

TUNICA AT DUSK AND DAWN:
LANGUAGE CHANGE IN OBSOLESCENCE AND REVIVAL

AN ABSTRACT

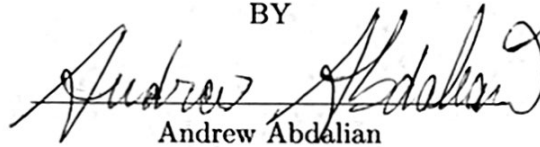
DEFENDED ON THE FOURTH DAY OF MAY 2023
TO THE INTERDISCIPLINARY PROGRAM IN LINGUISTICS
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE SCHOOL OF LIBERAL ARTS OF TULANE UNIVERSITY

FOR THE DEGREE

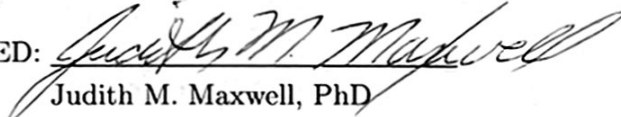
OF

DOCTOR OF PHILOSOPHY

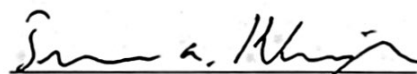
BY



Andrew Abdalian

APPROVED:


Judith M. Maxwell, PhD

Director


Thomas A. Klingler, PhD


Marc Zender, PhD

Abstract

The Tunica language is a language isolate whose last known native speaker died in 1948. Three separate linguists, working with three different Tunica speakers, documented the language at three different time periods from 1886 to 1939. This dissertation's first goal is to quantify language change across these three periods of time, as well as in 2017, during Tunica's revival effort. To accomplish this, a Tunica language corpus was created from Tunica language documentation to examine language change during the period of Tunica language obsolescence and revival. The dissertation uses frequency analysis to analyze potential loci of language change, including the verb aspectual system, nominal gender marking, use of evidentials, syncope in possessive noun prefixes, and word order. The differences between the subcorpora of documented historical Tunica yielded no conclusive evidence of gradual language change over time, revealing instead that the three Tunica speakers varied widely in their use of the language, but in no clear direction. Comparing these three subcorpora to the Modern Tunica subcorpus reveals certain language changes that would be expected in both obsolescence and early revival, including a higher use of less complex and more analytical verb forms and much lower use of marked structures. This dissertation's second goal is to provide information and tools that will enable greater access to and easier analysis of all Tunica language documentation. Tunica is a heritage language of the Tunica-Biloxi Tribe and has been the subject of community-based revival efforts since the 1970s. In 2010, the tribe partnered with Tulane University and created a more formal language revival organization based on principles of community-engaged scholarship. Language revival is a herculean task requiring community buy-in and the marshaling of as many resources as possible. In any language revitalization effort, easily accessible documentation is valuable for language workers, language learners, and the wider community. The work that underlies this dissertation has resulted in the creation of a more complete corpus of parsed Tunica texts, as well as methods and procedures that make it easier to investigate questions about the language.

TUNICA AT DUSK AND DAWN:
LANGUAGE CHANGE IN OBSOLESCENCE AND REVIVAL

A DISSERTATION

DEFENDED ON THE FOURTH DAY OF MAY 2023

TO THE INTERDISCIPLINARY PROGRAM IN LINGUISTICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

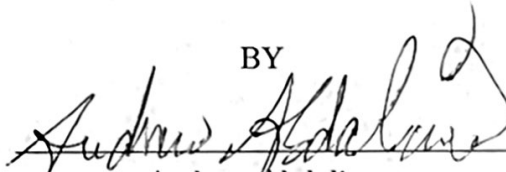
OF THE SCHOOL OF LIBERAL ARTS OF TULANE UNIVERSITY

FOR THE DEGREE

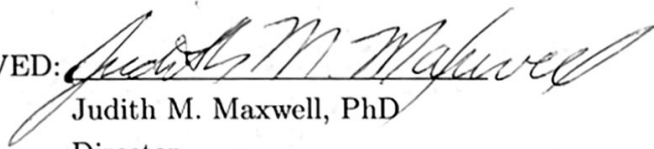
OF

DOCTOR OF PHILOSOPHY

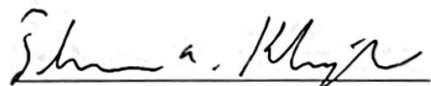
BY

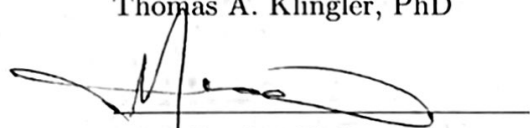

Andrew Abdalian

APPROVED:


Judith M. Maxwell, PhD

Director


Thomas A. Klingler, PhD


Marc Zender, PhD

© Copyright by Andrew Abdalian, 2023
All Rights Reserved

ACKNOWLEDGMENTS

First and foremost, I would like to thank the members of the Tunica-Biloxi Language & Culture Revitalization Program past and present: Donna Pierite, Elisabeth Pierite, John Barbry, Brenda Lintinger, Teyanna Pierite-Simon, Ryan Lopez, Sidney Barbry, Juston Broussard, and Summer Leleux. Your questions, counsel, and friendship have been invaluable to my understanding of Tunica. *Tikahch panu.*

Thanks to my advisor and director of my committee, Dr. Judith Maxwell, who introduced me to the Tunica Language Project and provided constant advice, feedback, and encouragement throughout my writing process. Your wisdom and humor made the long road of dissertation-writing more pleasant. To my other committee members, Dr. Tom Klingler and Dr. Marc Zender for providing valuable feedback on my work that improved the final product.

Members of the Tunica Language Working Group—*Kuhpani Yoyani Luhchi Yoroni* or KYLY—provided a network of scholars with whom I could discuss the finer points of Tunica. Thanks to Nathan Wendte, Tom Lewis, Craig Alcantara, Raina Heaton, Patricia Anderson, Mackenzie Walters, Lisa Sprowls, and Brett C. Nelson.

Special thanks to the fellowship of KYLY members who have worked as LCRP linguists: Tyler Whitaker, Meg Harvey, and Rebecca Moore.

The Tulanians that shared this dissertation-writing journey with me, Hawazen Al-Rowais and Jessica Glass, provided commiseration and support and kept me accountable.

Megan Mills and Robert Abdalian served as my intrepid proofreaders, discarding my extraneous commas with great aplomb. Any errors that remain are my fault alone.

Writing a dissertation is by its nature a solitary and lonely endeavor. Writing a dissertation during a pandemic compounds the issue. I am extremely fortunate to have had family nearby and friends both near and far to help me through. These are perhaps the people most responsible for me completing my dissertation.

My parents Robert and Sue Ellen, my sister Zarouhie and her partner Drew Rosenzweig were immensely helpful, offering to help with whatever I needed.

Craig Alcantara provided not only valuable ideas for my dissertation, but valuable friendship as well. Alice Noel provided heroic amounts of support throughout the dissertation-writing process, for which I am immensely grateful.

Christina Temes provided advice from afar from the perspective of someone who has been through the grad school crucible and has come out the other side. And Megan Mills, who was half a world away, provided abundant wit, humor, and encouragement in addition to her proofreading prowess.

Table of Contents

List of Figures	vii
List of Tables	viii
1 Introduction	1
1.1 Research questions	4
1.2 Structure of the dissertation	5
2 Tunica language documentation	7
2.1 The Tunica people	7
2.2 Tunica language classification	12
2.3 Analysis of early documentation	12
2.3.1 The documenters	15
2.3.2 The speakers	22
2.4 Conclusion	26
3 Language obsolescence	28
3.1 Structural consequences of language obsolescence	31
3.1.1 Common structural changes in obsolescence	32
3.2 Rate of change in obsolescence	34
3.3 Conclusion	35
4 Tunica revival	37
4.1 <i>Kuhpani Yoyani Luhchi Yoroni</i> (KYLY): The Tunica Language Working Group	37

