DIST attitude** uh tha-- h--,

that language attitude uh, hab a 'am a s-'ape.
that language attitude uh, thus PART LOC PART STAT-good.
'Yes, it is true that-- that O'odham are cautious about their language, but, that attitude, uh, that language attitude, it is OK. (*#3*: 565-569)

The next discourse deictic example is **cataphoric**. In the example, the narrator explains that his father can't remember the details of a specific war that occurred in his father's lifetime because he was too young. What he can remember, however, are the long

distances he had to walk behind his older brother. The speaker uses the demonstrative *hega'i* to verbally point ahead to this statement which appears in the adjacent clause.

(64) **Hega'i** s-cegito, m-at hebai me:k k 'an
DIST STAT-remember SBDR-3.AUX.PFV somewhere far CONJ LOC
that only he remembers that they somewhere hihi [sic] and there

'e-si:s 'oid.
3.REFL-older.brother behind
his older brother behind.
'[My father does not remember this war.] He only remembers how far he walked behind his older brother [i.e. he only remembers that: how far he walked…].'
(*Antonio*: 222-223)

The historical example departs from the findings attested for in English, at least, [37] in that a distal demonstrative is used while in English proximal demonstratives are preferred for cataphoric reference. Proximal demonstratives, however, can also be used as a cataphoric device, as shown in the modern example (65) below. In this case, *'i:da* 'this', points ahead specifically to what the O'odham people will do with their language. They will speak, write, and communicate in O'odham with computers. The speaker is making the case that it is up to the O'odham community to use the language in these capacities in the future. The speaker utilizes the demonstrative to cataphorically refer to the entire proposition included in the adjacent sentence.

(65) 'A:pim O'odham, 'am hab o ju: **'i:da**.
2PL O'odham LOC thus FUT do.PFV **this**.

'A:pim ḍ 'em-ñe'ok, mt 'a:pim hab o 'o'ohañ
2PL COP 2PL.POSS-language 2PL.AUX.PFV 2PL thus FUT write.IPFV

[37] Diessel only provides English examples for cataphoric reference.

c o ñeok-ad, o we:maj ñeo g computer…
CONJ FUT speak-FUT.IPFV FUT together.with speak.PFV ART computer
'You O'odham will make this happen [lit. do this]. It is your language that you will write it and speak it and will communicate with the computer…' (#3: 584)

Pronominal, discourse deictic demonstratives may gradually develop into connectives, especially when they co-occur with other lexical items such as adpositions or adverbs (Diessel, 1999, pp. 125-127). Given that discourse deictic demonstratives function to conjoin two different propositions in discourse, it is hypothesized that over time pronominal demonstratives that co-occur with adposition or adverbs can gradually develop a connective function. They may be interpreted as carrying meanings such as *therefore, because of, as a result of,* etc. There is also evidence in O'odham that demonstrative-postposition combinations may evolve into specialized referring postpositions. I will address this topic at a later point.

Tracking examples

Tracking demonstratives anaphorically indicate a non-topical referent that carries semi-active or inactive status in the hearer's consciousness. Tracking demonstratives may mark a topic shift as well as reactive referents that have not been mentioned for some time in the discourse (see discussion in §6.1.3 and coding in (50), §6.4.2). Based on previous observations (Diessel, 1999; Himmelmann, 1996), either pronominal or adnominal demonstratives may serve the tracking function, but Himmelman's findings suggest a possible preference for adnominal constructions. This appears to be the case in O'odham, as well, as illustrated in the examples that follow. Cross-linguistically, languages also appear to show a preference for one demonstrative type over an another – in this case, either distal or proximal. While proximal demonstratives are possible in O'odham, the distal type is more pervasive, although

this may be related to the fact that distal demonstratives are, overall, more frequent in O'odham (see Figure 14 in §6.6.2).

In the first O'odham example of a tracking demonstrative, a mother is talking to her son or *'e-maḍ* 'her child' or 'child of hers'. She explains to him why his father is no longer alive – because somebody killed him. At this point, while the hearer likely has an open discourse file for the boy since he was mentioned recently, he is only semi-active as the mother talks about the father. Then in the final clause of the segment, the narrator uses an adnominal, distal construction to anaphorically refer to the boy once again.

(66) Ku-ṣ wenog m-a-ṣ 'am hab 'i 'a:gid-a-him
CONN-QUOT then SBDR-3.AUX.IPFV-QUOT LOC thus PUNC tell-EP-PROG

g 'uwi **'e-maḍ**, "M-am-t 'ab ju:
ART woman **3.POSS-child** SBDR-2PL.AUX.PFV[38] LOC do.PFV

g m-o:g k 'ab mea."
ART 2SG.POSS-father CONN LOC kill.PFV

Ku-ṣ 'am haha 'i ṣoak-i-him **hegai** **wiappo'oge'el**.
CONN-QUOT LOC then INCEP cry-EP-PROG **DIST** **boy**
'And so it is said then the woman was telling her child, "That's what somebody did to your father, he was killed." And so then that boy began crying. (*Boy:* 43-45).

In fact, the tracking function indicates a topic shift to the boy. In the subsequent lines (not shown here), we are told that upon learning the news about his father, the boy declares that one day he will avenge his father and that he begins training himself to be a fast runner and hunter in preparation. The entire story, in fact, is primarily about how the boy grew up to be a warrior.

[38] Carries the meaning of 'someone' or 'somebody' in this case.

The next set of examples, (67)-(69), show several different yet related discourse segments from a narrative. The segments illustrate how a demonstrative can be used as a discourse tool to track referents throughout discourse. In the lines leading up to what is included in (67) below, the narrator is talking about an O'odham man named Sohonimhimḍam who was captured by some Mexican soldiers. The captive is staring at the ground, scraping it with a stick. While he scratches around, he finds a stone, which the narrator introduces for the first time as *g hodai* with the article *g*.

(67) 'am **g** **hodai** hema 'i wu:ṣad k pi hab masma
LOC **ART** **stone** one PUNC take.out CONJ NEG in.this.certain.manner
there the stone one he took out and not the same as

g hohodai 'amai. Wa cem pi ge'ej haṣaba s-we:c.
ART rock.PL right.there. although NEG large but STAT-heavy.
the rocks there. although not large but was heavy.
'He uncovered one stone and it was not like other rocks from there. It was not large, but it was heavy.' (*War*: 212-213)

At this point, the hearer does not have a previously opened discourse file for the referent, *g hodai*. The introduction of the referent activates the file. The narrator then briefly describes some features of the rock. In the immediate subsequent lines not included here, the narrator talks about how, at this time, gold had not been seen anywhere, but Sohonimhimḍam believes what he found may be gold. He digs a hole and buries the rock. Next, the narrator describes how the soldiers were preparing to leave and take Sohonimhimḍam away with them as reflected in (68) below. At this point, the rock has been backgrounded and is semi-active. The narrator then briefly brings up the semi-active referent again when Sohonimhimḍam notes where he buried it, referring to it with the distal, adnominal construction *hegai s-namkigam hodai* 'that valuable rock'.

(68) 'I ma:si t o hihi g ṣontal k o
PUNC morning 3.AUX.PFV FUT go.PL.PFV ART soldier CONJ FUT
in the morning they will go the soldiers and will

'i bei g Sohonimhimḍam k
PUNC take.PFV ART Sohonimhimḍam CONJ
take aways the Man who walks like he's going to get married and

'am si 'a'amicud g jeweḍ mo hebai
LOC INTS figure.out.IPFV ART ground SBDR-3.AUX.IPFV where
there he made a careful notice of the ground where

ke:k g 'al 'u:ṣ c 'am heg weco hiaṣp
stood.IPFV ART little bush CONJ LOC that under bury.IPFV
stood the little bush and there that under it is buried (cont'd)

hegai s-namkigam hodai.
DIST STAT-valuable rock
that valuable rock.
'In the morning the soldiers were leaving and taking Sohonimhimḍam and there he carefully noted the ground where the little bush stood and there under it is buried that valuable rock.' (*War*: 217)

The story then takes another shift. The soldiers take Sohonimhimḍam away, he eventually escapes when the soldiers are not paying attention, and he then returns home. The *hodai* 'rock' is not brought up again for over 20 lines and numerous clauses. The referent essentially obtains inactive status and the story shifts to the extent that the narrator's audience unlikely knows whether or not the rock will ever re-emerge in the story again.

Up until this point, the narrator has been telling the story from the perspective of his grandfather. He then switches perspectives, telling it from his own vantage point. He eventually tells us about a time when he was younger and about a little old man who used to frequently visit his home as explained in the segment in (69) below.

(69) Ḍ g 'al ke:li, 'ia, jijiwa t-ki: 'am.
COP ART little old.man here arrive.REP.PFV 1PL.POSS-home LOC
there was a little old man here repeatedly he arrived our home there

Ḍ o ge cuhug 'e-je:ñgĭ g O'odham.
RMT FUT all night 3.REFL-discuss.at.meeting.PFV ART people
they would all night smoke and talk the people

Hebai wo hihi hasko, go:k 'o waik
sometimes FUT go.PL.PFV somewhere two or three
sometimes they would go somewhere two or three

taṣ 'ab. Gmhu 'i me:ko,
day(s) upon over.there PUNC move.far.away
days upon time went far

ñ hahawa ka: m-o wenog cem
1SG.AUX.IPFV then hear SBDR-3.AUX.IPFV then PART.try.in.vain
I then heard that in that time they were trying to

ga:g **hegai hodai** m-at 'am hiaṣ 'i:da 'al ke:li.
look.for.IPFV **DIST** **rock** SBDR-3.AUX.PFV LOC bury.IPFV this little old.man
find that rock that there buried this little old man
'A small elder man at this location would come to our home. The O'odham people would meet all night. Occasionally, they would go somewhere for two or three days. For a long while, and I then heard at that time, that they were trying to look for that rock that this little old man had buried.' (*War*: 240)

The narrator learns that members of his household were helping the little old man locate *hegai hodai* 'that rock' that he had buried a long time ago. The demonstrative is used to reactivate the referent that has otherwise become non-topical and inactive. The discourse segments in (67)-(69) illustrate how the demonstrative is co-referential with a previous mention of a referent, in one case as a semi-active referent, and then later to reactivate the previously introduced referent that has not been mentioned for some time in the narrative and in this manner is the referent that is unexpected.

Another example of the tracking function is also included in the segment in (69) above. This instance illustrates a topic shift with a second mention of the referent. The

narrator, as noted earlier, relays this portion of the story from the perspective of himself as a younger boy who at the beginning of the segment, we are told, meets *g 'al ke:li* 'a little old man'. The boy does not know who the man is. At this point, the audience also does not know who he is either, and in this sense, *g 'al ke:li* 'a little old man' is essentially a first mention.

The narrator introduces the new discourse participant with the article *g* and then later for the second mention, the narrator uses the demonstrative *'i:da 'al ke:li* 'this little old man' who at this juncture is non-topical given that the rock is currently of primary concern. In this instance, the second mention also signals a topic switch. The narrator finally reveals in the subsequent line that the old man is, in fact, Sohonimhimḍam. The story then briefly focuses on Sohonimhimḍam, and we are told about the difficulty he has in recognizing the land and is unable to locate the rock that he buried many years ago.

Recognitional examples

Recognitional demonstratives signal a known referent that has not been previously introduced in the discourse; it is inactive the hearer's consciousness. Commonly, speakers will add additional clarifying information in the form of a relative clause in order to gauge if the hearer is able to correctly recall the referent. The speaker might also use hesitations, pauses, tag questions, etc. which reveal the hearer's uncertainty about whether or not the hearer is able to retrieve the expected known referent (see full description in §6.1.3; coding criteria are listed in (51), §6.4.2). According to Diessel (1999), adnominal demonstratives are strongly associated with this function. Himmelmann (1996) also observes that distal demonstratives appear to be more common. O'odham tends to follow the same patterns, although there are exceptions which I will address below.

In the first example below from a historical text, the speaker appears to use a distal, adnominal demonstrative to signal that he assumes the hearer is familiar with the referent, which has not been mentioned prior in the discourse. The speaker assumes that any O'odham person who lived during the early 1900s will know about the O'odham people who reside at the location named *Gi:ki Wo'o*. After introducing the referent, the speaker then immediately continues with the story; the referent does not become a salient topic in the narrative.

(70)	Ḍ	wenog	'ab	'i	ji:a	**hegam**	**Gi:ki**	**Wo'o**	**'ab**	**O'odham**.
	COP	then	LOC	PUNC	arrive.PFV	**DIST.PL**	**Plow**	**Lying**	**LOC**	**people**
	it was	then	there	made	homes	those	Plow	Lying	at	people

'Then is when those Gi:ki Wo'o O'odham moved there.' (*Antonio*: 138)

Example (71) below illustrates an exception to the claim that it is most typically adnominal demonstratives that are employed for the recognitional function rather than the pronominal demonstrative (Diessel, 1999, p. 105). In this example, the speaker is being recorded for a video that will be posted on a YouTube channel geared towards an O'odham audience. The primary speaker in the video (i.e. the interviewee) has just been asked with whom he has been talking in the community by the interviewer, who is also a member of the O'odham community. The speaker answers the question and introduces a new referent to the discourse with the distal demonstrative *hegam* 'those'. The speaker does not immediately reveal who *hegam* refers to. Instead the speaker first provides the location of where *hegam* 'those' are situated – at the Gila River Governing Center. Presumably, the speaker does this to help his audience correctly identify the referent once he finally states who it is. This also suggests that the speaker presumes his audience will know the referent once he overtly introduces it. The speaker then repeats the demonstrative, *hegam*, followed by a relative clause ([*m*-AUX clause]$_{REL}$) which includes explicit information about *hegam* – the Four

Tribes Culture Resource Protection Committees. Note also the many hesitation markers (e.g. ah, eh, etc.) that are peppered throughout this short excerpt, suggesting that the speaker may be gauging whether at least the interviewer is able to correctly identify the referent. The hesitations may also suggest that the speaker is searching for the best way to introduce the referent, making it as identifiable as possible to the O'odham public audience. Finally, given that the speaker is addressing the O'odham public, it is likely that the speaker presumes that his audience shares his knowledge of the committees he speaks of.

(71) K 'amjeḍ **hegam** 'ep ha-ñei 'att gamhu ah,
CONJ then **DIST** also 3PL.OBJ-see.PFV 1PL.AUX.PFV over.there HES,

Akimel, uh, Gila River Governing Center, **hegam** m-a-ṣ
'River, uh, Gila River Governing Center, **DIST** SBDR-3.AUX.IPFV-QUOT

ḍ eh, Four Tribes ah, ah, Culture Resource Protection, ah, Committees.
COP HES, Four Tribes HES, HES, Culture Resource Protection, HES, Committees
'And then we also saw those over at, ah, Akimel, uh, Gila River Governing Center, those who are known as, eh, Four Tribes, ah, ah, Cultural Resource Protection, ah, Committees.' (*#3*: 125-131)

This example is an exception to the predominant cross-linguistic pattern of adnominal demonstratives being employed for this function discussed in Diessel (1999). It appears in this instance that the speaker uses the pronominal form to introduce the referent and then clarifies who the known referent is with different types of identifying information.

Lastly, Table 39 below summarizes how O'odham compares to the cross-linguistic patterns regarding discourse functions, demonstrative type (proximal or distal) and syntactic formation or construction type (pronominal or adnominal) discussed in §6.1.3 (see Table 31). The list shows *tendencies*; there are exceptions.

Subtype	Cross-linguistic tendencies	O'odham tendencies
Discourse deictic	*Anaphoric:* proximal or distal (English)	Proximal or distal
	Cataphoric: Proximal (English)	Proximal or distal
	Pronominal	Pronominal
Tracking	*Either* proximal *or* distal	Distal
	Adnominal	Adnominal
Recognitional	Distal	Distal
	Adnominal	Adnominal

Table 39. Comparison between O'odham and cross-linguistic tendencies for demonstrative type (based on Diessel, 1999 / Himmelmann, 1996).

Overall, it appears that O'odham does not differ much from what previous findings show. The main departure are discourse deictic demonstratives with cataphoric reference; whereas English, at least, shows a strong preference for proximal demonstratives, O'odham may employ either the proximal or distal types.

6.5.2 More grammaticalized functions of demonstratives

The last group of examples demonstrate a number of features of more grammaticalized uses of demonstratives in O'odham based on the coding laid out in §6.4.3. First, I consider the categories which occur most frequently in the data. These are third-person pronouns, determinatives, connectives, and referring postpositions. Next, I discuss the less frequent yet clear cases of the remaining two categories: specific indefinite articles and placeholders.

Third-person pronouns

Pronominal tracking demonstratives can eventually extend to third-person pronouns (Croft, 2019; Diessel, 1999; Givón, 2001; Himmelmann, 1996; Lehmann, 2002). This occurs when pronominal demonstratives move up the Accessibility Scale to indicate active referents (see (11) and Figure 12, §6.1.4). Adopting pronominal demonstratives for the third-person pronoun function is a common cross-linguistic strategy (p. 74), and O'odham appears to do this, as well; however, sometimes the distinction between a tracking and third-person pronoun is ambiguous given that, in both cases, the demonstrative is co-referential with a previously introduced referent and can be translated with either tracking or third-person pronoun meanings. In the data, I classify tokens as third-person pronouns on the basis that they are pronominal, co-referential with a noun introduced at a prior point in the discourse, and indicate referents with active status (see identifying guidelines in (52), §6.4.3).

The first three examples utilize the distal demonstratives, which more often reflect third-person pronoun functions and meanings in the data. In the first clause shown in example (72), the narrator states that some people in the story want to tell something to a man named Shawañ Uwe. In the contiguous clause, the speaker then employs the demonstrative *heg,* which Juan Dolores glosses as 'he', to anaphorically refer to Shawañ Uwe. The referent is active given its continuity in the discourse segment.

(72)

…t	aṣaba	we:peg	o	'a:g	Shawañ Uwe
3.AUX.PFV	but	first	FUT	tell.IPFV	Smell.Like.Crow
but they must		first	will	tell	Smell Like Crow

t	**heg**	o	ha-'apecudac	'am	ha-wui
3.AUX.PFV	**he**	FUT	3PL.OBJ-prepare.for.PFV	LOC	3PL.OBJ-toward
	he	will	make good for them	there	toward

g ṣontal
ART soldiers
the soldiers …
'… but they must first tell Shawañ Uwe, and he will prepare it for the soldiers…
(*War*: 309)

In (73), a character called Coyote in the story grabs *'e-maccuḍ* 'his grindstone'. In the next reference to the grindstone, the narrator employs *heg* 'it', to anaphorically refer to the grindstone once again, which Coyote ties around his body so that he can dive deep into a pool of water.

(73) K 'am aha meḍk 'am aha 'ep bei g 'e-maccuḍ.
CONJ LOC then run-COMPL LOC then also take.PFV ART 3.POSS-grindstone

'Am be-k **heg** 'im aha woi 'e-'o: 'am k 'an
LOC take-COMPL **it** LOC then lie.down.PFV 3.REFL-back LOC CONJ LOC

si wu: 'e-'eḍa.
INTS tie.up.PFV 3.REFL-in.the.middle
'Then he ran and got his grindstone. Once he had it, he laid it on his back and tied it firmly around [lit. in the middle of] himself.' (*Coyote Turkey*: 38, 68, 76)

Finally, (74) includes a parallel pattern in which the speaker first employs a noun phrase, *g Chairman* 'the Chairman' which in the next clause is referred to again with *heg* 'him'. In all three cases, the pronominal forms clearly refer to active referents given the recency of their previous mentions.

(74) Tt hig geh 'am haha o ñei g Chairman,
1PL.AUX.PFV should HES LOC then FUT see.PFV ART Chairman,

k **heg** 'am haha 'ab o we:maj ñeo,
CONJ **him** LOC then LOC FUT together.with talk.PFV
'We should then go see the Chairman and speak with him.' (*#2:* 109-110)

In all three cases above, it is also worth noting that the distal demonstrative is given in its reduced form. This has significance in that it may suggest a developing split in morphological form between pronominal demonstratives which refer to semi- or inactive referents versus active referents. This pattern for third-person pronouns will either be confirmed or refuted in the quantitative portion of this study (§6.6.4).

The data also support the suggestion in previous studies (Mathiot, 1973; Saxton, 1983) that proximal demonstratives can also sometimes carry third-person pronoun meaning, as in the example below (see §6.2). The speaker in this example is quoting another person who said that they did not like non-O'odham persons to use "*'i:da*", which an O'odham speaker stated could be translated as either 'this' or 'it'. The proximal demonstrative in this case refers to the O'odham language.

(75) "'A:ñi ḍ ñe'ok 'i:da kuñ pi ṣa'i hoho'id
1SG COP language PROX CONN-1SG.AUX.IPFV NEG in.the.least like.IPFV

m-at g Miligan o hekaj **'i:da**"
SBDR-3.AUX.PFV DET Americans FUT by.means.of **it/this**.
'[Someone said,] "This is my language and I don't like any Americans to use it/this.' (*#3:*140)

The data suggests that pronominal, tracking demonstratives – more often of the distal type but the proximal type is also possible – have been extended to express third-person pronoun meaning. That is, they have extended to refer to entities with active status in the discourse.

Determinatives

Next, there is evidence that both distal and proximal demonstratives may serve as determinatives, although the former is much more common in the data, and either adnominal or pronominal constructions are possible, with the data showing a preference for the latter.

The demonstrative in an adnominal construction marks the head of a relative clause while a pronominal construction serves as the head of the clause, itself. I briefly introduced this pattern in §6.2 and describe it in more detail in this section. The hypothesis is that determinatives develop from recognitional demonstratives (Diessel, 1999). Instead of the demonstrative indicating a specific referent that the speaker presumes the hearer knows as well, the determinative indicates a *type* or *category* of referent that is characterized by the relative clause, as I argued in §6.1.4. Furthermore, the referent is typically a first mention as well as non-salient; that is, once the referent is introduced to the discourse, it does not continue as a central discourse participant. In what follows, I analyze several examples from the data, describing both their syntactic pattern as well as their functions in greater detail (see coding criteria laid out in (54), §6.4.3).

The first determinative example occurs in a video that was recorded for future placement on YouTube. The intended audience is the O'odham community. In the excerpt, the same plural, distal demonstrative, *hegam* is used twice, and in both cases it is followed by [*m*-AUX CLAUSE]$_{\textbf{REL}}$. While the speaker has been talking about the fact that a fence was erected, this is the first time that the speaker has directly referenced whoever did it; this is a first mention. Based on the text, nobody knew at the time who erected the unwanted fence, and no other indicators are present that suggest that the speaker is attempting to establish a known referent or who the specific person or persons were who constructed the fence. The speaker is simply stating that he and a colleague informed the government about the event – that somebody illegally put up a fence. The plural, pronominal demonstrative in this case stands in for an unspecified referent that the relative clause characterizes and in this manner plays the role of determinative. Once the referent is introduced, the speaker then goes on to

state that the report was recorded by an official who would later share the information with other officials. In other words, the referent does not continue as salient.

(76) … 'am ha-gagida **hegam** m-at-p heḍai ḍ 'i
LOC 3PL.OBJ-accuse **DIST.PL** SBDR-3.PFV-ASSUMP who COP PUNC

hegam m-at-p 'am wawañ hegai ko:li.
DIST.PL SBDR-3.AUX.PFV-ASSUMP LOC stretch.IPFV DIST fence
'…we reported those who were, those who put up that fence.' (*#2*: 138)

The example above is contrasted with the recognitional demonstrative example in (77) below. This comparison highlights the differences between the two functions given that they are closely related (see Table 32 in §6.1.4) with determinatives having hypothetically developed from recognitional demonstratives. In (77) below, the plural form of the distal, adnominal demonstrative refers to 'those O'odham' who specifically reside on the Tohono O'odham Nation. This is the first time that the speaker has mentioned this particular sub-group of the O'odham community; however, based on context as well as linguistic cues such as the many hesitations, it appears that the speaker is attempting to clearly establish this specific group of people whom his audience will undoubtedly know. The speaker is not referring to the general O'odham public but rather those O'odham who reside on the Tohono O'odham Nation or reservation, not on the other reservations. Note also that the speaker begins by simply referring to those O'odham who live on the 'Nation' and then further clarifies the known referent by following up with 'Tohono'. This is in contrast to the previous example in which the speaker does not appear to have a particular group of people in mind who erected the fence (see §6.5.1 for additional examples of recognitional demonstratives).

(77) (Recognitional demonstrative)

Uh, eh, **hegam O'odham** m-o gmhu ki: ah, eh, eh,
HES, HES, **DIST.PL O'odham** SBDR-3.AUX.IPFV over.there live.IPFV HES, HES, HES,

Nation, ah, Tohono t'am k 'ia 'am ha'ic m-o 'ia
Nation, HES, Tohono ABS-at CONJ here LOC some SBDR-3.AUX.IPFV here

'ep 'e-maṣcam 'e-- 'ep 'ia Cukṣon.
Also 3.REFL-teach.IPFV 'e-- also here Tucson

Hegam 'at-p hab 'ap[39] 'am a cem s-ma:c 'am
DIST.PL 3.AUX.PFV-ASSUMP thus also LOC indeed try STAT-know LOC

cem s-'amicud, has masma, 'am 'am cem s-juñ-im-k
PART STAT-understand.IPFV somehow LOC LOC try STAT-do-DESID-STAT

g 'e-cipkan 'am ha-we:hejeḍ g O'odham.
ART 2.REFL-work LOC 3.PL.OBJ-for ART O'odham
'Uh, eh, those O'odham who live over at the, ah, ah, Nation, ah, Tohono [O'odham Nation], and also some who are here studying also he—here in Tucson. They too must know, must understand, in what manner, you are trying to do your work for the O'odham people.' (*#3*: 333-339)

Another example which demonstrates the determinative rather than the recognitional use of demonstratives is shown in (78) below. The excerpt is taken from a historical text in which a mother is speaking to her son whose father is no longer alive. The mother is aware that a group of men will go hunting, and she wants her son to go with them and observe how they hunt. She then explains the reason why in the segment shown below.

(78) "…heghekaj m-añ 'a:ñ ḍ 'uwi, kc pi ma:c
DIST.because.of SBDR-1SG.AUX.IPFV 1SG COP woman CONJ NEG know.IPFV

m-añ-s has ju-k o m-maṣca
SBDR-1SG.AUX.IPFV-DUB do.something-COMPL FUT 2SG.OBJ-teach.PFV

[39] Variant of *'ep* 'also'.

hegai m-o ḍ ceoj himdag."
DIST SBDR-3.AUX.IPFV COP man way.of.doing
'"…because I am a woman and do not know how I will teach you what men's responsibilities are [lit. that which are man's ways of doing]."' (*Boy*: 27)

The speaker in the above example employs the singular, pronominal demonstrative, which is then immediately followed by a relative clause which defines what the referent is – in this case, the general responsibilities that were typically carried out by men at that point in time. The pronominal demonstrative is used as a first mention – it cannot be said to directly point back to a previously mentioned referent – and it also does not evoke a specific referent that is expected to be known by the hearer. Once the referent is defined, the narrator then goes on to tell the audience in the subsequent lines (not shown here) that the boy leaves with the group of men.

One other detail to note is that the singular demonstrative is in its phonologically full rather than reduced form (plural demonstratives are always full). This is the case throughout the data whether the demonstrative is distal, as in the example above, or proximal (see the singular, proximal demonstrative example in (80) below). This suggests that at least at this stage, the form does not show any phonological reduction associated with this grammatical function.

It is also possible for proximal demonstratives to function as determinatives, as in examples (79) and (80).

(79) Uh, **'i:dam** m-o 'am daḍha, m-o 'an
Uh, **PROX.PL** SBDR-3.AUX.IPFV LOC sit.IPFV SBDR-3.AUX.IPFV LOC

ha-maliom 'amai…
3.PL.POSS-supervisor right.there…
' Uh, these who are sitting in office, who are their supervisors there…' (*#1:* 4).

(80) …'ep **'i:da** m-o hab 'a'aga Mi:lgan the economy,
also **PROX** SBDR-3.AUX.IPFV thus call.IPFV Americans the economy

m-o cem pi ha'icu g lial 'o hemu s-kawk
SBDR-3.AUX.IPFV little NEG thing ART money or now STAT-difficult

g lial, himdag.
ART money way.of.doing
'…also this that the Americans call the economy, that there isn't much money or now the money is tight…' (*#4*: 127-129)

In (79), *'idam* 'these' precedes two modifying relative clauses in this case, the first of which explains that *'idam* refers to the people who hold positions in political office and the second which tells us they are also supervisors of some type. The proximal demonstrative does not indicate a previous referent, and the example also does not refer to a specific, known referent. Similarly, in (80), the singular proximal demonstrative, *'i:da* 'this' serves as a determinative of the relative clause, being neither co-referential with a prior referent not indicating a specific known referent. The relative clause is used to define what precisely *'i:da* represents.

Connectives and referring postpositions

In this portion of the descriptive analysis I describe and discuss both connectives and referring postpositions in O'odham. I combine my discussion of these two demonstrative uses due to the fact that, in some cases, the constructions clearly share certain similarities and, in other cases, the specific ways in which they contrast adds clarity to how each category should be defined. (See specific coding procedures in (56) and (57), §6.4.1).

While I briefly discussed the syntactic make up of these constructions earlier in §6.4.1, more in-depth description is necessary in order to draw the readers' attention to certain details. I will address function more directly later. I begin first with a set of examples which illustrate the frequent co-occurrence of demonstratives and postpositions. There are

two possible patterns: [DEM POSTP] or [DEM POSTP N]. The first pair of examples below shows pronominal demonstratives with the [DEM POSTP] pattern.

(81) Ku-mt 'aṣ 'ab a 'a:ñ 'am 'aṣ o ñ-dagito **heg'eḍa…**
CONN-2PL.AUX.PFV just LOC surely 1SG LOC just FUT 1SG.OBJ-leave **PROX.in**
"'And so you just leave me there [lit. in that]…"' (*Boy*: 95)

(82) Hemu hab a 'i masma […] m-ant 'am
now thus PART PUNC in.this.way […] SBDR-1SG.AUX.PFV

o ha-hemapa **'id'eḍa**.
FUT 3PL.OBJ-gather.PFV **PROX.inside**.
'[Coyote said to his wife,] "Now the best thing for us to do […] is to gather the animals in it."' (*Coyote Skunk*: 30, 62, 79).

In (81) above, someone is giving orders to another group of people, all of whom are in the midst of fighting a battle. The person tells everyone else to leave him *heg'eḍa* 'in that', meaning to leave him behind in their current location. In (82), Coyote and his wife have just finished building their home. Coyote turns to his wife and tells her that they should trap all of the other animals *'id'eḍa* 'inside this' or 'it'. Note that in both cases, the demonstratives are in their reduced forms, *heg* and *'id*. Additionally, both demonstrative tokens are written together with the postpositions in each example as single, unified words. In the case of (81), an example from the historical data, this is following Juan Dolores's convention for this specific combination of lexical items (i.e. *heg* and *'eḍa*), which are otherwise considered independent lexical items. Example (82) reflects Mathiot's practice. In fact, in a majority of demonstrative-postposition combinations, Mathiot records them as a unit. Both of these details may suggest that the demonstrative and postpositions have fused to some extent, an observation which Diessel (1999) argues is one indicator of grammaticalization (see outline included in (10), §6.1.4).

Examples (83) and (84) below illustrate cases in which a demonstrative and postposition co-occur with the referent of the demonstrative. In such cases, a specialized construction is used, [DEM POSTP N], rather than the basic adnominal construction that would normally be employed.

(83) "… k 'a:pi, wo mua m-at 'ia wo meḍ-ad
CONJ 2SG, FUT kill.PFV SBDR-3.AUX.PFV here FUT run.SG-FUT.IPFV
and you will kill it that here will be running

heg wui ṣa'i 'am t-weco."
DIST toward brush LOC 1PL.OBJ-under."
that toward brush there under us
'"… and [you] shoot him when he runs toward the [lit. that] brush below us!" (*JD*: 254, 306)

(84) Ku-nt hemu o 'ab 'abṣ 'i hi: **'idwui**
CONN-1SG.AUX.PFV now FUT LOC just here come.PFV **PROX.to**

t-nawoj mi:tol.
1PL.POSS-friend cat
'[And] now I just came here to [this] our friend the cat.' (*Coyote Cat*: 27, 49, 58)

In (83) above the distal demonstrative precedes the postposition *wui* 'toward,' which is followed by its nominal referent *ṣa'i* 'brush'. The same syntactic pattern, [DEM POSTP N], is true of the proximal example in (84) which again Mathiot also records as a unified word, *'idwui*. What this and some of the other examples may suggest is the development of referring postpositions in which the singular demonstratives have become an obligatory part (see further discussion in §6.5.2).

The next two examples are from Saxton (1982) and Langacker (1977), which further illustrate the interpretation that the entities represented by the nouns are, in fact, the referents of the demonstratives in DEM POSTP N constructions.

(85) **heg we:m** ce'oj
that with man
'with that man' (Saxton, 1982, p. 189)

(86) **heg-da:m** do'ag
that-on mountain
'on that mountain' (Langacker, 1977, p. 103; modified)

In (85) and (86), the demonstratives are followed by postpositions *we:m* 'with' and *'da:m* 'on', respectively, which are in turn followed by nouns. In both cases, the nouns *ce'oj* 'man' and *do'ag* 'mountain' are the nouns to which the demonstratives refer. In cases where the referent is plural, the same syntactic pattern is used.

Examples (83)-(86) above contrast with circumstances in which a demonstrative is not present. In these instances, the word order is typically [N POSTP] (see §3.7.2 for full discussion on adpositions), as shown in (87) below.

(87) m-o gahu **San Juan** **'amjeḍ** hihim g ṣontal…
SBDR-3.AUX.IPFV over.there **San Juan** **from** come.PL.IPFV ART soliders
that over there San Juan there from are coming the soldiers
'[Someone told them] that the soldiers are coming from San Juan…' (*War*: 61)

In the example above, *'amjeḍ* 'from' is in postposition in relation to San Juan, the name of a locale. The difference between (87) and the previous set examples also suggests a specialized construction has developed in the presence of a demonstrative.

Moving on to function, I begin with connectives. Connectives mark a "semantic relationship between the conjoined propositions" (Diessel, 1999, p. 125). One hypothesized cross-linguistic source for connectives are pronominal, discourse deictic demonstratives that co-occur with either an adposition or adverb. Connectives and discourse deixis are similar in that they both play the role of connecting two chunks of discourse. Over time, the pronominal demonstrative and adposition or adverb that are employed for the connective

function may show signs of fusion, one of the phonological characteristics of grammaticalization (Diessel, 1999, pp. 118-119).

While demonstratives can hypothetically co-occur with any of the postpositions in the language, the connective meaning appears to be limited to contexts in which the demonstrative occurs with two postpositions in particular; *hekaj* 'with, by means of, because of' and *'amjeḍ* 'from'. This becomes evident after identifying the tokens that have the features listed above in (55), §6.4.3. Possible connective meanings that can be attributed to *hekaj* in particular are included in the following dictionary entry from Mathiot (n.d.a), which the author has labeled as a conjunction (*Conj*) (p. 229).

hekaj, heg-hekaj, g-hekaj Conj: therefore, consequently, because, in order to

Some examples in the data further illustrate some different connective meanings. The first example in (88) was collected by Juan Dolores. In it, the distal, pronominal demonstrative occurs adjacent to the postposition *hekaj* (recorded here as *ekaj*). The example reflects the strong tendency for the construction to be written as a single word and for the demonstrative to occur in reduced form (i.e. *heg* instead of *hegai*). As pointed out in §6.1.4, fusion of component parts is one possible indicator that a construction is more advanced on a grammaticalization continuum. In line with this view, Mr. Cruz, the primary language consultant for this study, states that the demonstratives cannot be omitted in examples like those shown below (personal communication), suggesting that it is obligatory. Dolores provides the gloss 'for that reason' for the construction as a whole, and this meaning is also reflected in the free translation, "As a consequence."

(88) "Ñ-we:mkam 'o s-paḍma waṣ kokṣa cuckagad
1SG.OBJ-companion 3.AUX.IPFV STAT-lazy just sleep.REP nightly
my partner is lazy just sleeps in the nights

m-ac wapga **hegekaj** 'e-papḍacud g
SBDR.1SG.AUX.IPFV irrigate.DISTR **DIST.because.of** 3.REFL-spoil.DISTR ART
that we irrigate **for that reason** spoils the

wapaikka c heg 'amjeḍ mu'i ñ-cikpan-a-cud
ditch.PL CONJ DIST from much 1SG.OBJ-work-EP-CAUS.IPFV
ditches and that from much works me
'"My partner is lazy. He just sleeps the nights that we irrigate. **As a consequence** the ditches get damaged and I have a lot of work.'" (*JD*: 259, 307).

The construction in this case conjoins the first proposition – that the speaker's working partner is lazy and sleeps when they need to be irrigating the land – with the proposition that follows; namely, as a result of his laziness, the ditches are ruined which leads to more work for the speaker. An anaphoric, discourse deictic sense can also be detected; *hegekaj* refers to the proposition preceding it.

A cataphoric sense is also present in some cases with the identical construction, *hegekaj*. In the lines leading up to example (89) the narrator is talking about the type of work his father had him regularly do when he was younger. One of his duties was to lead a horse that had a bucket attached to it which the horse would then pull along behind it. The narrator then states, as shown in (89), that the work was not interesting or desirable *hegekaj* 'for the reason' or 'because' you don't actually go anywhere with the horse. The construction as a whole clearly links one proposition with another. Furthermore, *hegekaj* points forward in the narrative in this case. *Why is the work uninteresting? I am going to tell you why: because it leads you anywhere*.

(89) Pi tahoho'idam 'i:da cipkan **hegekaj** pi hebai 'e-hihim,
NEG interesting PROX work **DIST.because.of** nowhere 2.REFL-go.REP.IPFV
not interesting this work **for the reason** nowhere (you) go
'This work is undesirable **because** it leads nowhere.' (*War*: 260-261)

Example (90) below similarly illustrates a connective meaning that is cataphoric in nature. In the text immediately preceding what is shown below, the speaker discusses the need for community approval before specific government officials can receive a wage increase. Then, in (90), the speaker essentially states that only at that point – if they receive community approval – will they raise wages, and the reason for this argumentation is stated following the sentence connective *hegekaj* – because the officials work for the people.

(90) …ah m-att 'o'i ceṣaj o **hegekaj**
HES SBDR-1PL.AUX.PFV pursue.PFV raise.PFV FUT **DIST.because.of**

m-o g hemajkam ha-wui cipkan.
SBDR-3.AUX.IPFV ART people 3PL.OBJ-toward work.IPFV
'…ah, we will raise it **because** it is they who are working for the people.' *(#4:* 57)

In other cases, the distal demonstrative component is further reduced to *g* (i.e. *hegai* → *heg* → ***g***), as shown in example (91) below, although I have observed this only in the midcentury texts. This form may be associated specifically with the Totoguañ sub-dialect of Tohono O'odham (see §4.2.2). In the example below, as with the previous examples, the construction links one proposition with another, in this instance anaphorically. A story character is invited to an event because of what is stated in the earlier proposition; because the king's son is getting married.

(91) a m-at-ṣ o ha-ho:nt g lai 'alidag
oh! SBDR-3.AUX.PFV-QUOT FUT 3PL.OBJ-wife.make ART king man's.child

ghekij 'am m-waid ...
DIST.because.of LOC 2SG.OBJ-invite.IPFV …
'"The son of the king is getting married. **That's why** you are invited…"' (*Coyote Cat*: 2, 37, 54).

On occasion, the proximal demonstrative, *'i:da* 'this', also co-occurs with the postposition *hekaj*, usually in the reduced from *'id*, and a similar connective meaning is expressed by the construction. In example (92), the narrator, Antonio Lopez, is describing advice that he was given as a child. He is told to show generosity to others, to exercise regularly, to practice perseverance, among other things. He is told that *'idhekaj* 'because of this', meaning if he follows these words of advice which are referred to anaphorically, he will have a good life.

(92) …m-at o s-ñennaṣañ-k, wo s-hohot-k,
SBDR-3.AUX.PFV FUT STAT-be.awake.PLS-COMPL FUT STAT-be.fast.PL-STAT
that they will be wide awake will be fast in action

wo s-cu-nankog-ad, wo s-cecojima-k,
FUT STAT-STAT-endure.REP.IPFV-FUT.IPFV FUT STAT-generous-STAT
would have great endurance will be generous

hab cem hekid ḍ o tacui-k 'am 'e-hajuñ ha-ki:k 'am,
thus always COP FUT beloved-STAT LOC 2REFL-relative 3PL.OBJ-house.PL LOC
thus always will be desired there your relations houses at

'idhekaj hab hebai s-'ap o ki:ḍag.
PROX.because.of thus somewhere STAT-good FUT home-STAT
for these reasons somewhere good will be your living
'[Perhaps all children are raised in the same way:] to be early risers, to be quick, to persevere, to be generous, so that there will always be love at your relatives homes, and **because of this**, home life will be good wherever you are.' (*Antonio*: 101-105)

Based on the examples occurring in the data for this study, the other context in which demonstratives appear to be employed for a connective function is when the reduced, distal

demonstrative *heg* co-occurs with the postposition *'amjeḍ* 'from'. This is illustrated in (93) below. The meaning of *heg 'amjeḍ* in the example is along the lines of 'as a result of that' – in this case, as a result of becoming a runner, the person being spoken to will receive many good benefits. In the same vein, Saxton et al. (1983) include an entry in their dictionary which lists *'amjeḍ* as an adverb with the meanings 'then' or 'subsequently' (p. 3).

(93)

Hab	o	melid	t	**heg**	**'amjeḍ**
thus	FUT	become.able.to.run	3.AUX.IPFV	**DET**	**from**
thus	(you) will	become a good runner		**that**	**there from**

mu'i	ha'icu	wo	s-'e-m-'ap-k.
many	thing	FUT	stat-3.REFL-2SG.OBJ-good-STAT
many	things	will	be good for you

'[He was told,] "You will become a runner and **as a result of that** [lit. from that] you will benefit in many ways."' (*Antonio*: 13)

Example (88) from above repeated below as (94) also includes *heg 'amjeḍ.* The person being quoted explains that he has a lot of extra work 'as a result of' the ditches being ruined. That is, the demonstrative and postposition connect one idea to the other in a discourse deictic manner that is, in this instance, anaphoric.

(94)

"Ñ-we:mkam	'o	s-paḍma	waṣ	kokṣa	cuckagad
1SG.OBJ-companion	3.AUX.IPFV	STAT-lazy	just	sleep.REP	nightly
my partner	is	lazy	just	sleeps	in the nights

m-ac	wapga	hegekaj	'e-papḍacud	g
SBDR.1SG.AUX.IPFV	irrigate.DISTR	for.that.reason	3.REFL-spoil.DISTR	ART
that we	irrigate	for that reason	spoils	the

wapaikka	c	**heg**	**'amjeḍ**	mu'i	ñ-cikpan-a-cud
ditch.PL	CONJ	**DIST**	**from**	much	1SG.OBJ-work-EP-CAUS.IPFV
ditches	and	**that**	**from**	much	works me

'"My partner is lazy. He just sleeps the nights that we irrigate. **As a consequence** the ditches get damaged and [from that] I have a lot of work."' (*JD,* D&M, pp. 259, 307).

Earlier in this section, I discussed several examples (see (83) and (84) above) that are similar to the one shown in (95) below where the noun referent of the demonstrative is present in the sentence. In these cases, the postposition follows the demonstrative rather than the nominal referent to which the demonstrative refers. That is, the word order in such cases is [DEM POSTP N], not *[DEM N POSTP].

(95) Eh, 'aṣ 'i gahi 'i wo'iwa **heg wei** Ju:kam ñe'ok.
HES, just PUNC sideways PUNC lie.down **DIST towards** Mexican language
Eh, it just shifted towards [lit. lie down towards] the Spanish [lit. Mexican] language.' (*#3*: 48)

The function of the construction in this case is not connective but rather to verbally point to the nominal referent, which in this case is *Ju:kam ñe'ok,* 'lit. Mexican language'. There are no instances in the data where the pattern [DEM POSTP N] appears with the connective function. Given the clues that suggest connectives are closely related to and develop out of the discourse deictic function where the demonstrative refers to entire propositions, not nominal referents, it follows then that connectives would not appear with the [DEM POSTP N] pattern. Instead, this example illustrates a referring postposition.

The following supplemental example is borrowed from Mathiot's (n.d.a) dictionary to further illustrate that connectives and referring postpositions are distinct from one another. The example is made up of the same components (i.e. *heg/g* + *hekaj/ekaj*) as several of the earlier sentence connective examples, specifically (88)-(91), but it clearly differs in crucial ways from these cases. First, the proximal demonstrative refers to the noun, *'u:s* 'stick', and the example reflects the word order ([DEM POSTP N]) that is frequently employed in cases where a postposition is also present. In the earlier examples in (88)-(91), the function of the construction is to connect one proposition to another, either anaphorically or cataphorically.

The other major difference is that *hekaj* in this instance clearly expresses the postpositional meaning, 'with' or 'by means of' rather than combining with the demonstrative to link one sentence or proposition to another. Constructions like the one in (96) instead indicate specialized, referring postpositional meanings.

(96) Ge:gewa-ñ **heghekaj** 'u:s!
hit-IMP **DIST.with** stick
Beat him with that stick! (Mathiot, n.d.a, p. 229)

Referring postpositions can also occur with the [DEM POSTP] construction in the data. In (97), the proximal demonstrative component *'id* in the construction, *'id'eḍa*, indicates the previously introduced referent *t-ki:ta* 'our house construction', and in (98) *heg* refers to the location mentioned earlier, *Taḍ Memelkuḍ.* In these cases, the constructions function in much the same way that tracking demonstratives do.