4.1.1	The Language & Culture Revitalization Program (LCRP) . . .	38
4.1.2	Organizing a community-engaged scholarship (CES) project .	40
4.1.3	KYLY goals	42
4.1.4	KYLY's revitalization philosophy	43
4.2	Competing ideologies in language revival	46
4.2.1	Continued reanalysis of documented Tunica	49
4.3	Conclusion	50
5	Methodology	52
5.1	Corpus creation	52
5.1.1	Documented Historical Tunica	52
5.1.2	Modern Tunica texts	55
5.1.3	The corpus	57
5.1.4	Digitizing the corpus	57
5.2	Automating frequency analysis	65
6	Areas of study	66
6.1	Phenomena noted by documenters and speakers	67
6.2	Verb aspect	70
6.2.1	Auxiliary verbs	71
6.2.2	Completive aspect	76
6.2.3	Habitual aspect	77
6.2.4	Stative verbs	79
6.3	Gender	80
6.4	Evidential = <i>ani</i>	85
6.5	Elision in possessive prefixes	87
6.6	Word order	91
6.7	Conclusion	92

7 Data	94
7.1 Historical phenomena noted by documenters and speakers	97
7.1.1 Swanton’s incorporated suffixes	97
7.1.2 Haas and Sesostrie Youchigant’s gender-number affix (GNA) analysis	99
7.1.3 Reanalysis of verbs of nonvolitional action	101
7.2 Verb inflection	103
7.3 Gender	105
7.3.1 Leveling of gender in the plural	105
7.3.2 Variable gender	107
7.3.3 Frequency of gender-number affixes over time	110
7.4 Evidential = <i>ani</i>	113
7.5 Elision in alienable possession	118
7.6 Word order	120
7.7 Conclusion	121
7.7.1 Future study	123
Appendix A Python script	125
Appendix B R script	130
Abbreviations	148
Glossing abbreviations	149
Works Cited	150

List of Figures

- 5.1 An interlinear sentence in Fieldworks Language Explorer (FLEX) . . . 59

List of Tables

4.1	Development of Tunica orthography from Haas to Modern Tunica . .	38
4.2	Tunica vowel blending chart	44
5.1	Sources of Historical Tunica	53
5.2	All transcribed Historical Tunica texts	54
5.3	Sources of Modern Tunica texts	56
5.4	Corpus texts	57
5.5	Transcription differences	64
6.1	Tunica auxiliary verbs	70
6.2	Class I completive	76
6.3	Class II completive	76
6.4	Class I habitual	77
6.5	Class II habitual	77
6.6	Tunica gender-number affixes	84
6.7	Tunica possessive prefixes	89
6.8	Haas (1946) word order chart	91
7.1	Corpus information	94
7.2	Tunica text categories in Haas (1950)	95
7.3	Corpus by text genre	96
7.4	Last consonant in verbal and nominal stems from Swanton (1921) . .	98
7.5	Last consonants in the current Tunica lexicon	99
7.6	Forms of third-person singular gender-number affixes (GNAs)	99

7.7	Verb inflection percentages by era	103
7.8	Gender and number distribution	105
7.9	Noun-marked gender by person, number, and subcorpus	106
7.10	Gender-marked plural nouns by animacy, gender, and subcorpus	106
7.11	Prefixes noun gender marking	110
7.12	Percentage of non-gender-marked prefixed nouns by genre and era	111
7.13	Frequency of gender marking of prefixed nouns in tellings of the Orphan Myth	112
7.14	Frequency of gender marking of prefixed nouns in tellings of the Flood Myth	112
7.15	Frequency of = <i>ani</i> by era	114
7.16	Frequency of = <i>ani</i> per text from Volsin Chiki	116
7.17	Percentage of sentences using = <i>ani</i> by genre and era	117
7.18	Elision of alienable possessors	119
7.19	Verb location by era	120

Chapter 1: Introduction

In January 2016, I traveled to Marksville, Louisiana, for a week-long Tunica language immersion program organized by the Tunica-Biloxi Tribe of Louisiana in collaboration with Tulane University’s linguistics program. Though I grew up in New Orleans, less than a three-hour drive from Marksville, I had never heard of the town, the tribe that lives there, or their heritage languages. What I saw that winter was a community-directed initiative to revitalize Tunica—an indigenous language with no current native speakers—aided by volunteers from an academic institution, with the long-term community goal of reintroducing the language as a language of everyday communication between tribal members. The possibilities of this project—and its challenges—interested me greatly.