(97) Hemu hab a 'i masma (m-att 'idañ 'am a
now thus PART PUNC in.a.certain.way SBDR-1PL.AUX.PFV now LOC indeed

na:to g t-ki:ta k hab ma:s)
complete ART 1PL.POSS-house.construction CONJ kind-of

m-ant 'am o ha-hemapa **'id'eḍa**.
SBDR-1SG.AUX.PFV FUT 3PL.OBJ-gather.PFV **PROX.inside**
'[Coyote said to his wife:] "Now, the best thing for us to do – since it looks like we are through with our house building – is to gather the animals in it [lit. this]." (*Coyote Skunk*: 30, 62, 79)

(98) 'Am hab ce:gig hemuc 'amai, m-a-ṣ 'am hab ḍ
LOC thus name now over.there, SBDR-3.AUX.IPFV-QUOT LOC thus COP

Taḍ Memelkuḍ 'am 'aṣ. 'Ab 'aṣ **heg 'oidk** me: hegai 'o:b.
Foot Running LOC just LOC just **DIST along** run.SG.PFV DIST enemy
'The place is now named, what is known as Taḍ Memelkuḍ [Foot Running], it is there. It is along there where that enemy ran.' (*Boy*: 76-78)

It may be the case that all demonstrative postpositional constructions of the types discussed above – whether they are connectives or referring postpositions – are grammaticalized to some extent. Signs of fusion such as phonological reduction as well as the fact that the demonstratives appear to be obligatory suggest that examples of both types should be viewed as fixed constructions in O'odham. As Diessel (1999) points out, signs of this nature suggest further advancement along a grammaticalization continuum.

The differences between the earlier examples in (92)-(94) and those in (95)-(98) above also suggest different paths of development, with the former cases having developed from discourse deictic demonstratives into a connective function, while the latter show signs of evolving from tracking demonstratives into postpositions with specialized meanings in O'odham. We would expect that both of these construction types would become more frequent over time as they become more strongly associated with a connective meaning in some cases and to indicate specific referents in others (see §6.6.4).

Specific indefinite articles

Evidence in the data suggests that O'odham always employs proximal demonstratives to mark specific yet indefinite referents which are unknown to the hearer. Once the specific but unknown referent is introduced to the discourse, it commonly remains "thematically important" (Wright & Givón, 1987, p. 16). I identified three cases of this use in the data (see (53), §6.4.3 for coding). This particular function of proximal demonstratives in O'odham, I hypothesize, developed due to external factors (i.e. influence from English). Several details point in this direction. First, I identify examples of this type only in the modern data, suggesting that this development occurred after prolonged contact with English. Secondly, as pointed out in §6.1.4, cross-linguistically it is more common for languages to develop a

separate specific indefinite article from the numeral one (e.g. Diessel, 1999; Wright & Givón, 1987). In contrast, English employs proximal *demonstratives* for this use (Diessel, 1999; Himmelmann, 1996; Wald, 1983; Wright & Givón, 1987). That is, proximal demonstratives have grammaticalized at least in colloquial English to serve this purpose in certain discourse settings. O'odham also uses its proximal demonstrative, *'i:da*, for a similar function. Finally, the fact that O'odham employs its proximal demonstrative – as in colloquial English – rather than its distal demonstrative also suggests external influence from English. I discuss all three examples in the descriptive analysis below.

The first example in (99) is from a YouTube video. In it, the speaker introduces the referent, *'i:da doctoral degree* 'this doctoral degree' for the first time in the recording.

(99) … Ṣhaba hab 'i ñ-'el ñt hig o 'am
however thus PUNC 1SG.REFL-think.PFV 1SG.AUX.PFV should FUT LOC

ṣa'i ñei 'ep m-a-s hascu ḍ 'i:da, ah,
in.the.least see.PFV also SBDR-3.AUX.IPFV-DUB what COP PROX, HES,

m-añ hemu 'am 'oid **'i:da doctoral degree** …
SBDR-1SG.AUX.IPFV now LOC pursue.IPFV **PROX doctoral degree**
'…so, I thought, I should check out what this is, ah, what I am currently pursuing, this doctoral degree…' (*#3*: 20-22)

In the example above, the demonstrative cannot be said to play a tracking function. In fact, the example occurs within the speaker's first few lines of the YouTube video recording in which the speaker is sharing some background information about himself for his future viewers. The demonstrative also does not refer anaphorically or cataphorically to a specific chunk of discourse. Furthermore, given that the speaker is being recorded for a general O'odham audience, many of whom do not know the speaker, it is also unlikely that he would employ the demonstrative to signal a known referent in relation to his own life. Additional

cues also suggest that the speaker does not expect the audience to have knowledge about the specific referent. The speaker first briefly prefaces the upcoming referent by telling his audience that it will relate to something that the speaker is currently pursuing; the speaker is pointing out why the upcoming referent will be relevant to the discussion. The speaker then states what the specific referent is, *'i:da* or 'this' *doctoral degree*, using the proximal demonstrative in order to momentarily highlight the referent. Once the speaker introduces the referent in the discourse, the speaker goes on afterwards to explain what he has learned during his time pursuing the degree. The referent is thematically important in that it is the primary scenario in which the speaker gains specific types of knowledge that leads him to get involved in an entire other set of activities. It seems unlikely that the speaker would use the demonstrative if he did not want to draw attention to the referent for some discourse purpose, especially given that the referent is discourse-new. The speaker employs the proximal demonstrative for a reason, and the details suggest that the example falls in line with a specific indefinite use.

The two remaining tokens, in (100) and (101), which are also from YouTube, illustrate some of the same patterns. In (100), the speaker is talking about the rights of Indigenous people, and in (101), the speaker is answering a question about the general O'odham public versus Tribal Council approval for a language revitalization project.

(100) hemu 'id 'an kaidag uh, Human Rights k,
now PROX LOC make.a.sound.IPFV uh Human Rights CONJ

'i:da United Nations Declaration on the Rights of Indigenous Peoples,
PROX United Nations Declaration on the Rights of Indigenous Peoples,

Article 13 and 14 and 31, m-o specific heg 'am hab si
Article 13 and 14 and 31 SBDR-3.AUX.IPFV specific DIST LOC thus INTS

cu'ig heg 'eḍa, …
be.a.certain.way DIST in
'Now, we hear about this, uh, Human Rights and this United Nations Declaration on the Rights of Indigenous Peoples – Article 13 and 14 and 31 – where it is specifically [stated] in that Article…' (*#3*: 167-171)

(101) …ha'i **'idam** **communities**, 'amai, uh,
some **PROX.PL** **communities** over.there uh,

Wa'a Ki: hegam 'o ha'ic hab kaij s-hohoid
Casablanca DIST.PL 3.AUX.IPFV some thus say.IPFV STAT-like

'ac taccu c…
1PL.AUX.IPFV want 1PL.AUX.IPFV
'…some of these communities at uh, Wa'a Ki:, some of them did say we like it and we want it.' (*#3:* 203-205)

In (100) the speaker brings up the United Nations Declaration of Human Rights for the first time in the discourse. Again, given that the speaker is being recorded for a general O'odham audience, there is no reason to suspect that the speaker would assume his future viewership would be familiar with this referent, which is based on a very specific type of knowledge. After introducing the referent, the speaker goes on to specify certain articles in the Declaration which states the rights that Indigenous people have to revitalize and protect their respective languages and cultures, how the people have a right to protect their childrens' human rights, etc. Details associated with the Declaration continue to be highly salient to this segment of discourse and the speaker uses the proximal demonstrative as a specific indefinite article to highlight this fact.

Similarly, in (101) above, the speaker begins by stating *'idam* 'these' *communities*, which refers to specific communities that the speaker has not yet introduced into the discourse. The speaker then talks about *Wa'a Ki:* (referred to in English as Casablanca), and how the people there approve of the language revitalization project. The speaker then continues to talk about the youth community, the elder community, etc. and their feelings

about the project. Again, the speaker utilizes the proximal demonstrative with the functions associated with the specific indefinite use in order to signal how central the various communities will be to his discussion.

In these examples, it is the proximal demonstrative that is preferred over the distal one. The proximal demonstrative functions as a specific indefinite article, which marks new and unknown referents that are at the same time specific and central to the discourse segment much as proximal demonstratives have been shown to do in English, as well (Diessel, 1999; Himmelmann, 1996; Wald, 1983; Wright & Givón, 1987).

Placeholders

Finally, cross-linguistically, demonstratives often evolve into placeholders (see §6.1.4). Speakers employ the demonstratives at a point in the discourse when they have trouble recalling a lexical item, often a noun (e.g. Hayashi & Yoon, 2010). The speaker uses the demonstrative as a stand-in for the intended lexical item. Additionally, the demonstrative usually mirrors the morphosyntactic features such as number distinctions of the anticipated word (Podlesskaya, 2010). This use is of particular interest in the context of language shift. Among speakers who have relatively few opportunities to use O'odham, placeholders may be an indication of uncertainty in the retrieval of lexical items.

About half a dozen clear instances of the placeholder function appear specifically in the modern data in a single speaker's discourse. (I discuss coding criteria in (57), §6.4.3.) I provide three of them below. In all cases, the proximal demonstrative plays the placeholder function, as reflected in examples (102)-(104) below.

(102) … Ṣhaba hab 'i ñ-'el ñt hig o 'am
however thus PUNC 1SG.REFL-think.PFV 1SG.AUX.PFV should FUT LOC

ṣa'i ñei 'ep m-a-s hascu ḍ **'i:da**, ah,
in.the.least see.PFV also SBDR-3.AUX.IPFV-DUB what COP **PROX**, HES,

m-añ hemu 'am 'oid 'i:da doctoral degree …
SBDR-1SG.AUX.IPFV now LOC pursue.IPFV PROX doctoral degree
'…so, I thought, I should check out what this, ah, what I am currently pursuing, this doctoral degree…' (*#3*: 20-21)

(103) kc hema 'ep 1955 t-'amjeḍ hemu m-añ 'ia cekc g
CONJ now also 1955 ABS-from now SBDR-1SG.AUX.IPFV here have ART

eh, eh, **'i:da** (…) thumb drive uh, flash drive ah, ah,
HES, HES, **PROX**(…) thumb drive HES, flash drive HES, HES,

m-o 'an 'o'ohaḍag g Madeleine Mathiot, ah, Dictionary
SBDR-3.AUX.IPFV LOC writing ART Madeleine Mathiot, HES, Dictionary
'…and also another from 1955, which I now have here on the, eh, eh, this (2.0) thumb drive, uh, flash drive, ah, ah, where the writing, the eh, Madeleine Mathiot, ah, Dictionary.' (*#3:* 421-428)

(104) … k eh, heg hekaj 'at-p hab 'e-ju:
CONJ HES DIST because.of 3.AUX.PFV-ASSUMP thus 3.REFL.happen.IPFV

m-o, m-o hab 'a'aga **'idam**, ah, linguistics, ah,
SBDR-3.AUX.IPFV SBDR-3.AUX.IPFV thus call.IPFV **PROX.PL** HES linguistics HES

ah, studies or, or, ah, academics, ah, language shift…
HES studies or or HES academics HES language shift
'… and, eh, the reason why it probably happened – what is, what is referred to in these, ah, linguistics, ah, ah, studies, or, or, ah, academics, ah, "language shift"…' (*#3:* 37-39)

In all three examples above, the proximal demonstrative is employed as a substitute for specific, anticipated lexical items. In (102), the referent is *doctoral degree*, in (103), *thumb drive/flash drive*, and in (104), *linguistics studies/academics*. Notably, in all three cases, the referents are in English which is the case with all of the examples displaying this function. It may be that the speaker is searching for a way to state these items in O'odham – either

because the speaker knows O'odham terms exist for these concepts but cannot recall them or is trying to create a way to state the terms in O'odham on the spot but is unable to do so. Whatever the reason, in the end, the speaker opts for the English terms. Note also that, in all three examples, the proximal demonstratives are accompanied by a series of hesitation markers and/or pauses. Both are additional indicators that the speaker is in the midst of searching for the appropriate lexical items with some difficulty. It could be that this is the result of the limited opportunities to use the O'odham language, resulting in fewer chances to learn the appropriate O'odham equivalents for the terms in the examples above. It may also be the case that the O'odham speaking community as a whole has few opportunities to create new O'odham terms for the specialized items in the excerpts.

Finally, also note that the demonstratives match the number of their referents, a hallmark feature of demonstrative placeholders. In (102) and (103), the singular forms of the proximal demonstrative, *'i:da* 'this', matching their singular nominal referents and, in (104) the speaker uses the plural form *'idam* 'these' as a stand-in for *linguistics studies* and/or *academics*, which are also plural.

6.6 Results: Quantitative analysis

In §6.5, I presented the results of my descriptive analysis regarding the wide range of functions that demonstratives play in O'odham. In the final portion of this study, I carry out a comparative, quantitative analysis of the data based on the distribution of token frequencies regarding different variables, such as demonstrative type, construction type, phonological form, and function. These are "the features that characterize the grammaticalization of demonstratives" (Diessel, 1999, p. 118). For example, distal demonstratives tend to grammaticalize more often than proximal demonstratives (p. 118). If this is the case, then

demonstratives will probably also gradually appear more frequently in the data over time. In this manner, assessing frequencies allows me to identify possible grammaticalization trends, if they exist in the data, based on previously observed crosslinguistic tendencies (see §§6.1.3 and 6.1.4). This assessment also allows me to determine the impact decreased domains of language use and/or contact has had on the development of demonstratives, if any, over approximately the last century. I present a few possible predictions of what the outcomes could be below.

6.6.1 Possible outcomes

First, if a particular demonstrative pattern does occur more frequently over time, the increase in frequency may suggest that, even though the settings in which O'odham is regularly used is diminishing, decreased domains of use has had little impact on the evolution of at least the particular area of grammar under examination. As an example, we would expect pronominal demonstratives to correlate with the third-person pronoun function (i.e. to indicate active referents in the discourse) more often as time passes, signaling that pronominal demonstratives continue to expand towards this more grammaticalized function. If this is the case in the data, this would suggest that pronominal demonstratives have so far continued to evolve despite language shift.

In some cases, the possibility also exists that contact with English (i.e. externally-motived change) has promoted certain developments, particularly during the second half of the twentieth century during which the O'odham community speaks more English with each passing generation. Any clues that point in this direction require cautious consideration.

If distribution frequencies are similar across time periods, this probably indicates that the demonstrative function has remained stable, having changed very little over the last

century or at least that the changes are occurring at a rate that is not detectable within a single century.

Finally, if specific types of patterns significantly decrease over time, this suggests that the evolutionary process may have been disrupted at some point, preventing the construction from evolving as it otherwise may have. To illustrate, I return to the earlier example in which pronominal demonstratives are expected to increasingly correlate with a third-person pronoun meaning over time. If this relationship instead decreases over time, it may be the case that, as O'odham is spoken less, this also leads to a less stable environment for the construction to evolve as it would be expected to.

The organization of this section is as follows: I present the distributions for the basic construction types (pronominal and adnominal), demonstrative types (distal and proximal) and phonological form (full or reduced) in §6.6.2. This is followed in §6.6.3 by an analysis of various distributions associated with exophoric, basic endophoric, and more grammatical demonstratives, and finally in §6.6.4, I analyze how the more grammaticalized uses of demonstratives are distributed in the O'odham data.

6.6.2 Distributions for basic construction types, demonstrative types, and forms

In this section, I present the overall distributions of all of the O'odham demonstratives in the data specifically with regard to the basic construction types (pronominal versus adnominal), demonstrative type (distal versus proximal), and phonological form (full versus reduced). The goal of this portion of the study is to provide a global view of the distributions over the three data periods and what they reveal in terms of grammaticalization trends.

Distributions of pronominal versus adnominal demonstratives

Pronominal and adnominal demonstratives both start off expressing the exophoric function, referring to entities in the surrounding physical context. After they evolve beyond this point, pronominal and adnominal demonstratives come to refer to entities stated within text itself and then later to serve more grammaticalized functions (see §§ 6.1.3 and 6.1.4). As either construction extends towards a wider range of functions and as those functions continue to grammaticalize, they are also expected to occur more frequently in the language over time.

A total of 634 tokens of demonstratives occur in the data (see Table 34 in §5.4). Figure 13 below depicts the frequency distribution in percentages of pronominal and adnominal demonstratives across the historical, midcentury, and modern data.

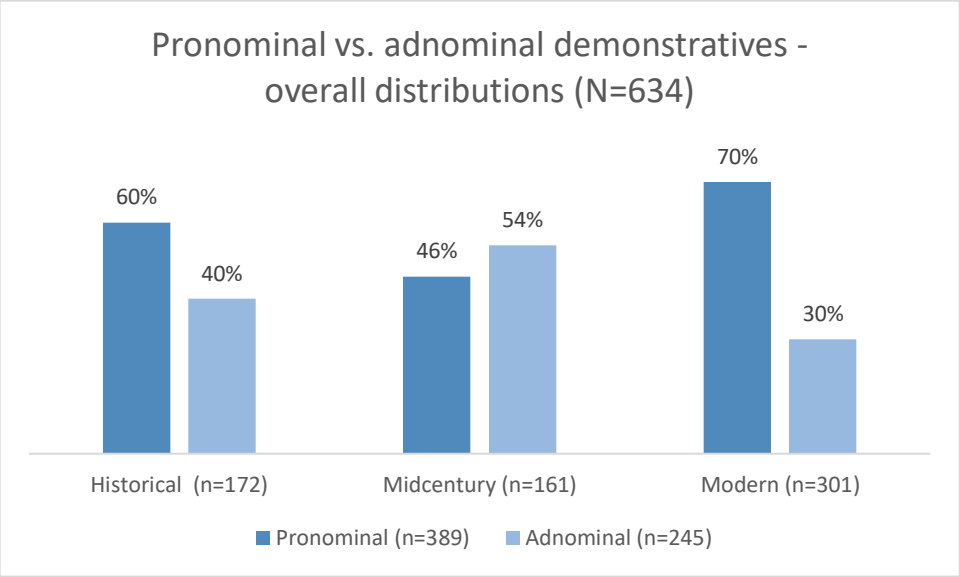


Figure 13. Overall percentages for pronominal vs. adnominal demonstratives.

While the historical and modern data appear to be similar in that pronominal demonstratives are clearly more common in both cases, the findings also show a 10% increase of the pronominal construction in the modern data (up from 60% to 70%). The midcentury data is somewhat of an outlier with the distribution of pronominal (46%) and adnominal

demonstratives (54%) being roughly equal. One possible explanation for this difference in the midcentury data may be due to the fact that they are comprised of a single collection of texts (*Coyote Stories,* Pancho & Mathiot, 1959-1960) told by a single speaker. These factors may accentuate certain patterns, thus skewing the results. In contrast, the historical and modern data include a variety of texts as well as multiple speakers. Looking at the overall numbers for all three periods, the increase in pronominal demonstratives in the modern data suggests that it has either come to be used for a wider range of functions and/or that those functions have strengthened and become more frequent. This is expected given that pronominal demonstratives serve several different more grammatical functions in O'odham: third-person pronouns, determinatives, connectives, referring postpositions, as well as placeholders. Subsequent findings show that third-person pronouns, determinatives, and place holders are the clearest contributors to this result. I address these details in §6.6.4.

I also implement a Chi-Square test on all three data periods combined in order to determine if there is a statistically significant association between construction type and data period. The null hypothesis is that no significant relationship exists between these two variables. The results are provided below.

The Chi-Square statistic is 26.7041. The *p*-value is < 0.00001. The result is significant at $p < .05$.

The results indicate a *p*-value less than .05 – a rejection of the null hypothesis, meaning in this case that the relationship between construction type and data period is significant.

Because the aberrant patterns described earlier for the midcentury data may be confounding the results, I also carry out a second test on only the historical-modern data pair. I want to determine if the difference in the distribution of construction types between the

historical and modern periods is significant, as this would confirm or reject whether or not the 10% increase in pronominal constructions is meaningful.

Historical and Modern data:
The Chi-Square statistic is 5.4745. The *p*-value is .019296. The result is significant at $p < .05$.

The *p*-value in this instance is less than .05 (.019296), rejecting the null hypothesis. The difference in frequencies of pronominal and adnominal construction between the historical and modern data is significant. The increase in pronominal demonstratives is probably not due to chance.

Distributions of distal versus proximal demonstratives

Next, I look at the distribution of distal and proximal demonstratives over the historical, midcentury, and modern data. Grammaticalization may also be reflected in the gradual rigidification of relationships between function and demonstrative type (Diessel, 1999, p. 118). As the relationships strengthen, they would also hypothetically occur more often in the data over time. Diessel's (1999) crosslinguistic study suggests that it is distal demonstratives that predominantly extend towards more grammaticalized functions (p. 118), although certain functions may align with proximal demonstratives instead. Figure 14 below displays the distributions in percentages for both demonstrative types across the three data periods.

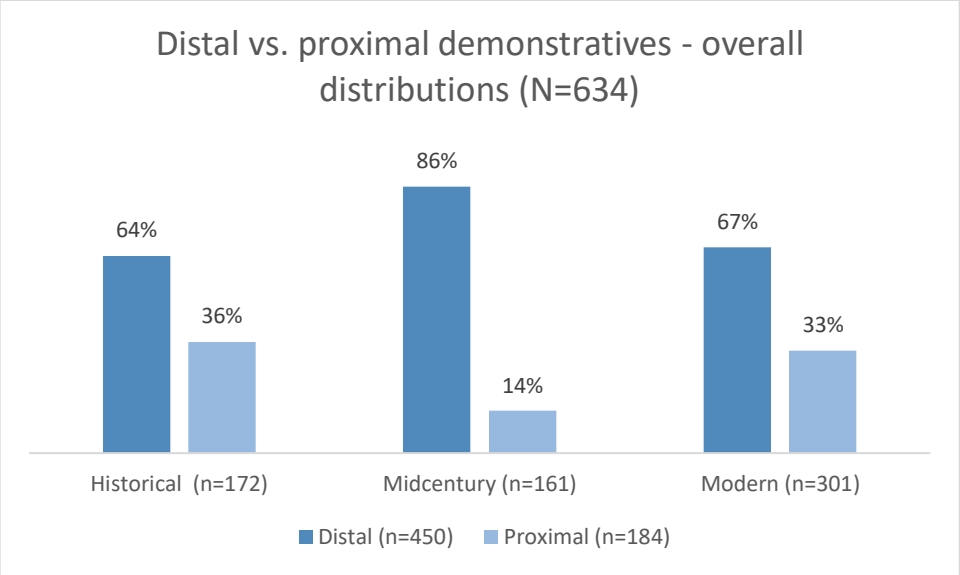


Figure 14. Overall percentages for distal vs. proximal demonstratives.

Overall, distal demonstratives occur with the greatest frequency for the historical (64%), midcentury (86%), and modern data (67%), and proximal demonstratives occur with lower frequencies (36%, 14%, and 33%, respectively). Distal demonstratives appear to serve a wider range of functions in O'odham than proximal demonstratives do in all three periods. This is in line with previous findings (Diessel, 1999). The most notable visual difference across the three data collections is that the modern and historical data show very similar distributions while the midcentury data employ distal demonstratives roughly 20% more frequently than the other two periods. Again, this difference in the midcentury data may be due to the nature of the texts – a single collection all told by the same speaker. It may the case that certain patterns are overrepresented, especially given the conspicuous similarity when only the historical and modern data are compared. For this reason, I run two different Chi-Square tests as I did with construction type in the previous section.

First, when I implement a Chi-Square test first for all of the data periods, the *p*-value is less than .05 (0.00001), showing that relationship between demonstrative type and the

periods is significant when all three data collections are tested. I also want to determine if the same result holds if I test the historical and modern data in isolation – again, because the midcentury data could be skewing the findings. The results for the historical and modern data are as follows:

Historical and modern data only:
The Chi-Square statistic is 0.3876. The *p*-value is .533563. The result is *not* significant at $p < .05$.