I became more involved with the revitalization effort, attending the yearly immersion workshop and summer camp and serving as the tribe’s on-site linguist from 2017 to 2018. As I engaged with both the language and the community, I noticed the community’s keen interest in earlier documentation of Tunica obtained from native speakers. One of the aims of the revitalization program is to make the language documentation more accessible, and I make that a goal in all of the work I do with Tunica. Though the project has benefited immeasurably from the documentation work of Tunica speakers working with three well-regarded linguists over a period from 1886 to 1939, this work was not written for a present-day lay audience. Most

of the published work is technical and contains specialized terms, some of which are no longer commonly used in linguistics, or whose main linguistic uses now refer to different phenomena.

This dissertation examines language obsolescence in Tunica through an analysis of the native speaker documentation. It also analyzes language change during the revival of Tunica. The work done for this dissertation has also made earlier documentation more accessible to members of the Tunica-Biloxi Tribe and other interested scholars through improved digitization and searchability, as well as by developing procedures for analyzing the documentation. With more accessible documentary materials, members of the revival project will be able to examine the language as spoken by native speakers more quickly, and the community at large will be able to engage with the documentation more easily. By highlighting language variation and change, the dissertation shows that language change has always been a reality of Tunica, as it is for all languages; this knowledge may help ease concerns about the authenticity of Modern Tunica¹ and allows language workers to be more confident in the decision-making process that they use to expand the language.

The study of the documented linguistic history of the Tunica language also provides a valuable case study on language obsolescence and revival because of its distinctiveness. First, Tunica is a language isolate from a region, the American Southeast, that has not been the subject of much research in the areas of obso-

1. This dissertation uses the term Modern Tunica to refer to Tunica as spoken after revival.

lescence or revival. Language isolates present a unique challenge; when looking at obsolescence, there are no related languages to help determine the changes in the obsolescent language as compared to a related, non-obsolescing language.

Revival of a language isolate is even more challenging, as language workers cannot rely on information from related languages to fill in lexical gaps or to corroborate theories of grammatical analysis. Analysis of an isolate provides the only possible information for an entire linguistic family. Examples of language revival are rare, and academic analysis of that revival is rarer still. Those instances of revival that have been studied are part of larger language families (e.g., Manx, Myaamia, Wôpanâak, Hebrew), and the most successful instance of language revival, Hebrew, was in continual use for liturgical purposes.² Few studies have examined the revival of language isolates. Looking at the revival of an isolate provides a unique example of a revival undertaken without the benefit of any documentation of related languages.

Second, Tunica's language revival initiative is ongoing and long-term. Tribal members' personal efforts to revive Tunica date back to the 1960s, when Chief Joe Pierite contacted linguist Mary Haas to request copies of her work because "[w]e would like the younger ones of our tribe to learn the language and songs" (Pierite 1964). In 2010, the Tunica-Biloxi tribe and Tulane University formed a partnership, Kuhpani Yoyani Luhchi Yoroni (KYLY) (Tunica Language Working Group), to formalize language revival efforts, and the revival effort achieved departmental

2. Hebrew revival also had the funding and organization of a nation-state, and tremendous cultural cachet.

status within the tribe's governmental structure in 2014 with the creation of the Tunica-Biloxi Language & Culture Revitalization Program (LCRP). This sustained revival effort and the resulting revived language are both worthy of study.

Third, Tunica is relatively well-documented for an indigenous language whose last known native speaker died in 1948. Three linguists documented Tunica over three distinct timespans when Tunica still had native speakers. The documentation is not exhaustive, and the different linguists worked with different understandings of linguistic theory based on the time at which they examined the language. Despite these shortcomings, the documentation at different times allows for a diachronic analysis of Tunica as it became less frequently spoken between 1886 and 1939. To have documentation of the obsolescence of language during that timespan—particularly an indigenous North American language—is unusual.

This dissertation seeks to contribute both to the community revival effort and to the linguistic understanding of language obsolescence and revival by providing an example with unique characteristics and detailed documentation.

1.1 Research questions

I undertook this study to contribute to the linguistic community and to the Tunica language revitalization effort. Though my research questions respond to the former goal, the byproducts of my research include more consistent and community-accessible language documentation, which furthers the latter goal.