The *p*-value in this case supports the null hypothesis. When only the historical and modern data are tested, the results show no significant relationship between the variables of demonstrative type and data period in this case. I interpret this to suggest that the patterns for distal and proximal demonstratives are more-or-less stable over time.

Distributions of phonological form

As demonstratives advance further along a grammaticalization path and their original meanings become more general in the sense that they occur in more generalized contexts, they also tend to become more reduced in form. In other words, the expectation is that demonstratives would gradually correlate more often with phonologically reduced forms as time passes. The findings for O'odham are shown in the bar graph in Figure 15 below. These numbers include singular, pronominal demonstratives only (N=305). I exclude both plural demonstratives and all adnominal constructions given that they nearly always manifest in full forms in the data (see Table 33 in §6.2 and Table 37 in §6.4.1).

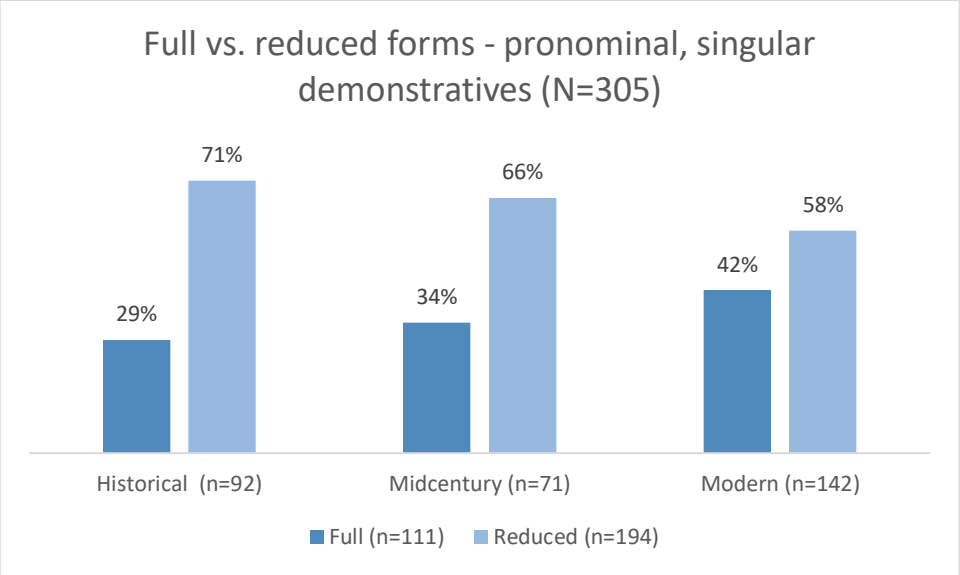


Figure 15. Overall percentages for full vs. reduced demonstratives.

Figure 15 reveals that the distribution of forms steadily changes. Full forms progressively *increase* (29%, 35%, and 42%, respectively) while reduced forms *decrease* (71%, 65%, and 58%). This goes against the expected pattern. Advancements in grammaticalization typically correlate with a *reduction* in form.

I again conduct two different Chi-Square tests, in this case not because the midcentury differs so much from the historical and modern data, but because the changes may be too gradual to detect over the span of the 50-year intervals. In fact, in the first test which I apply to all of the data, the *p*-value is .129216 in favor of the null hypothesis. No significant relationship exists between form and data period. The results of a separate Chi-Square test on only the historical and modern on the other hand show a statistically significant relationship between categorical variables:

Historical and Modern data:
The Chi-Square statistic is 3.9812. The p-value is .04601. The result is significant at $p < .05$.

While the changes between the historical and modern data are too gradual to be significant, the changes are detectable over a 100-year period. This suggests that the increase in full forms over time is not due to chance and that, while it may be that pronominal constructions have continued to grammaticalize, phonological reduction as an indicator of this process overall is not present. The question that remains is why full forms would consistently become more frequent and why *fluent* speakers would favor phonologically full forms over reduced forms, especially since either is an option to them. I will address this topic further after considering more of the results in §§6.6.3 and 6.6.4.

6.6.3 Quantitative analysis of discourse functions and more grammatical demonstratives

Next, I conduct a comparative analysis of the different distributions regarding the exophoric, basic endophoric, and more grammatical demonstratives across the three data periods. First, I look at how these three categories compare directly with one another, secondly, how the basic endophoric demonstratives (discourse deictic, tracking, and recognitional functions) distribute over time, and finally I consider phonological form once again, but in this case I look more closely at how the full and reduced phonological forms correlate separately with the categories of 'basic endophoric demonstratives' and 'more grammaticalized demonstratives.'

Overall distribution: exophoric, basic endophoric, and more grammatical demonstratives

I begin with the overall distribution of exophoric, basic endophoric functions (i.e. discourse deixis, tracking, and recognitional uses), and the more grammaticalized demonstratives in Figure 16 below. This sheds light on whether or not the overall frequency trends point towards more grammaticalized uses over time. This is the expected pattern based on the

hypothesis that as demonstratives evolve, they become more general in meaning and, in turn, more frequent. Based on this premise, exophoric demonstratives should be the least frequent in any period given that they express the narrowest of functions – to refer deictically to entities in the interlocutors' physical surroundings.

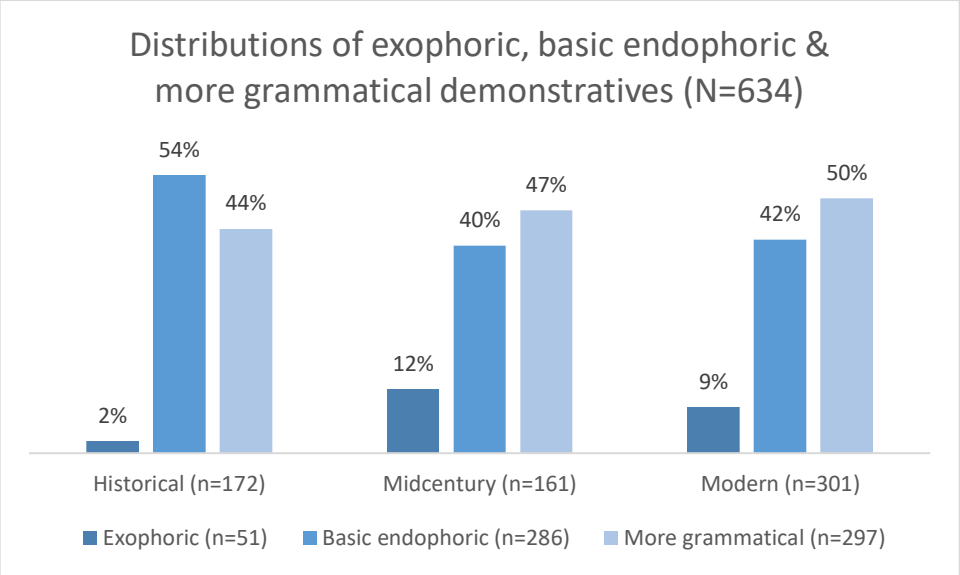


Figure 16. Percentages for exophoric, basic endophoric, and more grammatical demonstratives.

First, the comparative frequencies in Figure 16 above first confirm that exophoric demonstratives are, overall, the least frequent. While there is an increase in the frequency of exophoric demonstratives in the midcentury and modern data (12% and 9% of the total tokens, respectively) when compared with the historical data (2% of the total tokens), this likely can be attributed to the context and not necessarily due to exophoric demonstratives being less common in the language when the historical texts were recorded in the early 1900s (i.e. it is unlikely that the explanation is that speakers referred to physical entities less often in the past than speakers do today). Secondly, with regard to the distribution of basic endophoric demonstratives, they occur with the greatest frequency in the historical data –

over half (54%) of the total tokens for that period. This might suggest that while demonstratives were already clearly used for more grammatical purposes at that time, they may not have been quite as advanced, or certain functions that are used in the later periods are not used in the historical data. Thirdly, the more grammatical demonstratives show an incremental increase, comprising 44% of the tokens in the historical data, and 47% and 49% in the midcentury and modern data, respectively.

A Chi-Square test assesses if the changes in distribution are significant. This test is based on the frequencies for the basic endophoric and more grammatical demonstratives *only* (i.e. I exclude exophoric cases). At this point, I want to determine if a relationship exists between the data periods and changes in the distribution specifically for the basic endophoric and more grammaticalized demonstratives. Exophoric demonstratives do not play a vital role in the trajectory that demonstratives may take as they increasingly grammaticalize, while the basic endophoric demonstratives are the likely source from which all of the more grammatical uses evolve (see §6.1.4).

Again, I carry out two separate Chi-Square tests – first across all three data points and then for the historical and modern data only. While Figure 16 shows that the frequencies for more grammaticalized demonstratives gradually increase, the degree to which they do so may be too marginal for them to be statistically significant for the 50-year intervals, but the differences may be significant when I consider only the historical and modern data periods.

The first test on all of the three data periods is not significant (the *p*-value is .152361). In contrast, the second test which compares data over a 100-year period *is* statistically significant:

Historical and Modern data:
The Chi-Square statistic is 3.9812. The *p*-value is .04601. The result is significant at $p < .05$.

The result of the Chi-Square statistic is a *p*-value is less than .05, rejecting the null hypothesis; the grammatical status of the demonstratives and data period have a significant relationship. This indicates that demonstratives overall continue on a trajectory towards more grammatical uses based on the token frequency distributions.

Distributions for the basic endophoric demonstratives

Next, I present the distributions for the basic endophoric, discourse-pragmatic functions (discourse deixis, tracking, and recognitional) in Figure 17 below. The goal here is simply to provide a picture of each function's frequency in the data and what these numbers contribute towards the whole picture for O'odham demonstratives. These functions represent the earliest stages of the grammaticalization of demonstratives, which later expand to a number of more grammatical functions (see §§6.1.3 and 6.1.4).

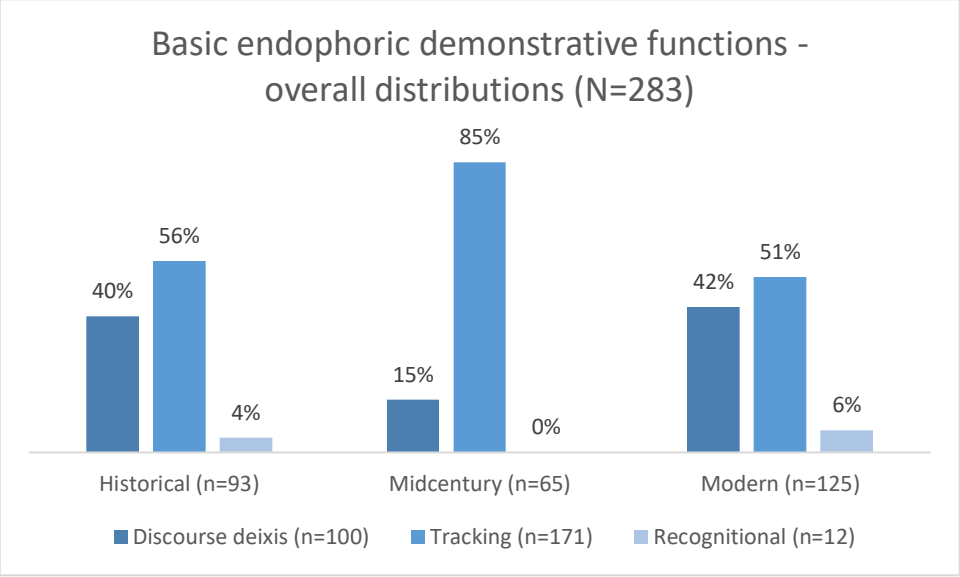


Figure 17. Percentages for basic endophoric demonstrative functions.

First, the distributions for the historical and modern data are quite similar with tracking demonstratives having the highest frequencies (40% and 42%, respectively), followed by

discourse deictic demonstratives (56% and 51%) and only a few tokens of recognitional demonstratives (4% and 6%). The midcentury data essentially follows this pattern but include a much larger percentage of tracking demonstratives (85%) and no examples of recognitional demonstratives.

Distribution of form for basic endophoric and more grammaticalized demonstratives

Due to the unexpected findings for the overall distributions for phonological form for all of the data in Figure 15 (§6.6.2) above, I also examine the distribution of full and reduced forms in a couple of additional ways. In this case, I separate the basic endophoric tokens (discourse deictic, tracking, and recognitional demonstratives combined) from the more grammaticalized tokens in Figure 18 and Figure 19 below to determine if this will produce different findings. As stated earlier, more grammaticalized structures tend to undergo more phonological reduction as they become more general and frequent in discourse. Based on this hypothesis, the *less* grammaticalized or basic endophoric (i.e. discourse deictic, tracking, and recognitional) demonstratives should have a higher number of full forms overall. The *more* grammaticalized uses, on the other hand, should have a higher rate of reduced forms. Regarding the data periods, the historical data are expected to have the lowest rates of reduced forms, the modern data the highest rates, and the midcentury data should fall somewhere in between. In both cases, I examine the frequencies solely for singular, pronominal demonstratives. Both plural and adnominal demonstratives nearly always manifest in their full forms (see Table 33 in §6.2). Including them in this portion of the analysis would confound the results.

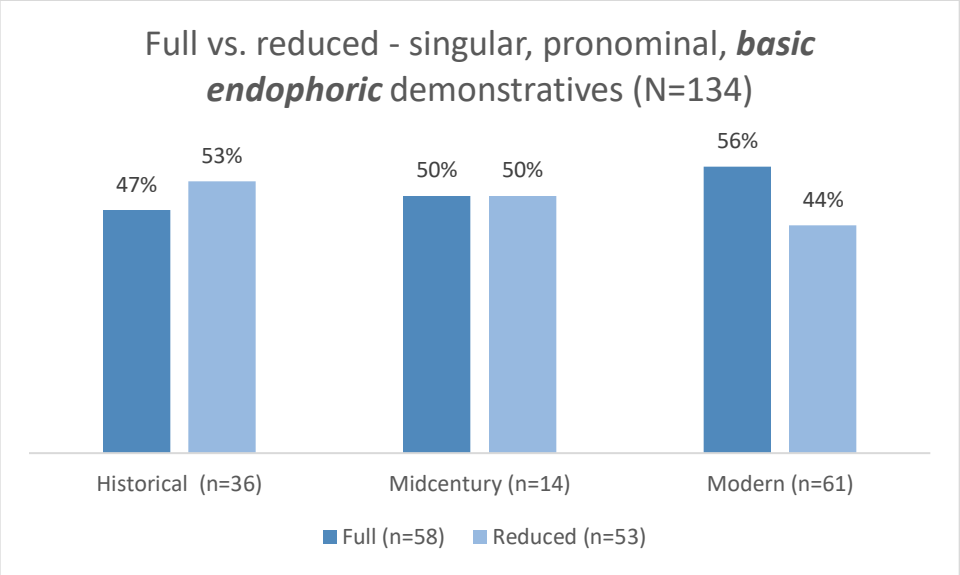


Figure 18. Overall percentages for full vs. reduced *basic endophoric* demonstratives.

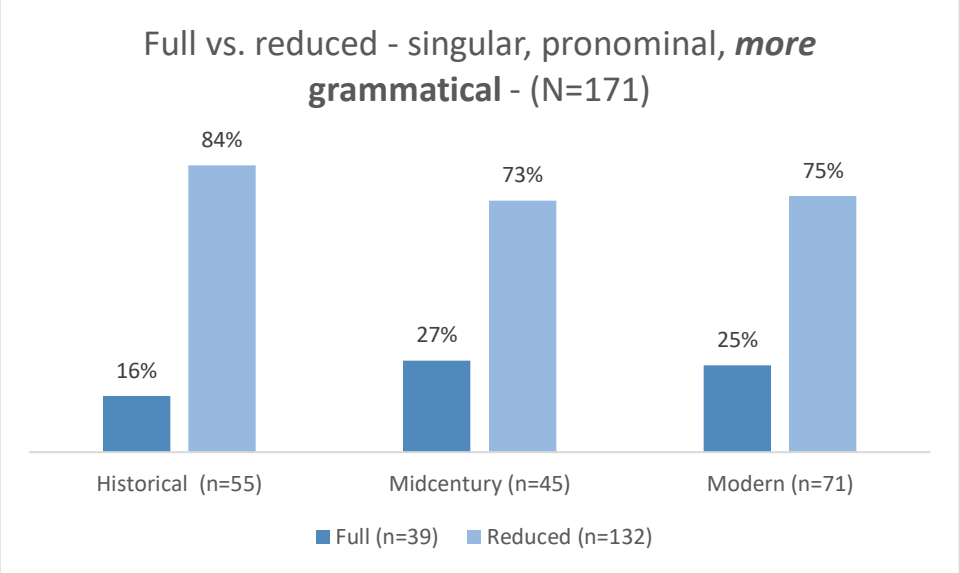


Figure 19. Overall percentages for full vs. reduced *more* grammatical demonstratives.

The distribution of form is roughly equal for the basic endophoric functions in Figure 18, while reduced forms clearly occur more frequently – overall – with the more grammaticalized demonstrative functions in Figure 19, as expected. In terms of the three data periods, the distributions appear fairly stable across time for the basic endophoric demonstratives. In the case of the more grammaticalized functions, however, full forms occur

more frequently, counter to what would be expected, in the midcentury and modern data for the more grammatical functions. This pattern is similar to some extent with the early findings for the overall distribution of forms in Figure 15 (§6.6.2).

In order to assess if any of the relationships are statistically significant, I carry out a two-way Chi-Square test for both the basic endophoric and more grammatical demonstrative functions. When all three periods are tested, the results are statistically insignificant for both basic endophoric functions and form, and more grammatical functions and form (the *p*-values are .70802 and .379271, respectively, both in favor of the null hypothesis). The separate results for only the historical and modern data are shown below:

Basic endophoric demonstratives and form (historical and modern data):
The Chi-Square statistic is 0.6584. The *p*-value is .417123. The result is *not* significant at $p < .05$.

More grammaticalized demonstratives and form (historical and modern data):
The Chi-Square statistic is 1.4872. The *p*-value is .222654. The result is *not* significant at $p < .05$.

In both cases, again, no significant relationship exists between form and the data period with regard to the basic endophoric demonstratives and the more grammaticalized functions when only the historical and modern data are tested. This suggests that, while O'odham clearly utilizes demonstratives for a series of functions, several of which are at more advanced stages in their evolution, phonological form has changed very little when looked at narrowly for both basic endophoric and more grammaticalized functions. That said, the fact that the pattern for form and the more grammatical demonstratives in particular in Figure 19 more-or-less aligns with what we saw in Figure 15 (§6.6.2) where the *overall* distribution of phonological form are statistically significant, the question remains as to why this trend would appear. Again, this counters the expectation that more grammaticalized functions should correlate with more phonological reduction over time. Before providing a possible

hypothesis for why this might be the case, next I look in greater detail at the more grammaticalized functions, and I also consider how full forms distribute even more narrowly with third-person pronouns to determine if a similar trend continues to hold (see Figure 21 below).

6.6.4 A more detailed analysis of the more grammaticalized demonstratives

In the previous sections (§§6.6.2 and 6.6.3), I looked broadly at token distributions and various correlations and what they might suggest about grammaticalization trends. In this section, I look more narrowly at the distributions of some of the more grammaticalized demonstrative functions in conjunction with basic endophoric demonstratives out of which the former probably evolve (see §6.1.4) as well as with other variables (e.g. demonstrative type, construction, form, etc.), where relevant. Like the previous quantitative examinations, the goal is to determine if these patterns strengthen over time in a manner that would be expected based on how demonstratives tend to evolve. If the patterns are counter to what is expected, different possible explanations can be explored, including the possible effects of language shift. For a few of the grammatical functions, the token frequencies are low, making it difficult to make meaningful generalizations. That said, some important observations about the data can be made in some cases based on the period(s) in which the tokens occur, construction types, form, etc. Once again, the more grammaticalized demonstrative categories in O'odham include: 1) third-person pronominal reference, 2) determinatives, 3) connectives, 4) referring postpositions, 5) specific-indefinite demonstratives and, finally, 6) placeholders. Earlier in § 6.4.3, I laid out the criteria I use to identify each of these functions in O'odham and, in §6.5.2, I presented numerous examples of each function based on my descriptive analysis.

Third-person pronouns

Third-person pronouns likely evolve from pronominal tracking demonstratives. The former anaphorically indicate active referents while the latter anaphorically point to semi-active, and in some cases, inactive referents in the hearer's consciousness (see §6.1.4). The distinction between third-person pronouns and tracking demonstratives is sometimes ambiguous in O'odham, suggesting that the evolution of pronominal demonstratives in this regard continues to be an ongoing process (see §6.2). In the descriptive portion of results (§6.5.2), I illustrate that both distal and proximal demonstrative types may play this function, although the former are more common. This finding is likely due to the fact that distal demonstratives, overall, have a higher token frequency (see Figure 14 in §6.6.2). In order to assess ongoing trends and the degree to which pronominal demonstratives have extended towards a third-person pronoun meaning over an approximately 100-year timespan, I look at the distribution of third-person pronouns in comparison to all of the basic endophoric types, and I also examine how third-person pronouns correlate over time with full and reduced phonological forms.

Beginning with third-person pronouns and the less grammaticalized, basic endophoric demonstratives, Figure 20 below reflects the frequency distributions in terms of percentages. Diessel argues that all more grammaticalized demonstratives evolve directly from the basic endophoric functions. This comparison provides a big picture view of whether or not third-person pronouns have increased in frequency, a possible indicator of that they have continued to advance towards more grammaticalized status.

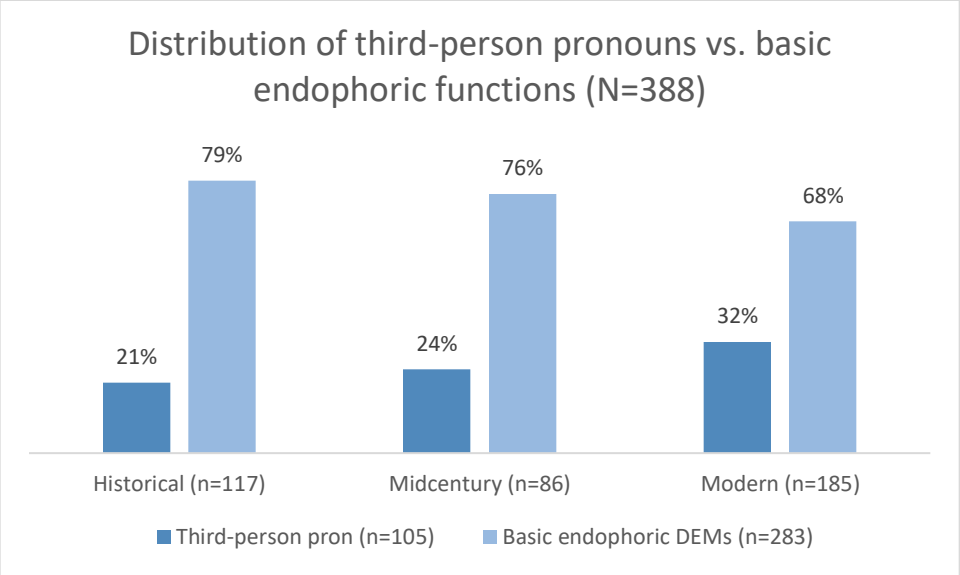


Figure 20. Percentages for third-person pronouns vs. basic endophoric functions.

Figure 20 shows that third-person pronouns are on a steady upward trend with a 21% token frequency the historical data followed by 24% in the midcentury data and 32% in the modern data. This suggests that pronominal forms continue to advance further towards a third-person pronoun function and may account, in part, for the finding in Figure 13 (§6.6.2), which shows an increase in the overall distribution of pronominal constructions.

The results of a Chi-Square test for all of the data show no significant relationship between the demonstrative functions and data period. The *p*-value is .062347. Once again, when we test only the historical and modern data, however, which represent a timespan of 100 years instead of just 50-year increments, the results *are* significant.

Historical and Modern data:
The Chi-Square statistic is 5.0717. The *p*-value is .02432. The result is significant at $p < .05$.

Pronominal demonstratives appear to continue their expansion towards the third-person pronoun function, which indicates active referents in discourse.

Next, I examine the distribution of phonological form – full versus reduced – for third-person pronouns. The analysis is limited to singular pronouns, which manifest in either full or reduced forms. A strong relationship between the third-person pronoun meaning and a phonologically reduced form would suggest more advanced stages of grammaticalization based on the hypothesis that, as meaning becomes more bleached and functions expand, forms also tend to gradually reduce over time (Bybee et al., 1994, p. 19). It would also suggest a developing morphological split between pronominal demonstratives that serve the tracking function and the more grammaticalized third-person pronoun function. Figure 21 below presents the percentages for full and reduced singular forms of third-person pronouns for each data period.

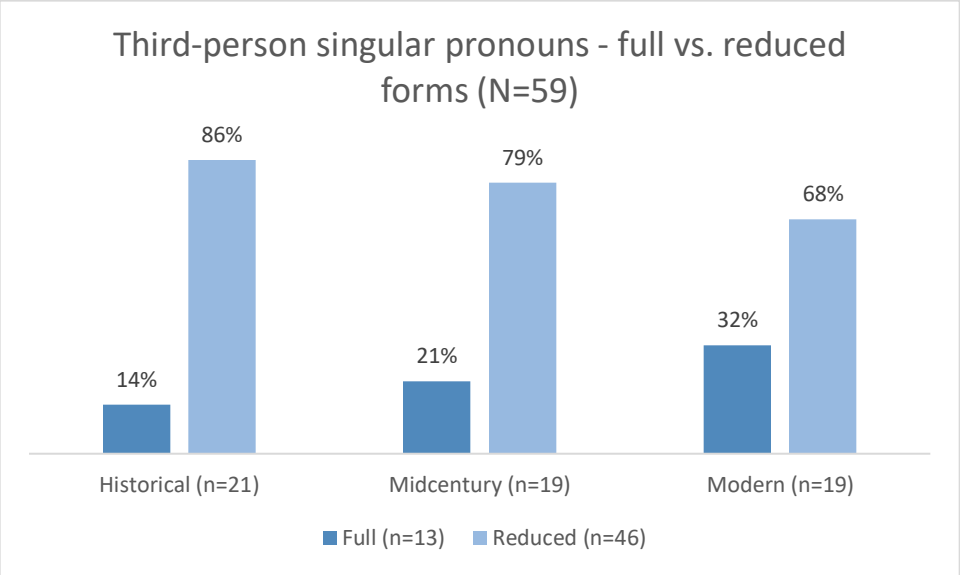


Figure 21. Percentages for full vs. reduced forms for singular third-person pronouns.

Across all three data periods, phonologically reduced demonstratives (86%, 79%, and 68%) are clearly more frequent than their full counterparts (14%, 21%, and 32%); however, reduced forms gradually decrease in frequency over time, and full forms increase, contrary to what we would expect. This again mirrors the overall trend we saw earlier in Figure 15 and

to some extent in Figure 19. That said, statistically, the relationship between phonological form and data collection is not significant (*p*-value is .416408). When only the historical and modern data are tested, the result is the same: not significant, as reflected below.

Historical and modern data:
The Chi-Square statistic is 1.7107. The *p*-value is .190891. The result is *not* significant at $p < .05$.

Statistically, the full and reduced forms are not significantly different, although the more frequently occurring full forms contributes to the earlier finding for all of the data (Figure 15 in §6.6.2) and also to the finding for all of the more grammaticalized functions combined (Figure 19). The fact that this pattern emerges when the phonological form is looked at in three different ways is worthy of further examination. I address this issue more extensively in the next section.

Further consideration of the increase in phonologically full forms

In the preceding sections, the findings suggest that phonologically full demonstrative forms occur more frequently over time rather than reduced forms (see Figure 15, Figure 19, and Figure 21). This is counter to the other finding that demonstratives appear to continue to grammaticalize in O'odham (Figure 16; also Figure 20 above), based on the premise that more advanced stages of grammaticalization more often correlate with a gradual reduction in form. The question that remains is why this unexpected finding would be the case. In order to explore this further, I first consider possible morphosyntactic as well as phonological explanations. If neither of these offers a clear explanation for the finding, then sociolinguistic factors can be considered.

Earlier, I pointed out that morphosyntactic distinctions such as subject and object functions can eventually develop a contrast in form (see end of §6.4.1). One function may

gradually reduce in form as it grammaticalizes while the other does not reduce or at least not to the same degree (e.g. Bybee, 2015). A potential outcome is that different forms may gradually correlate with specific morphosyntactic functions; however, this does not appear to be the case in O'odham (see §6.4.1). Reduced forms can function as either subjects or objects throughout the data (see (72)-(74) in §6.5.2). In order to test this further, I look more closely at how full and reduced forms distribute with subject and object third-person pronouns as shown below in Figure 22 and Figure 23.

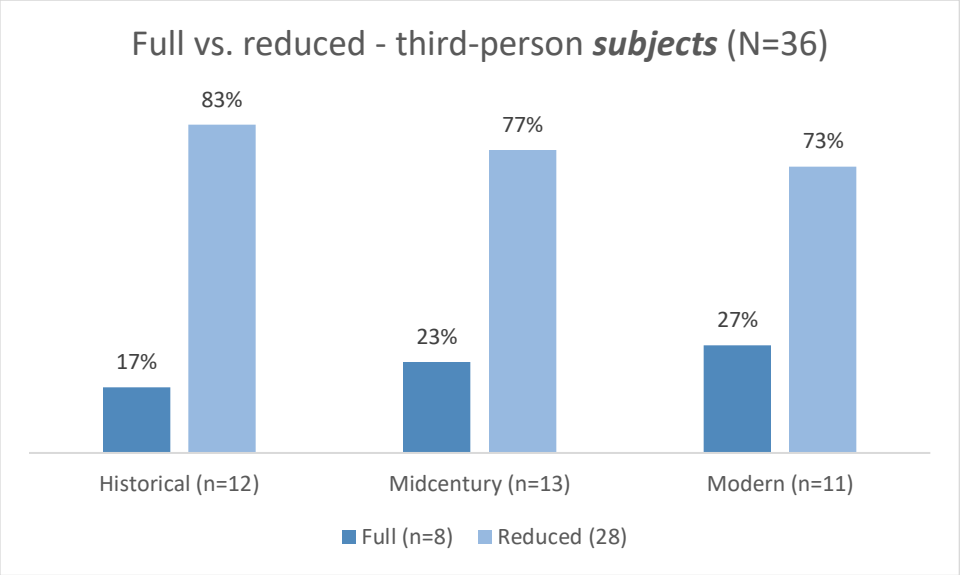


Figure 22. Percentages for full vs. reduced forms for singular, third-person *subjects*.

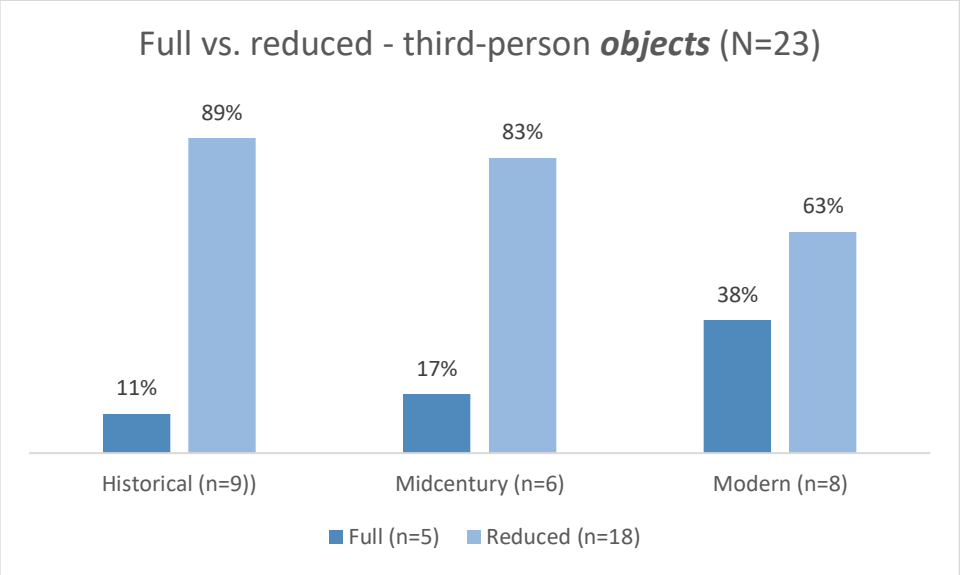


Figure 23. Percentages for full vs. reduced forms for singular, third-person *objects*.

As the figures above reflect, subjects and objects show little difference with respect to phonological form. In both cases, reduced forms are preferred, suggesting that form is attached to neither subject nor object categories. Furthermore, the same trend that was seen earlier across the three data periods is also observed here; full forms increase slightly over time whether for subjects or objects. This also suggests that changes to form in association with either morphosyntactic category is not present.

The other possibility that must be considered is whether or not form can be attributed to the location in the utterance. Previous studies show that words (including function words) in utterance initial and final positions tend to be longer while words in medial positions tend to be shorter (Bell et al., 2002; also Klatt, 1975). Based on this premise, we want to determine if phonologically full forms of demonstratives occur most frequently in utterance initial and final positions in especially in the modern data. If they do, then we can possibly attribute the patterns seen in Figure 21 (and by way of implication to the findings in Figure 15 and Figure 19) to prosodic rather than social explanations related to language shift.

In order to determine if this is the case, I examine all third-person pronoun tokens for their position in an utterance. I follow Bell et al.'s (2002) definition for **utterances** which are "sentence-like units which often make up spoken conversation, and hence are defined with respect to … syntactic coherence" (p. 1019) which in turn provides an approximation of "large intonation boundaries" (p. 1019). The authors base this relationship on Croft's (2005) finding that grammatical boundaries often correlate with intonation boundaries. In the results below, I group together initial and final positions based on previous findings that full forms are more likely to occur in both of these positions and in more reduced forms in medial position. Figure 24 below shows the findings for full forms and, for comparison, Figure 25 includes the findings for reduced forms.

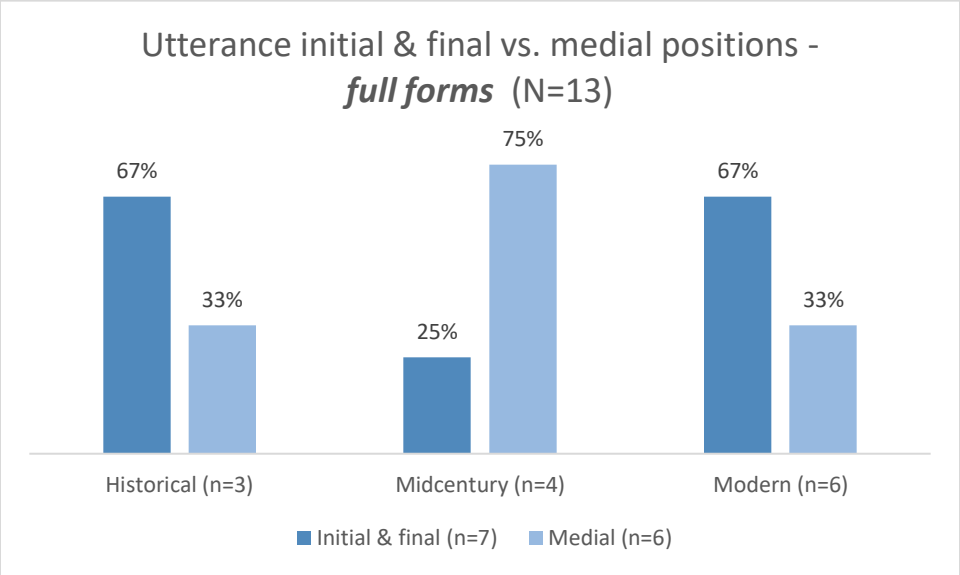


Figure 24. Percentages for utterance initial & final vs. medial positions *full forms* (singular, third-person pronouns).

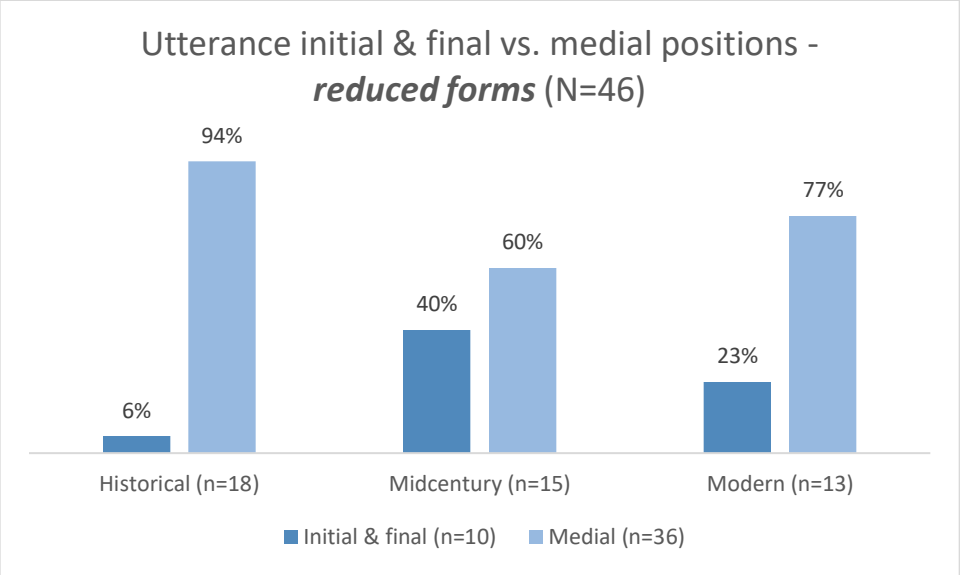


Figure 25. Percentages for utterance initial & final vs. medial positions *reduced forms* (singular, third-person pronouns).

Overall, full forms occur infrequently for third-person pronouns (N=13) as noted earlier for Figure 21. Regarding their distributions in utterances, Figure 24 shows that full forms occur the least in the midcentury data and in the historical and modern data equally; they do not appear more often in the two prominent utterance positions, initial and final, in the modern data than they do in the historical data. Figure 25 show us that reduced forms occur most often in medial position across the data, as expected. This sampling suggests that the frequency with which full forms are produced in prominent positions probably does not account for why full forms increase over time with more grammaticalized functions overall and for third-person pronouns more specifically.

After eliminating these two variables – morphosyntactic functions and utterance placement of full forms – social explanations due to decreased use of O'odham may account for the gradual increase in phonologically full forms. Given that this study deals with fluent speakers, a possible hypothesis is that this finding may be due to the speakers accommodating their speech patterns, particularly in the modern data, with the knowledge

that many O'odham people are less proficient in O'odham today than they once were. Previous studies have shown, for example, that speakers who intentionally employ "clear speech" – that is, speech patterns that are slower, louder, and with more careful articulation – benefits non-native speakers' comprehension (e.g. English: Bradlow & Bent, 2002; Croatian and English: Smiljanić & Bradlow, 2005). By extension, the same may also be true of less fluent native speakers of O'odham. Based on the fact that fully fluent speakers of O'odham today are more often in settings where they must modify their speech in order to increase comprehension than in the past, this could be a possible explanation for the findings that demonstratives increasingly are produced in their phonologically full forms.[40] The fact that the modern data is largely comprised of YouTube recordings, speakers may have to some extent used clear speech while being at least somewhat conscious of the fact that different listeners with a wide ranges of fluency levels would likely be viewing the videos in the future. It would be interesting to investigate if this effect shows up elsewhere in O'odham – not just with demonstratives but in additional contexts.

Determinatives

Determinatives, as illustrated in §6.5.2, are obligatory markers of either the head of a relative clause or function as the head of the relative clause itself (Diessel, 1999; also see §§6.1.4 and 6.4.3). Schematically, the construction is represented as DEM or DEM N + [*m*-AUX CLAUSE]$_{\textbf{REL}}$. Determinatives are devoid of any other discourse pragmatic meanings such as the discourse deictic, tracking, or recognitional functions (i.e. they are purely grammatical at this stage) and are hypothesized to evolve from the latter. The distal demonstrative is preferred in nearly

[40] Thank you to Naomi Shin and Jill Morford for suggesting and discussing this idea with me and for the suggested readings.

all cases across the data, with the exception of 6 tokens. Furthermore, pronominal demonstrative constructions are pervasive for determinatives across the data, again with only a few exceptions (1 out of 9 tokens in the historical data, 0 tokens in the midcentury data, and 4 out of 21 tokens in the modern data are *not* pronominal). These are another contributor to the findings for the overall distribution of pronominal demonstratives shown in Figure 13 (§6.6.2). The demonstrative always occurs in its phonologically full form, showing no change in this regard. Because these variables are fairly consistent for determinatives, in this case I am only concerned with how they distribute in comparison to the basic endophoric categories in Figure 26 below.

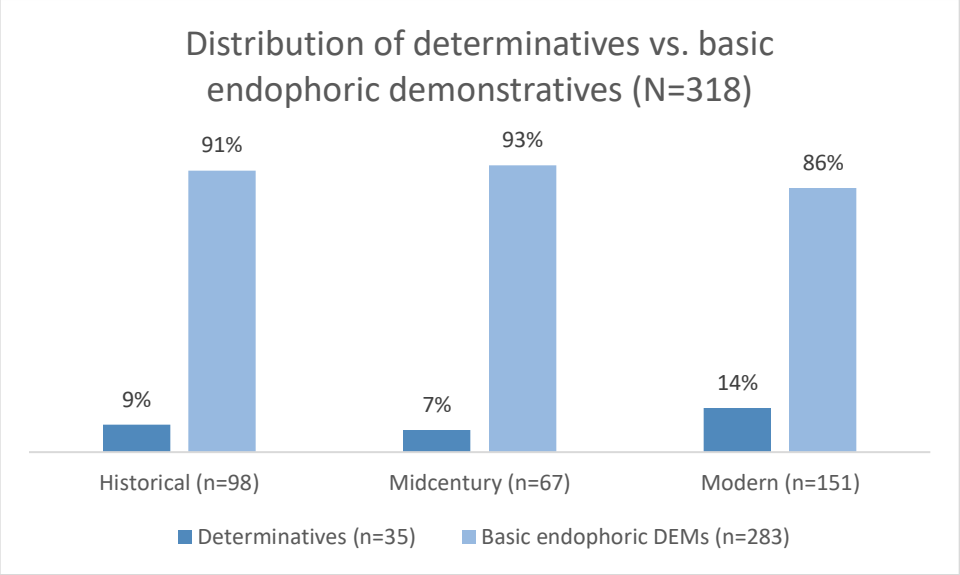


Figure 26. Percentage of determinatives and endophoric demonstratives.

The figure above shows relatively similar distributions of determinatives with 9% and 7% in the historical and midcentury data, respective, and then a slight increase to 14% in the modern data, which may suggest that the function continues to spread in O'odham, although this claim is highly tentative. More conservatively, determinatives are stable across data periods, which two different Chi-Square tests support. The results of the first test on all three

periods is not significant (the p-value is .195636). Since the percentage of determinatives increases in the modern data up to 14%, I also test the historical data period with the modern data in isolation:

Historical and modern data:
The Chi-Square statistic is 1.7457. The *p*-value is .186412. The result is *not* significant at $p < .05$

The differences in distribution between the historical and modern data are not significant. The determinative function appears to be stable over time, showing very little change.

Connectives

Connectives link two propositions together and, in this manner, they are related to discourse deixis. In O'odham, they are composed of a demonstrative and postposition, most commonly *hekaj* which in isolation can be glossed as 'with', 'by means of', 'because of' and, on occasion, the postposition *'amjeḍ* 'from' or 'about'. When a demonstrative occurs with one of these postpositions, the meanings that are conveyed are, 'as a result of', 'as a consequence of ', etc. I describe several examples in §6.5.2. Connectives always appear in the fixed construction DEM POSTP. This finding is true across all of the data periods. Additionally, singular demonstratives are always in reduced form in the data, most often the distal forms *heg* or sometimes *g*, and in a couple of cases the proximal *'id*. Both the fixed construction type and reduced phonological forms suggest that that connectives are fairly advanced and that this may have been true even at the time the historical data was collected in the early 1900s. In this portion of the study, I only consider its overall distribution in comparison to the basic endophoric demonstratives.

A total of 48 tokens express connective functions. The overall distribution in percentages for the historical, midcentury, and modern data is reflected in Figure 27.

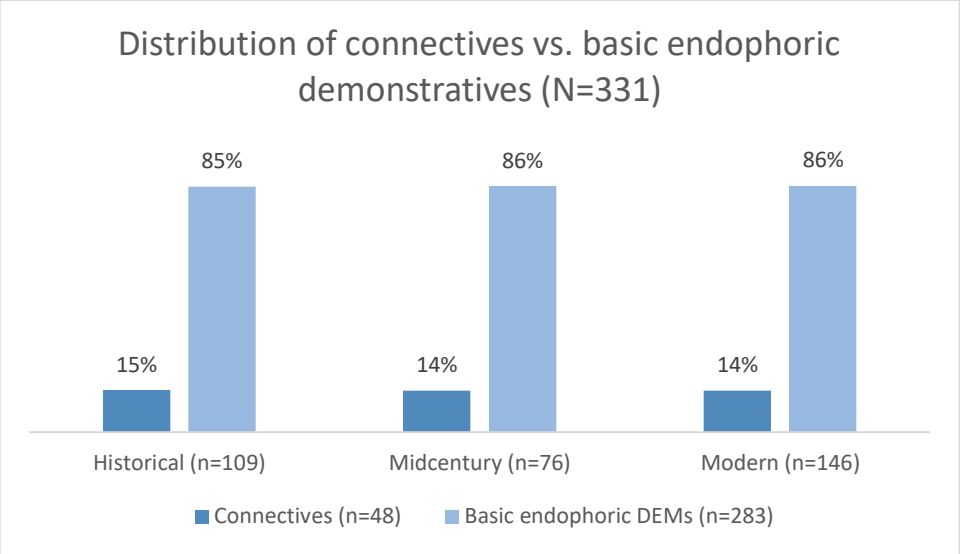


Figure 27. Percentages for connectives.

The distribution is unequivocally stable across the data with very little change in distributions, with 15% of the tokens in the historical data and 14% of the tokens for both the midcentury and modern data.