This study will address four questions:

1. How much did Tunica change during its documented obsolescence, and how much did it change in revival?
2. Which aspects of the language (phonetics, phonology, morphology, syntax, prosody) changed the most, and which changed the least?
3. How well do these changes comport with existing theories of language obsolescence and language revival?
4. How can the special case of Tunica inform an integrated theory of language change?

To answer questions 1 and 2, I use frequency analysis to analyze the change in several linguistic phenomena across texts that the three linguists (Albert Gatschet, John Swanton, and Mary Haas) gathered from native speakers, as well as one text created in the revived language in 2017. Using the analyzed data, I answer question 3 by showing how the case of Tunica language change during obsolescence and revival compares with existing examples in the literature. For question 4, I look at whether Tunica contributes to an integrated theory of language change.

1.2 Structure of the dissertation

Chapter 2 offers an overview of the history of the Tunica language and the Tunica people, situating the language in its proper cultural and geographic contexts, as well as in its linguistic environment. The chapter introduces the documentation of the language created from native speaker speech, as well as the work done by KYLY, the Tunica Language Working Group, which has endeavored to revive Tunica and make it once again a language of everyday communication.

Chapter 3 provides information on pertinent case studies that have given rise to theories of language obsolescence. By looking at where the field is today, we can

better determine where Tunica might fit into—or how it might confound—these theories.

Chapter 4 discusses the organizational structure and guiding principles that form the basis of the Tunica language revival effort. It also considers issues of language ideology that arise in revitalization and revival situations, and describes how KYLY works to resolve these issues.

Chapter 5 explains the methodology used to create, organize, and analyze the data in the Tunica language corpus. It enumerates the issues with a dataset from different sources and different times, during which the linguistic state of the art had changed, and the efforts that went into normalizing this data into a unified, comparable dataset.

Chapter 6 describes the features of Tunica to be analyzed for evidence of language change, and why each feature was chosen.

Chapter 7 presents the data and analysis, as well as a discussion of whether the changes in Tunica fit into existing theories of language change in obsolescence or revival.

Chapter 2: Tunica language documentation

2.1 The Tunica people

Boas (1942) states that “language is a reflex of culture and...there are everywhere linguistic devices that enable the language to follow the demands of culture” (181). Because language change may occur as a result of changing situations or cultural milieus, when examining language change—specifically in an endangered language—language contact, lifeways, and cultural exchange situate the language in context. This section provides a brief historical sketch of the Tunica people.

In the Tunica creation story, as told by Sesostrie Youchigant to Mary Haas, the Tunica people emerged from a hole in a mountain, *Tamaroha* (Haas 1950:18). Another story tells of the tribe’s migrations. Sesostrie Youchigant describes a tribal migration in several stages (132–135). The first migration came as a result of conflict with the English, but “the Tunica did not want to fight”, so they took their boats down the Mississippi and resettled.¹ The next migration followed a war with another Indian tribe resulting in the death of half the Tunica population. Again they migrated southward via the Mississippi River. The rope that tied the tribe’s boats together broke; half the tribe settled at a certain spot, while the other half

1. Brain (2017) posits that this may refer to a Chickasaw raid prompted by the English (see also Swanton 1911:311).

continued down the Mississippi to parts unknown. The half that stopped were attacked by an Indian tribe while they slept, but they succeeded in repelling the attack.² The final migration mentioned is the one that took the Tunica across the Mississippi River onto the Avoyelles prairie, where they made peace with the Avoyel tribe, settled, and remain to this day.

Much of the information in these stories accords with European accounts. The Tunica first appear in European written records in 1541, when Hernando de Soto's expedition came across the Mississippian settlement at Quizquiz, near the modern boundary between northwest Mississippi and Arkansas (Gatschet 1884). The chief of Quizquiz is described as old and sickly, but able to muster a force of three thousand men in war canoes in a matter of hours, resulting in de Soto suing for peace (Vega [1605] 1723:177–178).

Brain (1988) established the relationship of those at Quizquiz to the Tunica through archaeological comparisons. Luis Hernandez de Biedma, the king's agent on de Soto's expedition, wrote that the settlement "was tributary, like many others, to the sovereign of Pacaha" (Hernández de Biedma 1850:105). Based on the phonology and possible semantic interpretations of Pacaha personal names that Garcilaso de la Vega mentions in his account of de Soto's journey, Rankin (1990) posits that Pacaha was Tunica-speaking (566–567).³

2. This almost certainly refers to the Natchez attack on the Tunica after the Natchez asked for refuge, which the Tunica had granted (Swanton 1911:314; Brain 2017:4–5).

3. Rankin builds upon Swanton's intuition, contained in the final report of the United States De Soto Expedition Commission (1939), that Pacaha was a Tunican language.

In 1699, letters between French missionaries and the Seminary of Quebec mention the “Tonikas” as living at the confluence of the Mississippi and Yazoo rivers, well south of Quizquiz (Shea 1861; Swanton 1946:198). One of these missionaries, Father Antoine Davion, set up a mission at this location and remained there some 20 years, though he found few converts (Shea 1861:55n25). The Tunica people would become closely allied with the French, owing at least partially to this early experience with Davion (Swanton 1946:198).

Benard de la Harpe, writing based on Iberville and Bienville’s memoirs (and later personal experiences) also makes mention of Davion’s 1699 mission in his journal, though he implies that a mission house was already in existence at Davion’s arrival (French 1846:16).

Sometime around 1705, the Tunica migrated south from the confluence of Yazoo and Mississippi rivers to the confluence of the Red and Mississippi rivers (Swanton 1911:311).

Antoine-Simon Le Page du Pratz mentions the Tunica briefly in *Histoire de La Louisiane* (1758), a memoir of his time in Louisiana from 1718 to 1734. He locates the Tunica along the Mississippi near the Red River. Le Page du Pratz’s writings corroborate the Tunica alliance with the French: “[The Tunica nation] has always been very attached to the French; they have even made war alongside us; the chief of that nation was a true friend of ours” (Le Page du Pratz 1758:220).⁴ He

4. In the original French: “laquelle a toujours été très attachée aux François ; ils ont même fait la guerre avec nous ; le chef de cette Nation étoit le veritable ami de la notre.”

also provides the first written information (other than personal names) about the Tunica language, noting the phonetic distinctness of Tunica:

Cette Nation parle une Langue d'autant plus différente de celle des autres, que ces Nations n'ont point la lettre R, au lieu que celle-ci en a beaucoup ; elle a aussi des usages différens.

This Nation speaks a language all the more different from that of the others [in the area], in that those nations do not have the letter R, whereas this one has lots of them; it [Tunica] also has different usages.

Le Page du Pratz (1758:221)

Possibly due to growing British influence in the area that the Tunica inhabited on the eastern side of the Mississippi River after the end of the Seven Years' War (1763), the Tunica gradually began migrating to the Avoyelles Prairie on the western side of the Mississippi, which was under Spanish control. In return for pledging their allegiance to the Spanish, the Spanish granted them title to land near their current location in Marksville, Louisiana. Alliances with the Biloxi, Ofo, and Avoyel—who were all living on the Avoyelles Prairie by the late 1700s—resulted in an amalgamated tribe by 1810 (United States Department of the Interior 1980:3–5).

These accounts of the Tunica might imply that the tribe migrated often, but it may also imply that the Tunica were not all concentrated in one area and that there were multiple Tunica settlements along and near the Mississippi River and its tributaries.

Albert Gatschet, a Swiss-German anthropologist and ethnographer, wrote in 1884 that in addition to the Tunica community located in Avoyelles Parish, some

also lived at the time near Lake Charles.⁵ These two settlements would have been at a distance of over 90 miles (145 km) (Gatschet 1884:40–41). If these two contemporaneous Tunica communities were indeed living at such a distance from one another, there may have been still other communities that went unremarked upon, or were not known by outsiders to be Tunica.

White settlers lodged numerous challenges to Tunica and Biloxi land claims in Avoyelles Parish after the area came under American control with the Louisiana Purchase in 1803. These took the form of legal challenges in American courts, as well as a settler killing Tunica chief Melacon over a land dispute (United States Department of the Interior 1980:12). Despite a subsequent reduction in recognized land, the Tunica continued to live at Marksville as well as the nearby village of Indian Creek.