Referring postpositions

As described in §6.5.2, referring postpositions are similar to connectives in that a demonstrative and postposition always co-occur. The construction takes the shape of either DEM POSTP or, when the referent is explicitly stated, DEM POSTP N. A wider range of postpositions may be employed in comparison to connectives. Singular demonstratives are always in reduced form and, while proximal demonstratives are possible, distal demonstratives are the most pervasive in the data, with only a few exceptions.

The overall distribution of referring postpositions is shown in Figure 28 below. There are a total of 41 tokens of referring postpositions.

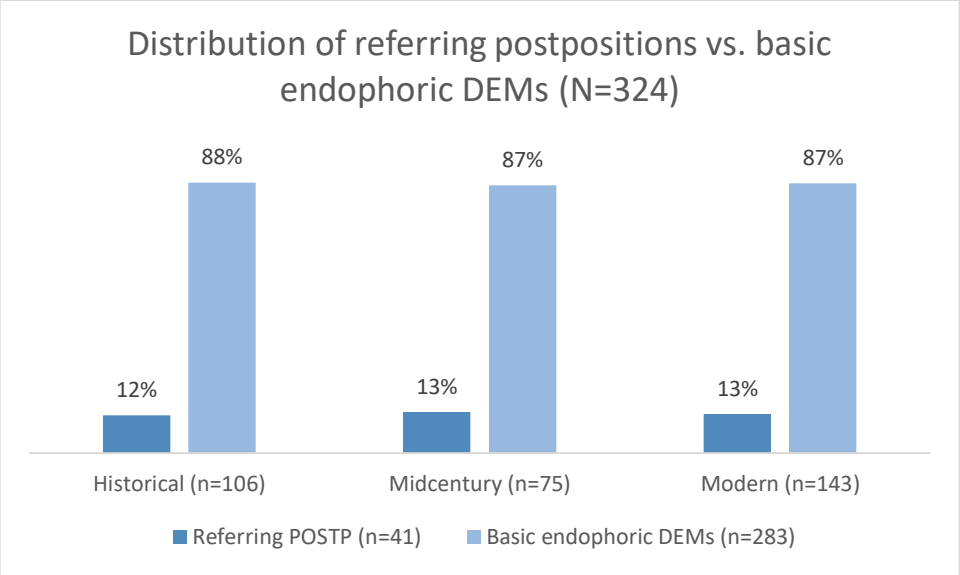


Figure 28. Percentages for referring postpositions.

Referring postpositions also appear to be stable, with very little change in their distributions over time.

Specific indefinite demonstratives and place holders

Finally, the last two more grammatical functions, specific indefinite demonstratives and placeholders, occur in limited contexts. Both occur with low token frequencies. Despite this fact, they show some noteworthy patterns, which I discuss in this section.

First, specific indefinite articles indicate a particular referent that is both new to the discourse and to the hearer (see §6.1.4). The function always employs an adnominal demonstrative construction in line with the notion of 'article.' Specific indefinite articles appear in the data a total of three cases, all of which occur in the modern data (I describe all three examples in §6.5.2). As discussed in §6.5.2, the development of this usage may be the result of contact with English. In all cases, the proximal demonstratives *'i:da* 'this' or *'idam* 'these' are adopted for this function in O'odham as it also is in spoken colloquial English even though cross-linguistically this development is extremely rare (Wald, 1983; see also Diessel,

1999); the numeral one most often grammaticalizes into this function (Givón, 1995; Wright and Givón, 1987). This fact coupled with the observation that this use only appears in the modern data points to a possible contact-induced development. A larger corpus would shed more light on this construction's evolution.

Finally, placeholders serve as a stand-in for a nominal referent that the speaker initially has difficulty retrieving but eventually is able to do so (see §6.1.4). This use occurs only in the modern data in six instances, several of which were described in detail in §6.5.2. The proximal demonstrative is always adopted in this case, and it is always in phonologically full form. Whether or not this function is a recent development is difficult to state with certainty. This may be the case, based on its appearance only in the modern data; however, given its function and the fact that the language is spoken less often in the late-shift period of which the modern data represents, there is another plausible explanation. Recall that in nearly all of the examples cited in the modern data (§6.5.2), after using the placeholder, the speakers utilize an English lexical item instead of an O'odham one. It may be that speakers today are more likely to experience lexical retrieval issues and/or simply do not know if an O'odham counterpart for a specific English word exists and in the end choose to use the English word. It is possible that placeholders may appear only in the modern not because it is a new development, but because there is an increased need for the placeholder function in discourse today. Speakers of O'odham have fewer opportunities to both hear and use the language, thus resulting in greater challenges in the lexical domain.

6.7 Discussion

The descriptive analysis in §6.5 illustrates that demonstratives have clearly expanded from their original exophoric functions to serve a number of additional functions with the basic

endophoric, discourse-pragmatic functions hypothetically representing the earlier stages of grammaticalization which then in turn extend towards a wide range of more grammaticalized functions. The quantitative analysis in §6.6 provides a more fine-grained analysis of how relevant patterns distribute across time, and to what extent demonstratives have continued to evolve over approximately the last century during which the O'odham population has increasingly shifted towards English. In this section, I discuss what these findings suggest regarding the effects, if any, this scenario has had on the expected chain of developments that demonstratives may move through based on previous crosslinguistic studies (see §6.1.4). The most pervasive finding is that language shift appears to have had little disruptive impact on the grammaticalization processes associated with demonstratives. In several cases, demonstratives appear to continue to evolve when frequency distributions are considered together with the statistical findings, and in most other cases, the results point to structures that remain stable over time. Only in one regard – phonological form – are the patterns not what we would expect. The best overall interpretation is that as language shift has progressed in the O'odham community this has had little impact on the evolution of demonstrative functions. The changes that may be a result of language shift are in the lexical domain and changes in production. In what follows, I discuss how the data leads to this conclusion.

The first piece of evidence that suggests that demonstratives do, in fact, continue to evolve in O'odham despite the shifting language scenario is the overall distribution of pronominal demonstratives (Figure 13). Pronominal demonstratives increase in frequency when compared to the historical and modern data, and the results of a Chi-Square test show that the relationship between construction type and data period is statistically significant. This is the case whether the statistical test is run on all the historical, midcentury, and modern

data combined or only on the historical-modern pair given the divergent patterns for the midcentury data which may be due to how uniform the texts are and that they are told by a single speaker, resulting in an overrepresentation of certain patterns. Based on the wide range of functions pronominal demonstratives tend to play in O'odham (third-person pronouns, determinatives, connectives, referring postpositions and placeholders), the significant increase in pronominal constructions is the expected pattern.

Another possible indicator of the ongoing evolution of demonstratives are the findings for the distribution of all of the grammaticalized functions combined (Figure 16). Their distribution shows an increase, and the results of a Chi-Square test on the historical and modern data are significant.

Finally, third-person pronouns (Figure 20) increase and, again, a Chi-Square statistics tests on the historical and modern data alone indicate that the differences between data periods are statistically significant. This finding also contributes, at least in part, to the overall distributional increase in pronominal constructions noted above.

Regarding stable structures, the distribution of demonstrative types (distal versus proximal) changes a negligible amount, specifically when a Chi-Square test is implemented on only the historical and modern data (see statistics in relation to Figure 14). The results suggest that the relationship between demonstrative type and data period is not significant. The differing patterns in the midcentury data may again be due to the nature of the texts.

Moving on to more grammaticalized functions, determinatives show a very slight increase in the modern data (up by 5 percentage points and 7 percentage points from the historical and midcentury data, respectively; see Figure 26) and a majority of tokens are

pronominal rather than adnominal. That said, the statistical results are not significant; these findings suggest that they change very little over time.

Next, connectives and referring postpositions show virtually no change in the data in several regards. The distribution of connectives (Figure 27) when compared with the basic endophoric functions show a 1% frequency decrease in the midcentury and modern data from the historical data, while referring postpositions (Figure 28) increase by 1% in the midcentury and modern data. Both functions always occur with reduced singular demonstrative forms, and their syntactic constructions appear to be rigidly fixed as the many examples in §6.5.2 illustrate. It may be the case that these constructions were already well established in O'odham prior to the historical period and they continue to play their functions as they have for some time.

Two other more grammatical functions which clearly increase over time are the specific-indefinite article and placeholders. O'odham speakers use both of these functions *only* in the modern data.

Beginning with the former, I identify three cases of specific-indefinite articles in the modern data. These may have developed based on a similar construction in colloquial English (i.e. new-*this*), indicating a contact effect. As I discussed in §§6.5.2 and 6.6.4, O'odham employs the proximal demonstrative to refer to specific, yet previously unmentioned referents as English does. Cross-linguistically, however, proximal demonstratives as a source for this function is uncommon. Usually it is the numeral 'one' that evolves into this function (Givón, 1995; Wright and Givón, 1987; also see Diessel, 1999).

Six tokens of the placeholder demonstrative function are utilized in the modern data. These are always pronominal (see end of §6.6.4) and also contribute in part to the earlier

noted increase in pronominal demonstratives in Figure 13 (§6.6.2). Whether this use is a relatively new development is not entirely clear. Given that the function of placeholders is to serve as a stand-in for a nominal lexical item that the speaker initially has difficulty retrieving, it may the case that its appearance in the modern data is simply due to circumstance – that is, due to the fact that O'odham is used less often than it was in the past, which leads to different types of issues in the lexical domain. With increasingly fewer opportunities to speak O'odham, speakers are more likely to a) experience difficulties recalling certain types of lexical items in O'odham, or b) have fewer opportunities to create new equivalent terms that exist in English, especially if they are more specialized terms as is the case in examples (102)-(104) in §6.5.2. Recall that in the end, terms are borrowed from English in the examples. In either case, the results suggest lexical issues that are likely due to the shifting linguistic landscape where English is more pervasive. More data for all three data periods may shed more light on whether or not this is a recent development in O'odham.

Finally, the most conspicuous and least expected findings are regarding phonological form. Full forms of singular, pronominal demonstratives increase while reduced forms decrease. This is the case when I examine all of the singular, pronominal demonstrative tokens across the data (Figure 15), when looking more narrowly at forms for all of the more grammaticalized functions combined (Figure 19), and also when examining third-person pronouns (Figure 21); however, only in the first case – overall distributions of form – are the results statistically significant. In the least, the increase in full forms in more grammaticalized constructions and third-person pronouns provide a partial explanation for the overall distributional increase of pronominal demonstratives across the data.

A look at a sample of full and reduced forms does not suggest that the morphosyntax is a motivating factor as to whether or not speakers produce one form over the other. Furthermore, speakers do not appear to use phonologically full demonstrative forms in prominent positions more often in the modern data when compared especially to the historical data. Instead, I posited that this finding could be due to fluent speakers accommodating their speech to the needs of hearers who may be less proficient in O'odham. Full forms could be the result of fluent speakers employing clear speech more often, as previous findings have suggested, for example, in the context of discourse between native and non-native speakers of a language (Bradlow & Bent, 2002; Smiljanić & Bradlow, 2005).

In sum, there are few indicators that demonstrative grammaticalization processes have been disrupted, or that the structures are deteriorating. These results may suggest that demonstratives may be more impervious to the impacts of severe language shift than other areas of grammar given the overall stability that they show in many different areas in which they operate.

Chapter Seven: Concluding remarks

7.0 Introduction

Over the last century, the O'odham community north of the US-Mexico border in Arizona has largely shifted towards English due to the coercive, assimilationist policies of the United States. Within this context, my primary goal for this usage-based study has been to closely look at whether or not this scenario of language shift may explain some of the changes or lack of changes that appear in O'odham. I have done so especially in light of how languages are expected to change based on what we know from previous cross-linguistic, typological studies on grammaticalization as well as changes that may require explanations outside of grammaticalization. In order to achieve a fine-grained analysis of change, I have looked closely at two specific areas of grammar – progressive aspect and demonstratives. In what follows, I summarize my primary findings for these studies and then discuss the contributions that this dissertation makes to the field and, finally, I explore possible avenues for future investigation.

7.1 Summary of findings

The findings have shown that the most pervasive pattern is that both progressive constructions and the myriad functions demonstratives can play either tend to change in predictable ways or remain stable over time, showing very little change. Speakers continue to use these structures as they have since at least the early 20th century. The one exceptional case may point to speaker-hearer accommodation. Evidence also points to a possible contact-induced development. Changes that may genuinely be due to infrequent language use are related to certain lexical tasks. These combined findings show that language shift has had

little impact on the evolution of progressive constructions and demonstratives and, while there may be some losses in the lexical domain, speakers are at the same time creating new structures to meet today's communicative needs. I elaborate on each of these points below.

7.1.1 Observed changes in progressive constructions

Progressive constructions in the data in most cases appear stable, changing very little over time. Progressive constructions favor highly agentive subjects, although non-agentive subjects are possible across all three data periods. Similarly, dynamic predicates co-occur with the progressive marker *-him* more often while stative predicates occur less often in all three data periods. Finally, I look at the distribution of motion-type meanings expressed by progressive constructions as a whole. While the distributions were similar across time for most of the motion types, progressive constructions with non-motion meanings occur most frequently in the modern data compared to the historical and midcentury data. Although the relationship between the categories of motion type and data period proves to be statistically insignificant, this finding may tentatively indicate that progressive constructions continue to expand and grammaticalize towards this more less specific meaning. This trend would be expected based on the hypothesis that grammaticalization often entails semantic generalization. Furthermore, the fact that motion-type meanings display the clearest indication of change may suggest that this is a stronger predictor of how the progressive aspect evolves when the source is a motion verb, as it is in O'odham.

Another change in the data occurs in the lexical domain; however the change is not restricted to the context of progressive constructions. Generic predicates co-occur with the progressive marker *-him* more frequently in the modern data, and the relationship between specific versus generic-meaning predicates and data period is statistically significant. The

same trend appears to be true of verbs across the language when a verb sample from the historical and modern data is examined. This indicates that it is not collocation trends between the progressive and generic versus specific-meaning verbs that has changed, but that the change appears to have taken place across the language as whole. This finding may be related to lexical retrieval tasks. Generic verbs can be used in a wide range of contexts and can often be used in lieu of a specific-meaning verb if enough contextual information is available to inform the hearer of the action that the generic verb represents. While speakers may have the appropriate verb with the narrower meaning stored in their lexicon, generic verbs are easier to access, especially if a person neither hears nor uses O'odham on a regular basis. Additional changes in the lexical domain also occur in the study of demonstratives. I will address these in the next section.

7.1.2 Observed changes in demonstratives

In the case of demonstratives, several of the more grammaticalized functions similarly appear to remain stable more often than not, in some cases showing virtually no changes in their distributions (i.e. connectives and referring postpositions). Where changes are visible among the demonstrative functions, they most often fall in line with what we know about how demonstrative constructions tend to evolve in the world's languages based on previous typological studies. More grammaticalized constructions when looked at in their entirety significantly increase in frequency over time, as do third-person pronouns when looked at in isolation. Although not statistically significant, determinatives also show a slight percentage increase in more recent data.

The most unexpected finding is that phonologically full demonstrative forms become more frequent as time passes while reduced forms become less frequent. This seemingly

contradicts the finding that demonstratives overall continue to expand towards more grammatical functions. However, when we consider the context that the modern fluent speakers of O'odham are using in the language – one where there are far fewer fully fluent speakers than in the past – it is possible that they are accommodating their speech patterns in order to increase the chances of being fully understood. Studies on the communication between native and non-native speakers, for example, provide evidence for more careful speech production in such contexts (Bradlow & Bent, 2002; Smiljanić & Bradlow, 2005).

Demonstratives also show a change in the lexical domain particularly with the placeholder function which appears solely in the modern data. Speakers utilize the placeholder demonstrative function (in any language) precisely because they momentarily cannot retrieve an intended nominal lexical item. In O'odham, the proximal demonstrative *'i:da* plays this function. In most cases, a speaker will eventually produce the nominal referent in English drawing on the other linguistic system that is available to them. This may be either because they know that an O'odham equivalent exists but they cannot recall it, they came to the conclusion that the equivalent probably does not exist in O'odham, or even that they were attempting to create an O'odham equivalent on the spot. Whatever the case, this is a sign that decreased use of O'odham in the community has led to either lexical loss in O'odham (which is compensated with English) or fewer opportunities to create equivalent terms for specialized concepts (e.g. *thumb drive, doctoral degree,* etc.).

The fact that the placeholder function only exists in the modern data can be explained in two possible ways. The first is this may be due to the fact that speakers need such a device more often today given the language shift that has taken place. Another possibility is that the function is, in fact, a relatively recent grammatical development.

A clearer instance of a new use of the demonstrative is the specific indefinite article which also is utilized only in the modern data. In fact, evidence suggests that this use function may have developed based on the existence of a similar construction in spoken colloquial English (described in Diessel, 1999; Himmelmann, 1996; Wald, 1983; Wright & Givón, 1987). While it is cross-linguistically rare for a specific indefinite article to grammaticalize from a demonstrative (Diessel, 1999; Wright & Givón, 1987), this is precisely what has occurred in both O'odham and English. Furthermore, it is the proximal demonstrative that is employed for this function in both languages. These parallel patterns point to a possible externally motivated development in O'odham.

If placeholders and the specific indefinite article are, in fact, recent developments, these then are examples of new structures being created despite fewer speakers and fewer domains in which O'odham is spoken. In the case that the specific indefinite article is based on its existence in English, this is an example of speakers actively and creatively combining information from two of the languages at their disposal.

7.2 Contributions of the dissertation

This dissertation offers several contributions to the field. First, in order to situate and contextualize this study, I have described a number of the historical events and policies that have resulted in the current endangered status of O'odham against the community's will. This provides a number of details that are specific to the O'odham case. Based on some of this information and on additional clues in earlier writings, I have also constructed a timeframe for the different phases of language shift, beginning with the early-shift period of the early 1900s, the middle-shift period of the mid-1900s, and finally the late-shift period of the late 1900s-early 2000s of which the O'odham community is currently in the midst. The timeline

provides a picture which suggests that while different colonizers of the region aimed to convert and assimilate O'odham people and other Native American groups, the education system correlates most strongly with the commencement of more accelerated loss of speakers in the O'odham community. While these findings will not be unsurprising to most, they offer clear evidence linking language loss to colonialism. The same chapter also shows that the O'odham community continues to highly value their language, viewing it as sacred, and are working to reverse the language shift that has transpired due to the trauma that several generations have experienced.

Next, I provided a detailed description of the typological features of the language. I described the language's phonological profile as well as its mildly synthetic morphological typology. I also described how O'odham codes for subjects and objects. While the former are coded in the auxiliary system, objects are marked on verbs. Next, I discuss how verbs mark for different aspectual categories. For example, the distinction between imperfective and perfective aspects is most commonly indicated by the presence or absence of word-final consonants/vowels. I addressed how O'odham shows a split nominative-accusative/ergative-absolutive system and also provided an overview of the language's relative free word order with regards to S, V, and O and other elements that are more fixed such as the second position of the auxiliary. Finally, I briefly described how overt morphosyntactic marking is located on the head of a phrase with regards to possessives as well as objects of transitive clauses. While all of these patterns have previously been observed in O'odham, and I largely based my discussion of O'odham structure on these previous descriptions, the information tends to be spread out across multiple sources. The description that I provide in this work not

only familiarizes the reader with the language, but also centralizes information on the language's typological characteristics for linguists who are interested in language universals.

I have also laid out in detail my methodology for how the materials I used for this dissertation were processed in collaboration with native speakers of O'odham. The improvements have not only made this dissertation possible, but they will also benefit community members on both sides of the US-Mexico border who may want to use them, for example, for educational purposes. The historical texts in particular offer a vantage point on O'odham history entirely from the O'odham perspective at a time when many community members feel that some types of older knowledge are being forgotten. The translations into both English and Spanish will make this knowledge available especially to those who do not speak O'odham or who speak very little of the language.

In cases where archival materials for other Native American languages exist, they can similarly be processed following the methods that I implemented. Such materials offer rich, culturally salient information that many communities seek to maintain. Additionally, based on the fact that many of the earliest records of other Native American language records were collected in the beginning of the twentieth century, processing archival materials has an added value for scholarship. It now may be possible to carry out more diachronic studies comparing historical records from approximately 100 years ago, as I did for this dissertation, with more recent language records of other Native American languages. Few studies of this nature exist, and Native American languages tend to be underrepresented in the literature when compared to studies on Indo-European languages, for example. Such work could increase Native American representation in the field.

Finally, the two studies that I have carried out in this work themselves contribute to our understanding of several different issues. First, it provides an in-depth description of progressive and demonstrative constructions in O'odham based entirely on authentic use of the language. Usage-based studies are otherwise few in number for O'odham. It also contributes to studies of grammaticalization processes in O'odham, which are also fairly limited in the literature, and it offers another account of what change in these specific constructions looks like in comparison to other languages that have been looked at in previous cross-linguistic studies. Finally, this dissertation has also explored areas of grammar that have been looked at very little in the context of language shift, offering insights into how this scenario may or may not impact how they change.

7.3 Topics for future research

In closing, I discuss two possible avenues for future research. First, in the study of demonstratives, I explain that one possible explanation for the increase in phonologically full forms instead of reduced forms over time is that speakers may be accommodating their speech in order to increase the comprehension of less proficient speakers. It would be worth carrying out a study on this topic alone. There are other areas of the language where alternate full and reduced forms are possible in O'odham. For example, O'odham's rich auxiliary system has a number of full and truncated variants as do first and second pronouns, both of which I described in §3.3. One question that remains is whether or not similar patterns would emerge with such structures in modern language use. While studies on 'careful speech' have been studied in the context of native speaker/non-native speaker communication, to my knowledge, this topic has not been considered in the context of endangered languages, particularly those that have a wide range of speaker types.

Finally, another area that I have not explored here that should be looked into in the future is speaker-specific patterns of change. For example, if I were to compare how individual speakers use progressive and demonstrative constructions, the findings might highlight a different set of details. Previous studies have suggested that when different speakers are compared to one another they tend to show more variation than usual on phonetic, phonological, as well as structural grounds (e.g. Campbell & Muntzel 1989; Tsunoda, 2005; Babel, 2008). This is interpreted to be a reflection of the fact that speakers have fewer chances to both hear, use, and practice a given language.

Appendix A: Comparative chart of O'odham orthographies[41]

IPA	Alvarez-Hale (1970)	Saxton (1982)	Saxton, Saxton & Enos (1983)	Mathiot (1973)	Mason (1950)	Dolores & Kroeber[42]	Tojono O'otam (Sonora)[43]
i iː	i iː	i ii	i ih	i ii	i iˑ	i, y[44] īˑ[45]	i ii
ɨ ɨː	e eː	ɨ ɨɨ	e eh	e ee	e eˑ	e, ü ē	e ee
u uː	u uː	u uu	u uh	u uu	u uˑ	u ū	u uu
ɔ ɔː	o oː	o oo	o oh	o oo	o oˑ	o ō	o oo
a aː	a aː	a aa	a ah	a aa	a aˑ	a ā	a aa

Table 40. Vowels.

[41] Special thanks to Andrés Sabogal who created an earlier version of both this and the consonantal chart. All errors are my own.
[42] Based primarily on Dolores and Mathiot (1991); also Mason (1950).
[43] Provided by Andrés Sabogal.
[44] Following or between vowels (Dolores & Mathiot, 1991, p. 236).
[45] Long vowels not always marked by Dolores (Dolores & Mathiot, 1991, p. 235).