In the 1930s, a lack of work in Avoyelles Parish led a significant portion of the tribe to travel to Texas to seek employment. Around this time, the tribe began efforts toward government recognition. The tribe did not meet with much success in this endeavor until the 1970s. In 1975, the Tunica-Biloxi Tribe of Louisiana received state recognition, and in 1981, it received federal recognition. The tribe has over 1200 enrolled members, most of whom live in Avoyelles and the surrounding parishes, with communities in Houston, Texas and Chicago, Illinois (Tunica-Biloxi Tribe of Louisiana 2016).

5. Gatschet placed a notice in the *Lake Charles Commercial* on February 7, 1885, asking for “reliable information on the *present* location of the *Tonica* or *Tunika Indians*, or of a part of that tribe” (emphasis in original). His search for Tunica speakers seems to have been in vain, so it remains unclear whether this community existed or not.

2.2 Tunica language classification

Tunica (ISO code: `tun`) is one of the heritage languages of the Tunica-Biloxi Tribe of Louisiana, a federally recognized tribe with land in Marksville, Louisiana (Tunica-Biloxi Tribe of Louisiana 2016; United States Department of the Interior 1980).⁶

Several linguists have attempted to classify Tunica as part of a language family: ethnographer Alphonse Pinart suspected it might be a Caddoan language (Gatschet 1884:41); Swanton thought Tunica belonged to a family that also included Atakapa and Chitimacha, which he named the “Tunican stock” (Swanton 1919:56). Haas for a time designated it as part of her proposed Gulf family, consisting of the Muskogean languages along with Natchez, Tunica, Chitimacha, and Atakapa (Haas 1951:71). Today, the consensus is that there is not enough evidence to classify Tunica as belonging to these or any other language family, and Tunica is thus regarded as a language isolate.⁷

2.3 Analysis of early documentation

Language documentation, while aiming for scientific accuracy and objectivity, is subject to the documenter’s perspective, training, and theoretical preconceptions, and is limited by the methods used at the time of documentation. Linguists can

6. The Tunica-Biloxi Tribe’s other heritage languages are Biloxi (Siouan), Ofo (Siouan), and Avoyel (undocumented).

7. Swanton, looking at early European descriptions of tribal groups in the Lower Mississippi Valley, suggests that Tunica or a related language may have been spoken by several other tribes whose language was never documented, including the Yazoo, Koroa, Tiou, and Grigra (Swanton 1919:7).

disagree on the relative worth of documentary methods and analytical frameworks. This was the case with those who documented Tunica. While Swanton thought Gatschet's fieldwork "exceptionally well done" (Swanton 1921:1), Haas noted that changes in linguistic theory had "almost completely revolutionized" field methods since Gatschet and Swanton's fieldwork (1886, and 1907–1910, respectively) (Haas 1953:175). Nevertheless, Haas was careful to note that she incorporated information obtained by Gatschet and Swanton where possible, rather than following what she saw as the trend at the time to disregard any research that came "before phonemics" (175).⁸

The knowledge and accuracy of the speaker from whom the language is recorded also influences the documentation. In most documentation contexts, the speakers will be native speakers who communicate regularly in their native language with other native speakers. The Tunica documentation context is different; even the earliest documentation took place when Tunica was no longer a widely spoken language, speakers were far from plentiful, and the frequency of its use within a regular speech community is uncertain. Gatschet wrote that his consultant, William Ely Johnson, was "the only Tunica far and wide" (Gatschet 1886a); Swanton wrote when he did his fieldwork that "[p]robably not more than half a dozen use it with anything approaching fluency" (Swanton 1921:1); and Haas noted "there is only one

8. Though the modern conception of the phoneme dates to the late 19th century, it was popularized in the Americanist tradition by Haas's mentor, Edward Sapir (1925). The changes in the linguistic state of the art are also responsible for Haas using phonemic transcription in the published texts she gathered from Sesostrie Youchigant, while Gatschet and Swanton's work was transcribed phonetically.

individual, Sesostrie Youchigant (born c. 1870), who speaks the language with any degree of fluency” (Haas 1940:9). This helps explain why each time frame features Tunica texts gathered from only one consultant.⁹

Because the data are from so few speakers and there is a lack of any analyzable audio recordings,¹⁰ a close look at the documenters’ backgrounds in Section 2.3.1 will be helpful in interpreting the data they created.