( ) = less frequently used orthographic representations

IPA	Alvarez-Hale (1970)	Saxton (1982)	Saxton, Saxton & Enos (1983)	Mathiot (1973)	Mason (1950)	Dolores & Kroeber	Tojono O'otam (Sonora)
p	p	p	p	p	p, (ʙ)[46]	p, (hp)	p
b	b	b	b	b	b	p, ('p)	
t̪	t	t	t	t	t,(D)	t, (ht)	t
t̪ː	tt	tt	tt	tt	--	't	
d̪	d	d	th	d	d	t, ('t)	
d	ḍ	ḍ	d	ḍ	ḍ	d, (t)	th
tʃ	c	č	ch	c	tc, (DJ)	tc	ch
dʒ	j	ǰ	j	j	dj	tc, 'tc	
k	k	k	k	k	k, (G)	k	k
g	g	g	g	g	g	k, q,[47] ('k)	
ʔ	'	'	'	'	'	' /not marked	'
s	s	s	s	s	s	s	s
ʂ	ṣ	ṣ	sh	x, ṣ [48]	c[49]	c	sh
h	h	h	h	h	h	h, (ȟ)	j
ɾ	l	l	l	l	l , (L)	l, ly, li	r
m	m	m	m	m	m, (M)	m	m
n	n	n	n	n	n, (D)	n	n
ɲ	ñ	ñ	ni	ñ	ñ, (Ñ)	n, ní, ny, (ñ)	ñ
w ~ v	w	w	w	v[50]	v, w	v, w, we	v
j	y	y	y	y	y	--	y

Table 41. Consonants.

[46] Mason (1950) used small capitals when voiced stops become voiceless in word-final position and preceded by a glottal stop (p. 9).
[47] "When a velar stop, either fortis or lenis, is followed by the vowel /i/ or /ï/, Dolores used the symbol *q*" (Dolores & Mathiot. 1991, p. 238).
[48] (Dolores & Mathiot, 1991, p. 237).
[49] Mason (1950) states that <c> presents a slightly further back sound than <s> (p. 9); presumably he was referring to the retroflex.
[50] Mathiot refers to this as a bilabial fricative (ɸ/β) (Dolores & Mathiot, 1991, p. 237).

Appendix B: Sample O'odham text

Title: *Wenog mat 'e-nu:nu:c g mo'abdam*
'At the time when Hunter Turned Into a Buzzard'
‘Cuando el Cazador se convirtió en Zopilote’

Narrated by: Robert Cruz[51]

Original video recording (*Hunter Turns into Buzzard)* available on YouTube:
https://youtu.be/_4y01V3wX_E

Date: April 2, 2012

English translation: Robert Cruz[52] and Keiko Beers
Spanish translation: Jacob Franco Hernández
Morphological glosses: Keiko Beers

(1) Ku-t-ṣ 'amhu hebai ṣa'i si me:k wipia-m
CONN-3.AUX.PFV-QUOT over.there sometime very INTS far hunt-go.PFV

g mu'abḍam.
ART hunter
'So, he, the Hunter, went on a very long hunt.'
‘Se dice que él, el Cazador, fue a una cacería muy larga.’

(2) T-ṣ hab hi:,
3.AUX.PFV-QUOT thus go.PFV

mu'i taṣ k 'ab pi ṣa'i 'abamk ha'icu-doakam.
many day CONJ LOC NEG very lucky animal
'So, many days passed and he was not lucky to encounter any wildlife.'
‘Así que pasaron muchos días y no tuvo suerte de encontrar algún animal salvaje.’

[51] The transcribed version of the story I provide here is based on Mr. Cruz's original transcription which is available on YouTube (link provided at top of page). I have changed some of his original spellings and added morphological breaks for the sole purpose of being consistent with the conventions I use throughout this work.
[52] Mr. Cruz determined what consisted of a line, indicated by the line numbers. Commas in the O'odham text indicate a non-clause final pause.

(3) 'Am him-c-ka-him ku-t-ṣ 'oia
LOC walk-CONC-STAT-PROG.IPFV CONN-3.AUX.PFV-QUOT consequently

ṣa'i si bihugim k 'p ṣa'i si tonom kc gewko.
very INTS hungry CONJ also very INTS thirsty CONJ get.tired.PFV
'As he was walking, he then consequently became very hungry and also very thirsty and very tired. '
'Por consiguiente cuando él estaba caminando le dio mucha hambre y también mucha sed y se cansó mucho.'

(4) 'I noḍ k 'u:hum 'i hi: 'am 'e-hemajkam ha-wei.
PUNC turn.PFV CONJ back PUNC walk.PFV LOC 3.POSS-people 3PL.OBJ-to
'He turned around and began to walk back to his people.
'Se dio entonces la vuelta y empezó a caminar de regreso hacia su gente.'

(5) 'Am him 'u:hum kc ḍ si bihugim-kam kc ḍ si gewkog-kam.
LOC go.IPFV back CONJ COP INTS hungry-NMLZ CONJ COP INTS get.tired-NMLZ
'On his way home he was very hungry and very tired.'
'Ya de regreso él estaba muy hambriento y muy cansado.'

(6) K 'oia 'am hu hebai m-at-ṣ,
CONJ then somewhere SBDR-3.AUX.PFV-QUOT

'e-ta:t g mo'abḍam,
3.REFL-feel.PFV ART hunter

m-at-ṣ aṣ 'i hawk g,
SBDR-3.AUX.PFV-QUOT just PUNC lightweight ART

him-ij.
walk-3SG.POSS
'And then at some point, Hunter just felt that his walking became effortless,'
'Y entonces en algún momento, el Cazador sintió que su caminar se aligeraba,'

(7) 'e-ta:t m-a-ṣ 'aṣ s-hawk.
3.REFL-feel.PFV SBDR-3.AUX.IPFV-QUOT just STAT-lightweight
'he felt that he was weightless.'
'él sentía que estaba muy liviano.'

(8) 'Am him-c 'aṣ 'i 'u:gka,
LOC go.IPFV-CONC just INCEP rise

m-a-ṣ 'am him g mo'abḍam.
SBDR-3.AUX.IPFV-QUOT LOC go.IPFV ART hunter
'As he walked the Hunter began to rise.'
'Entonces mientras él caminaba el Cazador empezó a surgir.'

(9) K 'am hu hebai m-at-ṣ 'e-ta:t
CONJ somewhere SBDR-3.AUX.PFV-QUOT 3.REFL-FEEL.PFV

m-at-ṣ 'aṣ 'u:g da:.
SBDR-3.AUX.PFV- QUOT just high fly.SG.PFV
'And at some point he felt himself fly high.'
‘Y en algún momento se sintió volar alto.’

(10) Ku-ṣ gnhu 'u:g da'a kc,
CONN-QUOT over.there high fly.SG.IPFV CONJ
'He was flying high and,
‘Él estaba volando alto y,

(11) Nhu 'u:g da'a kc haha 'aṣ,
over.there high fly.SG.IPFV CONJ then just

ñei-k,
see-COMPL

ḍhu 'e-weco m-a-ṣ g,
over.there 3.REFL-under SBDR-3.AUX.IPFV-QUOT ART

cu:hug gḍhu wo'o kc ṣa'i si s-'i'owim ma:s,
meat over.there be.lying.down CONJ extremely INTS STAT-tasty appear
'he was flying high up and then he noticed beneath him that meat was lying down there, and it looked very delicious,'
‘estaba volando alto y entonces se dio cuenta que debajo de él había carne tirada por ahí, y se veía muy sabrosa,’

(12) ṣa'i si s-wegĭ kc,
very INTS STAT-be.red CONJ

g,
ART

gi:gĭ-ij si s-toha.
tallow-3SG.POSS INTS STAT-white
'it was very red and the tallow was very white.'
‘estaba muy roja y la manteca era muy blanca.’

(13) Ku-t-ṣ 'am 'i huḍ,
CONN-3.AUX.PFV-QUOT LOC PUNC descend.PFV

g mo'abḍam k,
ART hunter CONJ

si hu: g cu:hug.
INTS eat.PFV ART meat
'So, he descended, and Hunter ate the meat to his satisfaction.'
‘Así que, él bajó, y el Cazador se comió la carne para su satisfacción.’

(14) 'Am 'i ko:wod k 'ab 'i dahiwa k 'am 'i ñei g,
LOC PUNC become.full CONJ LOC PUNC sit.down CONJ LOC PUNC see.PFV ART

cu:hug, ku-t-ṣ 'oia ce: g,
meat CONN-3.AUX.PFV-QUOT then discover.PFV ART

mo'abḍam ma:s g,
hunter appear ART

haiwañ jewalig 'am ko'ihim.
cow rotten.object LOC eat.PROG.IPFV
'He became full and he sat down and saw the meat, and Hunter saw that he had been eating a rotten cow.'
‘Se llenó y se sentó y observó la carne, entonces el Cazador se dio cuenta que se había estado comiendo una vaca podrida.’

(15) Hegai gi:g m-a-ṣ si s-toha ḍ 'aṣ waptop
DIST tallow SBDR-3.AUX.IPFV-QUOT INTS STAT-white COP just maggot.PL

m-at-p taṣ 'ama'i wo'o-ka-him hegai
SBDR-3.AUX.PFV-ASSUMP a.long.time there be.lying.down-STAT-PROG.IPFV DIST

haiwañ mu:k.
cow dead
'That tallow that was very white were just maggots, that it had probably been lying there for a long time, that dead cow.'
‘y que esa manteca que estaba muy blanca eran solo gusanos, y que probablemente esa vaca muerta había estado ahí por mucho tiempo.’

(16) Ñea,
see.DM

ku-t hab masma hab 'e-ju g
CONN-3.AUX.PFV in.this.certain.manner thus 3.REFL-do.PFV ART

mo'abḍam,
hunter,
'So, that's how it happened to Hunter,'
‘Y fue así como le sucedió al Cazador,

(17) m-at-ṣ 'e-ñuwi-c k aṣaba
SBDR-3.AUX.PFV-QUOT 3.REFL-turkey.buzzard-RES CONJ however

ge'e 'e-ṣo'ig.
much 3.REFL-suffer.PFV
'that he had become Buzzard but had suffered greatly.'
‘que él se convirtió en Zopilote pero habiendo sufrido mucho.’

(18) K ḍ si has ha'icu 'i:da,
CONJ COP INTS what.ABST thing(s) PROX

jeweḍ kegcud-dam.
land clean-NMLZ
'and so this is very important, this land-cleaner.'
‘de manera que éste es muy importante, este limpiador de la tierra.’

(19) Ñui 'ot si has ha'icu 'am t-wehejeḍ
turkey.buzzard 3.AUX.PFV INTS what.ABST thing(s) LOC 1PL.OBJ-for

'a:cim 'O'odham.
us 'O'odham
'Buzzard is very important for us O'odham.
‘El Zopilote es muy importante para nosotros los O’odham.’

References

Adams, D. W. (1995). *Education for extinction: American Indians and the boarding school experience, 1875-1928.* University Press of Kansas.

Aikhenvald, A. (2010). *Language contact in Amazonia.* Oxford University Press.

Aikhenvald, A. (2012). Language contact in language obsolescence. In I. Léglise & C. Chamoreau (Eds.), *Dynamics of contact-induced language change* (pp. 77-110).

American Indian Policy Review Commission. (1976). *Report on Indian education: Final report to the American Indian Policy Review Commission*. US Government Printing Office.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*(1), 65-87.

Ariel, M. (1990). *Accessing noun-phrase antecedents.* Routledge.

Atkins, J. D. C. (1888). *Correspondence on the subject of teaching the vernacular in Indian schools*. Government Printing Office. https://library.si.edu/digital-library/book/correspondence00unit

Auer, J. C. P. (1984). Referential problems in conversation. *Journal of Pragmatics*, *8*, 627-648.

Babel, M. (2008). The phonetic and phonological effects of moribundity. *University Of Pennsylvania Working Papers in Linguistics* 14(2), 25-34.

Bahr, D. M. (1983). Pima and Papago social organization. In A. Ortiz (Ed.), *Handbook of North American Indians: Southwest* (Vol. 10, pp. 178-200). Smithsonian Institution.

Batibo, H. (1992). The fate of ethnic languages in Tanzania. In M. Brenzinger (Ed.), *Language death: Factual and theoretical exploration with special reference to East Africa* (pp. 85-98). Mouton de Gruyter.

Bavin, E. L. (1991). Warlpiri in the 1980s: An overview of research into language variation and child language. In S. Romaine (Ed.), *Language in Australia.* Cambridge University Press.

Beers, K., Cruz, R., Hirrel, L., & Kerfoot, I. (2014). Describing reduplication patterns in Tohoho O'odham with language learners in mind. In B. Annible, K. Beers, L. Hirrel, & C. Occhino (Eds.), *Proceedings of the High Desert Linguistics Society conference* (Vol. 10, pp. 43-55). High Desert Linguistics Society.

Beers, K., Hirrel, L., Kerfoot, I., & Sabogal, A. (2011). *O'odham vowel processes: An investigation of vocalic change in reduplication and suffixation* [Unpublished manuscript]. Department of Linguistics, University of New Mexico, Albuquerque.

Bell, A., Jurafsky, D., Fosler-Lussier E., Girand, C., Gregory, M., & Gildea, D. (2002). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America, 113*(2), 1001-1024. doi: 10.1121/1.1534836#

Booth, P. M. (2005). If we gave up making of *nawait*, it would mean starvation: Saguaro wine defenders of Tohoho O'odham land and way-of-life. *The Journal of Arizona History*, *46*(4), 375-396.

Boudahmane, K., Manta, M., Antoine, F., Galliano, S., & Barras, C. (1998-2008). *Transcriber* [Computer Software]. http://trans.sourceforge.net/en/presentation.php

Brenneman. (2014). Bringing O'odham into the "Pimería Alta": Introduction. *Journal of the* Southwest, *56*(2), 205-218.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America, 112*(1), 272-284. https://doi.org/10.1121/1.1487837

Brown, C. H. (1999). *Lexical acculturation in Native American languages.* Oxford University Press.

Bruder, J. S. (1977). Changing patterns in Papago subsistence strategies: Archaeology and ethnohistory compared. *The Kiva*, *42*(3), 233-256.

Bühler, K. (1934). *Sprachtheorie.* Fischer.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.

Bybee, J. (2003a). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2, pp. 145-167). Lawrence Erlbaum Associates Publishers.

Bybee, J. (2003b). Mechanisms of change in grammaticization: The role of repetition. In R. Janda & B. Joseph (Eds.), *Handbook of historical linguistics* (pp. 602-623). Blackwell Publishers.

Bybee. J. (2015). *Language change.* Cambridge University Press.

Bybee, J., & Pagliuca, W. (1985). Cross-linguistic comparison and the development of grammatical meaning. In J. Fisiak (Ed.), *Historical semantics, historical word formation* (pp. 59-83). Mouton.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world.* University of Chicago Press.

California Language Archive. (2018) [Web page]. Retrieved Dec 10, 2018 from, http://cla.berkeley.edu/list.php

Campbell, L., & Muntzel, M. C. (1989). The structural consequences of language death. In N. Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death* (pp. 181-196). Cambridge University Press.

Chafe, W. L. (Ed.). (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Ablex.

Chafe, W. L. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 21-51). John Benjamins.

Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems.* Cambridge University Press.

Comrie, B. (1981). *Language universals and linguistic typology: Syntax and morphology.* University of Chicago Press.

Crawford, J. (1995). *Bilingual education: History, politics, theory, and practice* (3rd ed.). Bilingual Educational Services.

Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information.* University of Chicago Press.

Croft, W. (1995). Intonation units and grammatical structure. *Linguistics, 33*(5), 839-882.

Croft, W. (1998). The structure of events and the structure of language. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structures* (pp. 67-92). Lawrence Erlbaum Associates.

Croft, W. (2001). *Radical Construction Grammar: Syntactic theory in typological perspective.* Oxford University Press.

Croft, W. (2012). *Verbs: Aspect and causal structure*. Oxford University Press.

Croft, W. (2019). *Morphosyntax: Constructions of the world's languages.* Cambridge University Press. Manuscript in preparation.

Cruz, R. (2012). Am t ñe'ok et a:t o ce:ek t do'ibioda:lik [In our language is where we will find our liberation]. *Berkeley La Raza Law Journal*, *22*(1), 97-115.

Cruz, R. (2017). Why O'odham core values stop and emergent survivor O'odham vanished [Unpublished doctoral dissertation proposal]. University of New Mexico, Albuquerque.

Cyr, D. E. (1993a). Cross-linguistic Quantification: Definite articles vs. demonstratives. *Language Sciences, 15,* 195–229.

Cyr, D. E. (1993b). Definite articles and demonstratives in Plains Cree. In W. Cowan (Ed.), *Papers from the twenty-fourth Algonquian Conference* (pp. 64-80). Carleton University.

Cyr, D. E. (1996). Nikotwâsik iskwâhtêm, pâskihtêpayih!. In J. D. Nichols & A. C. Ogg (Eds.), *Studies in honor of H. C. Wolfart* (pp. 77-111). Memoir.

Dahl, Ö. (1985). *Tense and aspect systems.* Basil Blackwell.

Derbyshire, D. C. (1985). *Hixkaryana and linguistic typology.* Summer Institute of Linguistics and University of Texas at Arlington Press.

Diessel. (1999). *Demonstratives: Form, function, and grammaticalization.* John Benjamins.

Dixon, R. (1972). *The Dyirbal language of Queensland, North Australia.* Cambridge University Press.

Dixon, R. M. W. (1997). *The rise and fall of languages.* Cambridge University Press.

Dixon, R. M. W. (2003). Demonstratives: A cross-linguistic typology. *Studies in Language, 27*(1), 61-112.

Dobyns, H. F. (1950). *Papagos in the cotton fields.* University of Arizona Press.

Dobyns, H. F. (1963). Indian extinction in the middle Santa Cruz River Valley, Arizona. *New Mexico Historical Review, 38*(2), 163-181.

Dobyns, H. F. (1972). *The Papago people.* Indian Tribal Series.

Dolan, D. (1972). The Plomo papers. *Ethnohistory, 19*(4), 305-322.

Dolores, J. (1913). Papago verb stems. *University of California Publications in American Archaeology and Ethnology, 10*(4), 241-263.

Dolores, J. (1923). Papago nominal stems. *University of California Publications in American Archeology and Ethnology, 20.*

Dolores J., & Kroeber, A. (1909-1951). *Papago Language and Folklore* [Unpublished field notebooks]. (Available through Bancroft Library, University of California, Berkeley).

Dolores, J., & Mathiot, M. (1991). The reminiscences of Juan Dolores, an early O'odham linguist. *Anthropological Linguistics, 33*(3), 232-315.

Dolores J. (narrator) & Waterman, T. T. (researcher). (1909). *Narrative of the boy who became a warrior* or *Boy who avenged his father* [Wax cylinder recording]. The T. T. Waterman collection of California Indian sound recordings. (Available through the Phoebe A. Hearst Museum of Anthropology Repository, University of California, Berkeley).

Dorian, N. C. (1973). Grammatical change in a dying dialect. *Language, 49*(2), 413-438.

Dorian, N. C. (1976). A hierarchy of morphophonemic decay in Scottish Gaelic language death: The differential failure of lenition. *Word, 28*(1-2), 96-109.

Dorian, N. C. (1978). The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language*, *54*(3), 590-609.

Dorian, N. C. (1981). *The life cycle of a Scottish Gaelic dialects.* University of Pennsylvania Press.

Dorian, N. C. (1993). Internally and externally motivated change in language contact settings: Doubts about dichotomy. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 131-155). Longman.

Dressler, W. (1996). Language death. In R. Singh, *Towards a Critical Sociolinguistics* (pp. 195-210). John Benjamins.

Dyer, M. (1986). Primary objects, secondary objects and antidative. *Language*, 547-619.

Duarte, C. (2019, January 15). Stella Tucker dies; helped keep ancient O'odham art of saguaro fruit harvesting alive. *Arizona Daily Star*. https://tucson.com/news/local/stella-tucker-dies-helped-keep-ancient-o-odham-art-of/article_4aacd9e1-93a7-505e-9670-559d2e825679.html

Durie, M. (1986). The grammaticization of number as a verbal category. In V. Nikiforidou, M. VanClay, M. Niepokuj & D. Feder (Eds.), *Proceedings of the twelfth annual meeting of the Berkeley Linguistics Society* (Vol. 12, pp. 355-370).

Erickson, W. P. (1994). *Sharing the desert: The Tohoho O'odham in history.* University of Arizona Press.

Fillmore, C. M. (1997). *Lectures on deixis.* CSLI Publications.

Fitzgerald, C. M. (2002). Tohono O'odham stress in a single ranking. *Phonology*, *19*(2), 253-271.

Fitzgerald, C. M. (2003). How prosodically consistent is Tohoho O'odham? *Studies in Uto-Aztecan*, *5*, 55-74.

Fitzgerald, C. M. (2004). Language change and motion verbs in Tohoho O'odham [Conference handout]. LASSO XXXIII.

Fitzgerald, C. M. (2010). Language documentation in the Tohoho O'odham community. In L. Grenoble & N. L. Furbee (Eds.), *Language documentation: Practice and values* (pp. 9-11). John Benjamins.

Flores Farfán, J. A. (2004). Notes on Nahuatl typological change. *Sprachtypologie und Universalienforschung* (STUF) [Language Typology and Universals], *57*(1), 85-97.

Fontana, B. L. (1960). *Assimilative change: A Papago Indian case study.* [Unpublished doctoral dissertation]. University of Arizona, Tucson.

Fontana, B. L.(1974). The Papago tribe of Arizona. In D. A. Horr (Ed.), *American Indian Ethnohistory: Indians of the Southwest: Papago Indians III* (pp. 151-226).: Garland Publishing Inc.

Fontana, B. L. (1983a). History of the Papago. In A. Ortiz & W. C. Sturtevant (Eds.), *Handbook of North American Indians: Southwest* (Vol. 10, pp. 137-148). Smithsonian Institution.

Fontana, B. L. (1983b). Pima and Papago: Introduction. In W. C. Ortiz & A. Sturtevant (Eds.), *Handbook of North American Indians: Southwest* (Vol. 10, pp. 125-136). Smithsonian Institution.

Fontana, B. L. (1989). *Of earth and little rain: The Papago Indians.* University of Arizona Press.

Fox, B. A., Hayashi, M., & Jasperson, R. (1996). Resources and repair: A cross-linguistic study of syntax and repair. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 185-237). Cambridge University Press.

Frajzyngier, Z. (2011). Ergativity, number, and agreement. In *Proceedings of the annual meeting of the Berkeley linguistics society* (Vol. 11, pp. 96-106).

Franco Hernández, J. (2009, June 29-July 2). *Towards the maintenance of Sonoran O'otam: Some views on its current sociolinguistic situation* [Poster]. The 30th Anniversary Symposium Celebration of American Indian Language Development Institute (AILDI), Tucson, Arizona.

Franco Hernández, J., & Estrada Fernández, Z. (2016, April 27-29). *Clíticos de segunda posición en O'otam (Papago)* [Conference Handout]. XI Coloquio de Lingüística Escuela Nacional de Antropología e Historia (ENAH), Mexico City, Mexico.

Fuchs, A. (1993). *Remarks on deixis.* Groos.

Gaillard, D. D. (1894). The Papago of Arizona and Sonora. *American Anthropologist*, *7*(3), 293-296.

Ganenkov, D., Lander, Y., & Maisak, T. (2010). From interrogatives to placeholders in Udi and Gaul spontaneous narratives. In N. Amiridze, B. H. Davis & M. Maclagan (Eds.) *Fillers, pauses and placeholders* (pp. 96-118). John Benjamins.

Garcia, B. (n.d.). Interview [Audio recording]. On file with compiler R. Cruz, 2012.

Givón, T. (1983). *Topic continuity in discourse: An introduction.* John Benjamins.

Givón, T. (1990). *Syntax: A functional-typological introduction* (Vol. 2). John Benjamins.

Givón, T. (1995). *Functionalism and grammar.* John Benjamins.

Givón, T. (2001). *Syntax: An introduction* (Vol. 2). John Benjamins.

Gladwin, H. W. (1957). *A history of the ancient Southwest.* Bond Wheelwright .

Goddard, C. (1983). *A grammar of Yankunytjatjara.* Institute for Aboriginal Development.

Grinevald, C. G. (1998). Language contact and language degeneration. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 257-270). Blackwell.

Gregores, E., & Suárez, J. A. (1967). *A description of colloquial Guaraní.* Mouton.

Guillaume, A. (2016). Associated motion in South America: Typological and areal perspectives. *Linguistic Typology 20*(1), 81-177.

Gutelle, S. (2014, November 6). YouTube Millionaires: Mini Ladd enjoys 'moment in the spotlight. *Tubefilter*. https://www.tubefilter.com/2014/11/06/youtube-millionaires-mini-ladd-enjoys-moment-in-the-spotlight/

Guzman Betancourt, I. (1995). The linguistic policy of the company of Jesus in new Spain. *Revista Latina De Pensamiento y Lenguaje, 2*(2B), 381-390.

Hagège, C. (1993). *The language builder.* John Benjamins.

Hagan, M. W. (1959). *An educational history of the Pima and Papago peoples from the mid-seventeenth century to the mid-twentieth century.* [Unpublished doctoral dissertation]. University of Arizona.

Hale, K. (1959). *A Papago grammar* [Unpublished doctoral dissertation]. Indiana University.

Hale, K. (1983). Papago (k)c. *International Journal of American Linguistics*, *49*(3), 299-327.

Hale, K. (2002). On the position of the auxiliary in O'odham. *Anuario Del Seminario De Filología Vasca, 44*, 303-309.

Halliday, M. A., & Hasan, R. (1976a). *Cohesion in English.* Longman.

Hanks, W. F. (1990). *Referential practice.* University of Chicago Press.

Haugen, J. D. (2007). Denominal verbs and noun incorporation: Uto-Aztecan evidence for a unified syntactic account. In F. Hoyt, N. Seifert, A. Teodorescu, & J. White (Eds.), *Texas linguistics society 9: The morphosyntax of underrepresented languages* (pp. 145-165). CSLI Publications.

Haugen, J. D. (2008a). *Morphology at the interfaces: Reduplication and noun incorporation in Uto-Aztecan*. John Benjamins.

Haugen, J. D. (2008b). *Linguistik Aktuell/Linguistics Today : Morphology at the interfaces: Reduplication and noun incorporation in Uto-Aztecan* (Vol. 117). John Benjamins.

Hayashi, M. & Yoon, K. (2006). A cross-linguistic exploration of demonstratives in interaction: With particular reference to the context of word-formulation trouble. *Studies in Language, 30*, 485–540.

Hayashi, M., & Yoon, K. (2010). A cross-linguistic exploration of demonstratives in interaction with particular reference to the context of word-formation trouble. In M. Maclagan, B. H. Davis, & N. Amiridze (Eds.), *Fillers, pauses and placeholders* (pp. 33-66). John Benjamins.

Heine, B. (2003). Grammaticalization. In B. Joseph & R. Janda (Eds.), *The handbook of historical linguistics* (pp. 575-601). Blackwell Publishing.

Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization.* Cambridge University Press.

Heine, B., & Kuteva, T. (2003). On contact-induced grammaticalization. *Studies in Language*, *27*(3), 529-572.

Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change.* Cambridge University Press.

Heine, B., & Kuteva, T. (2010). Contact and grammaticalization. In R. Hickey (Ed.), *The handbook of language contact* (pp. 86-105). Blackwell Publishing.

Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework.* University of Chicago Press.

Heine, B., Song, K. (2011). On the grammaticalization of personal pronouns. *Journal of Linguistics, 47,* 587–630. doi:10.1017/S0022226711000016

Herrington, F. C., Woodward, E. S., & Alsberg, H. G. (1939). The Papago. *Arizona State Teachers College Bulletin, Flagstaff* (Vol. 20). Arizona State Teachers College. http://hdl.handle.net/2027/uc1.$b59456

Hickey, & Conde-Silvestre, J. C. (2012). Internally- and externally-motivated language change. In J. M. Hernández-Campo (Ed.), *The handbook of historical sociolinguistics* (pp. 387-407). Blackwell Publishing.

Hill, J. H. (1978). Language death, language contact and language evolution. In W. C. McCormack (Ed.), *Approaches to language: Anthropological issues* (pp. 45-75). Mouton de Gruyter.

Hill, J. H. (1983). Language death in Uto-Aztecan. *International Journal of American Linguistics*, *49*(3), 258-276.

Hill, J. H., & Zepeda, O. (1992). Derived words in Tohoho O'odham. *International Journal of American Linguistics*, *58*(4), 355-404.

Hill, J. H., & Zepeda, O. (1998). Tohono O'odham (Papago) plurals. *Anthropological Linguistics*, *40*(1), 1-42.

Himmelmann, N. P. (1996). Demonstratives in narrative discourse: A taxonomy of universal use. In B. A. Fox (Ed.), *Studies in anaphora* (pp. 205-254). John Benjamins.

Holisky, D. A. (1987). The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua*, *71*, 103-132.

Hopper, P. J., & Traugott, E. C. (1993). *Grammaticalization.* Cambridge University Press.

Hoy, B. (1994). War in Papaguería: Manuel Gándara's 1840-41 Papago expedition. *The Journal of Arizona History, 35*(2), 141-162.

Institute for Government Research. (1928). *The problem of Indian administration (Meriam report).* Johns Hopkins Press.

Jackson, R. H. (1981). The last Jesuit censuses of the Pimería Alta missions, 1761 and 1766. *KIVA, 46*(4), 243-272. doi:10.1080/00231940.1981.11757959

Joseph, A., Spicer, R. B., & Chesky, J. (1949). *The desert people: A study of the Papago Indians.* The University of Chicago Press.

Kelly, W. H. (1963). *The Papago Indians of Arizona: A population and economic study.* Bureau of Ethnic Research, Department of Anthropology, University of Arizona.

Kelly, W. H. (1974). The Papago Indians of Arizona. In D. A. Horr (Ed.), *American Indian Ethnohistory: Indians of the Southwest: Papago Indians III* (pp. 9-150). Garland Publishing Inc.

Kilpatrick, K. (2014, May 25). U.S.-Mexico border wreaks havoc on lives of an indigenous desert tribe. *Aljazeera America*. http://america.aljazeera.com/articles/2014/5/25/us-mexico-borderwreakshavocwithlivesofanindigenousdesertpeople.html

Klatt, D. H., (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics, 3,* 131-142.

Kroeber, A. L. (1949). Juan Dolores, 1880-1948. *American Anthropologist, 51*(1), 96-97.

Kroeber, A. L. (1950). Preface. In J. A. Mason (Ed.), *Language of the Papago of Arizona.* University Museum, University of Pennsylvania.

Kuryłowicz, J. (1975). The evolution of grammatical categories. In *Esquisses linguistiques* (Vol. 2, pp. 38-54). München: Fink. (Original work published 1965).

Labov, W. (1972). *Sociolinguistic patterns.* University of Pennsylvania Press.

Langacker, R. W. (1977). An overview of Uto-Aztecan grammar: Studies in Uto-Aztecan grammar 1. In *Studies in Uto-Aztecan grammar: An overview of Uto-Aztecan grammar* (Vol. 1). Summer Institute of Linguistics & University of Texas at Arlington.

Laury, R. (1997). *Demonstratives in interaction: The emergence of a definite article in Finnish.* John Benjamins.

Lehmann, C. (1985). Grammaticalization: Synchronic variation and diachronic change. *Lingua E Stile, 20*(3), 303-318.

Lehmann, C. (1992). Word order change by grammaticalization. In M. Gerritsen & D. Stein (Eds.), *Internal and external factors in syntactic change* (pp. 395-416). Mouton de Gruyter.

Lehmann, C. (2002). *Thoughts on grammaticalization* (2nd, revised ed.). Seminar für Sprachwissenschaft der Universität. https://www.christianlehmann.eu/publ/ASSidUE09.pdf.

Levinson, S. C. (1983). *Pragmatics.* Cambridge University Press.

Lichtenberk, F. (1988). The pragmatic nature of nominal anaphora in To'aba'ita. *Studies in Language, 12*, 229-344.

Lichtenberk, F. (1996). Patterns of anaphora in To'aba'ita narrative discourse. In B. A. Fox (Ed.), *Studies in anaphora* (pp. 379-411). John Benjamins.

Lindsey. (2000). Subordinate clauses and ergative patterns in Shoshoni . In V. P. de Guzman & B. Bender (Eds.), *Oceanic Linguistics Special Publication: Grammatical analysis: Morphology, syntax, and semantics: Studies in honor of Stanley Starosta* (pp. 274-292). University of Hawai'i Press.

Lopez, D. (1998). The parents have to do their part: Tohono O'odham language autobiography. *International Journal of the Sociology of Language*, *132*(1), 43-45.

Madsen, K. (2004). *Tohono O'odham Nation comprehensive education study.* Tohono O'odham Nation and Tohono O'odham Community College.

Madsen, K. D. (2005). *A nation across nations: The Tohoho O'odham and the United States-Mexico border* [Unpublished doctoral dissertation]. Arizona State University.

Marak, A. M., & Tuennerman, L. (2013). *At the border of empires: The Tohoho O'odham, gender, and assimilation, 1880-1934.* University of Arizona Press.

Mason, J. A. (1950). *The language of the Papago of Arizona.* University Museum, University of Pennsylvania.

Mathiot, M. (n.d.a). *Tohono 'O'odham-English dictionary (Vol. 1)* [PDF]. (Revised version of *The language of the Papago of Arizona, Vol. 1* published in 1973). http://www.acsu.buffalo.edu/~mathiotm/Mathiot/Volume%20I.pdf.

Mathiot, M. (n.d.b). *Tohono 'O'odham-English dictionary (Vol. 2)* [PDF]. (Revised version of *The language of the Papago of Arizona, Vol. 2* published in 1973). http://www.acsu.buffalo.edu/~mathiotm/Mathiot/Volume%20II.pdf

Mathiot, M. (1973). *A dictionary of Papago usage*. Indiana University Publications.

McCarthy, J. (1985). *A Papago traveler: The memories of James McCarthy.* Sun Tracks and the University of Arizona Press.

McCarty, T. L., & Watahomigie, L. J. (1998). Indigenous community-based language education in the USA. *Language, Culture and Curriculum, 11*(3), 309-324.

McCarty, T. L., Romero-Little, M. E., Warhol, L., & Zepeda, O. (2009). Indigenous youth as language policy makers. *Journal of Language, Identity, and Education, 8*, 291-306. doi:10.1080/15348450903305098

Meillet, A. (1912). L'evolution des formes grammaticales [The evolution of grammatical forms]. *Scientia (Rivista Di Scienzia) 26*(6), 284-400.

Mithun, M. (1989). The incipient obsolescence of polysynthesis: Cayuga in Ontario and Oklahoma. In Dorian (Ed.), *Investigating obsolescence: Studies in language contraction and death* (pp. 243-258). Cambridge University Press.

Mithun, M. (1990). Language obsolescence and grammatical description. *International Journal of American Linguistics, 56*(1), 1-26.

Mithun, M. (1991). Active/agentive case marking and its motivations. *Language*, *67*, 510-546.

Moseley, C. (2010). *The UNESCO atlas of the world's languages in danger* (3rd ed.) [Web page]. UNESCO Publishing. http://www.unesco.org/culture/en/endangeredlanguages/atlas

Nagaraja, K. S. (1985). *Khasi: A descriptive analysis.* Deccan College.

Nettle, D., & Romaine, S. (2000). *Vanishing voices: The extinction of the world's languages.* Oxford University Press.

Nichols, J., & Balthasar, B. (2013). Locus of marking in the clause. *The world atlas of language structures online*. Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/23

Niswander, J. D., Brown, K. S., Iba, B. Y., Leyshon, W. C., & Workman, P. L. (1970). Population studies on southwestern Indian tribes. I: History, culture, and genetics of the Papago. *American Journal of Human Genetics*, *22*(1), 7.

Ortega, M. (Creator & Compiler). (2013-present). *Tohonotv* [YouTube channel]. YouTube. https://www.youtube.com/channel/UCTb-pgRhgAmBEqN3Ty8l85Q

Pancho, J. & Mathiot, M. (1959-1960). *Coyote Stories.* https://www.acsu.buffalo.edu/~mathiotm/text.html

Payne, D. (1987). Information structuring in Papago narrative discourse. *Language*, *63*(4), 783-804.

Payne, D. (1992). Nonidentifiable information and pragmatic rules in O'odham. In D. Payne (Ed.), *Pragmatics of word order flexibility* (pp. 137-166). John Benjamins.

Payne, D. (2013). Alignment and Grammatical Relations. In S. Luraghi & C. Parodi, *The Bloomsbury Companion to Syntax* (218-239). Bloomsbury.

Perlman, A. (1969). "This" as a third article in American English. *American Speech*, *44*(1), 76-80.

Podlesskaya, V. I. (2010). Parameters for typological variation of placeholders. In M. Maclagan, B. H. Davis, & N. Amiridze (Eds.), *Fillers, pauses and placeholders* (pp. 11-32). John Benjamins.

Podlesskaya, V. I. & Kibrik, A. A. (2006). Korrekcija v ustnoj monologičeskoj reči po dannym korpusnogo issledovanija (A corpus study of repairs in spoken monologues). *Russkij jazyk v naučnom osveščenii 2*(12), 7–55.

Podlesskaya, V. I. & Kibrik, A. A. (2009). Rečevye sboi i zatrudnenija (Speech disfluencies). In A. A. Kibrik & V. I. Podlesskaja (Eds.), *Rasskazy o snovidenijax: Korpusnoe issledovanie ustnogo russkogo diskursa*, 177–218. Languages of Slavonic Culture.

Pratt, R. H. (1892). The advantages of mingling Indians with whites. In *Official report of the nineteenth annual conference of charities correction* (pp. 46-59). National Conference of Charities and Correction (U.S.) Session.

Prince, E. (1992). The ZPG letter: Subjects, definiteness and information-status. In S. A. Thompson & W. Mann (Eds.), *Discourse description: Diverse analyses of a fundraising text* (pp. 295-325). John Benjamins.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A grammar of contemporary English.* Longman.

Reeves, A. G. (2006). *To us they are butterflies: A case study of the educational experience at an urban indigenous-serving charter school* [Unpublished doctoral dissertation]. University of Arizona.

Reff, D. T. (2008). Sympathy for the devil: Devil sickness and lore among the Tohoho O'odham. *Journal of the Southwest*, *50*(4), 355-376.

Romaine, S. (2010). Contact and language death. In R. Hickey (Ed.), *The handbook of language contact* (pp. 320-339). Wiley-Blackwell.

Ruggles, L. (1867). *Letter to Arizona superintendent of Indian affairs George Dent* [Letter].

Sasse, H. J. (1992). Language decay and contact-induced change: Similarities and differences. In Brenzinger (Ed.). *Language death: Factual and theoretical explanations with special reference to East Africa (Contributions to the sociology of language)* (Vol. 64, pp. 59-80). Mouton de Gruyter.

Saxton, D. (1982). Papago. In R. W. Langacker (Ed.), *Studies in Uto-Aztecan grammar: Uto-Aztecan grammatical sketches* (Vol. 3, pp. 92-266). The Summer Institute of Linguistics & The University of Texas at Arlington.

Saxton, D., Saxton, L. & Enos, S. (1983). *Dictionary: Tohono O'odham/Pima to English-English to Tohono O'odham/Pima.* The University of Arizona Press.

Saxton, D., & Saxton, L. (1973). *Legends and lore of the Papago and Pima Indians.* The University of Arizona Press.

Schaengold, C. C. (2004). *Bilingual Navajo: Mixed codes, bilingualism, and language maintenance* [Unpublished doctoral manuscript]. Ohio State University.

Schlesinger, I. M. (1995). *Cognitive space and linguistic space.* Cambridge University Press.

Schmidt, A. (1985). Young people's Dyirbal: An example of language death from Australia. In *Young people's Dyirbal: An example of language death from Australia* (pp. 1-252). Cambridge University Press.

Segundo, T., Ladd, B., Krug, J. C., & Nichols, J. H. (1948). *U.S., Department of the Interior, Bureau of Indian Affairs, U.S: The Papago development program* [Pamphlet]. Haskell Institute Printing Shop.

Shaul, L. (2012). *Aspect and tense marking in Sonoran O'otam and the development of the category AUX in Tohoho O'odham* [handout]. http://www.wyominglinguistics.org/2012/09/06/aspect-and-tense-marking-in-sonoran-ootam-2/

Smiljanić, R., & Bradlow, A. R. (2005). Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America, 118*(3), 1677-1688. https://doi.org/10.1121/1.2000788

Smith, D. G. (1981). Admixture and population replacement of the Sells Papago Indians: Three strategies. *Social Biology, 28*(1-2), 126-144.

Steele, S. (2008). Situating morphology. *International Journal of American Linguistics, 74*(3), 313-364.

Stuart, P. (1977). United States Indian policy: From the Dawes Act to the American Indian policy review commission. *The Social Service Review, 51*(3), 451-463.

Summer Institute of Linguistics. (2017). *FieldWorks language explorer (FLEx)* [Computer Software]. Summer Institute of Linguistics.

Teague, L. S. (2009). Prehistory and the traditions of the O'odham and Hopi. *Kiva, 75*(2), 239-259. doi:10.1179/kiv.2009.75.2.011

Thomason, S. G. (2001). *Language contact.* Georgetown University Press.

Thomason, S. G. (2010). Contact explanations in linguistics. In R. Hickey (Ed.), *The handbook of language contact* (pp. 319-328). Blackwell.

Tohono O'odham Kekel Ha-Maṣcamakuḍ [Tohono O'odham Community College]. (2018, May 2). *Himdag at TOCC.* Tohono O'odham Kekel Ha-Maṣcamakuḍ. https://www.tocc.edu/himdag/

Tohono O'odham Department of Education & O'odham *Ne'ok* Revitalization Project. (2013). *O'odham Ñe'ok survey report* [Unpublished manuscript]. Tohono O'odham Department of Education.

Tohono O'odham Nation. (2016, November 10). *Location.* Official website of the Tohono O'odham Nation. http://www.tonation-nsn.gov/location/

Tohono O'odham Nation. (2017a, December 17). *Home.* Official website of the Tohono O'odham Nation. http://www.tonation-nsn.gov/

Tohono O'odham Nation. (2017b, December 17). *Tohono O'odham history.* Official website of the Tohono O'odham Nation. http://www.tonation-nsn.gov/tohono-oodham-history/

Tohono O'odham Nation. (2018, January 17). *Tohono O'odham Community College.* Official website of the Tohono O'odham Nation. http://www.tonation-nsn.gov/tohono-oodham-community-college/

Traugott, E. (2003). Constructions in grammaticalization. In R. D. Janda & B. D. Joseph (Eds.), *Handbook of historical linguistics* (Vol. , pp. 624-647). Blackwell Publishers.

Traugott, E. C., & König, E. (1991). The semantics-pragmatics of grammaticalization revisited. In E. C. Traugott & B. Heine (Eds.), *Approaches to grammaticalization* (Vol. 1, pp. 189-218). John Benjamins.

Trennert, Jr., R. (1988). *The phoenix Indian school: Forced assimilation in Arizona, 1981-1935.* University of Oklahoma Press.

Trudgill, P. (2010). Contact and sociolinguistic typology. In R. Hickey (Ed.), *The handbook of language contact* (pp. 299-319). Wiley-Blackwell.

Tsunoda, T. (2005). *Trends in linguistics: Studies and monographs: Language endangerment and language revitalization: An introduction* (Vol. 148). Mouton de Gruyter.

Tucker, C., Fitzgerald, C., & Miguel. (2012). Contemporary storytelling in Tohoho O'odham. In D. Kozak (Ed.), *Inside dazzling mountains: Contemporary translations of southwest native verbal arts* (pp. 391-406). University of Nebraska Press.

Underhill, R. M. (1939). *Social organization of the Papago Indians.* AMS Press.

Underhill, R. M. (1978). *The Papago Indians of Arizona and their relatives the Pimas.* AMS Press. (Original work published 1940).

Underhill, R. M. (1985). *Papago woman.* Waveland Press.

United States Office of Indian Affairs. (1887). *Annual report of the Commissioner of Indian Affairs.* Government Printing Office.

United States Office of Indian Affairs. (1890). *Annual report of the Commissioner of Indian Affairs.* (1890). Government Printing Office.

United States Office of Indian Affairs. (1891). *Annual report of the Commissioner of Indian Affairs.* Government Printing Office.

United States Office of Indian Affairs. (1905). *Annual report of the department of the interior for the fiscal year ended June 30, 1905, part 1*. Government Printing Office. http://digital.library.wisc.edu/1711.dl/History.AnnRep05p1

Wagoner, J. J. (1952). *The history of the cattle industry in southern Arizona, 1540-1940* [Unpublished master's thesis]. University of Arizona.

Wald, B. (1983). Referents and topics within and across discourse units: Observations from current vernacular English. In F. Klein-Andreu (Ed.), *Discourse perspectives on syntax* (pp. 91-116). Academic Press.

Wilkins, D. P. (1989). *Mparntwe Arrernte (Aranda): Studies in the structure and semantics of grammar* [Unpublished doctoral dissertation]. Australian National University.

Winford, D. (2003). *An introduction to contact linguistics.* Blackwell.

Wright, S. E., & Givón, T. (1987). The pragmatics of indefinite reference: Quantified text-based studies. *Studies in Language*, *11*, 1-33.

Zepeda, O. (1983). *A Papago grammar.* University of Arizona Press.

Zepeda, O. (1984). *Topics in Papago morphology* [Unpublished doctoral dissertation]. University of Arizona.

Zepeda, O. (1990). American Indian language policy. In K. L. Adams & D. T. Brink (Eds.), *Perspectives on official English: The campaign for English as the official language of the USA* (pp. 247-256). Mouton de Gruyter.

Zepeda, O. (1998). Voices in the desert: Contemporary approaches to language maintenance and survival of an ancient language, Tohoho O'odham. *International Journal of the Sociology of Language*, *132*(1), 47-58. doi:10.1515/ijsl.1998.132.47

Zipf, G. K. (1935). *The psych-biology of language.* Houghton Mifflin.