There is also uncertainty about speaker fluency in the Tunica context due to the very small number of remaining speakers and the fact that the speakers who were consulted may not have regularly used the language at the time the fieldwork was done. Swanton wrote in 1921 that “the language itself is fast falling into disuse” (Swanton 1921:1). Haas suggested in 1940 that Tunica “was probably in a more or less decadent state even two or three generations ago” and wrote that Sesostrie Youchigant “has had no occasion to converse in Tunica since the death of his mother in 1915”, 18 years before he spoke with Haas (Haas 1940:9). The biographies of the speakers in Section 2.3.2 are important to provide some idea of how frequently they might have spoken Tunica, and to determine possible influences on their speech from other languages they spoke.

9. Swanton gathered cultural information and checked Gatschet’s existing work with more than one consultant. He refers to Volsin Chiki, from whom he gathered Tunica texts, as “one of [his] principal informants” (Swanton 1911: plate 17a), and in the same work includes a photo of William Ely Johnson, whom he describes as “one of the few who still speak the old language” (plate 18a). However, the only texts he gathered at that time (1907-1910) are from Volsin Chiki. He later obtained texts from Sesostrie Youchigant in 1930 and 1931.

10. Mary Haas recorded Sesostrie Youchigant speaking Tunica on wax cylinders. The California Language Archive has digitized these recordings, but they are so degraded that making out words is difficult. Haas noted that even when first recorded, the recordings were “not of the best quality and some words are too blurred for identification” (Haas 1940:19n1).

The sections below provide brief backgrounds of those who documented the language and the speakers from whom they obtained it.

2.3.1 The documenters

Albert S. Gatschet

Albert Gatschet was a Swiss philologist and ethnographer. He studied at the Universities of Bern and Berlin, and wrote a book on the etymologies of Swiss placenames before moving to New York in 1868, where he wrote about and taught languages (“Albert Samuel Gatschet—1832–1907” 1907:561). In 1877, he was hired by the Geographical and Geological Survey, and was an original member of the Bureau of American Ethnology (BAE) upon its founding in 1879. He did fieldwork throughout the United States and Mexico, documenting and analyzing nearly 100 Native American languages, notably Klamath, about which he wrote one of the first detailed studies of a Native American language (562–563, 565). In the mid-1880s, he began documenting several Southeastern languages.

Having read many of the accounts of contact with the Tunica mentioned in section 2.1—including Le Page du Pratz’s description of the Tunica language as fundamentally different from surrounding languages—Gatschet thought Tunica of particular interest, as it might represent an unstudied language family (Gatschet 1884:41). Gatschet had worked with Louisiana tribes on trips in 1881–1882 and 1884–1885. On his second trip, he had tried to locate Tunica speakers near Lake Charles, Louisiana, placing a notice in the *Lake Charles Commercial* in February

1885. Presumably, this effort was unsuccessful. On his third trip to Louisiana, in the Fall of 1886, he visited Lecompte, Louisiana, with the intention of documenting Tunica (“Albert Samuel Gatschet—1832–1907” 1907:563; Gatschet 1886a).

On this trip, Gatschet located Tunica speaker William Ely Johnson (also known as Etienne Chiki) and worked with him in October and November 1886 to gather information about the Tunica language. This resulted in 259 handwritten pages of vocabulary, phrases, and texts, along with ethnographic information (Gatschet 1886b). Johnson and Gatschet both spoke varieties of French, facilitating communication as Gatschet elicited Tunica words and phrases from Johnson (Gatschet 1886a; Swanton 1921:1).

In a letter reporting the information he had gathered on Tunica (and Biloxi) to the director of the BAE, Gatschet wrote that he was not sure whether Tunica represented a “new stock” or whether it belonged to the “Pa’ni” (Caddoan) language family. He also made it explicit that his level of interest in the language was contingent on whether it was an isolate.

If it is a Pa'ni [Caddoan] language I would not attach so much importance to this discovery to as to devote weeks or months upon its study; but if it forms a new stock, it would be worth a profound study here on this spot...

Gatschet (1886a)

Only because Gatschet determined that Tunica was an isolate did he spend considerable effort documenting it.

Gatschet published articles that mention Tunica in passing, but the only publication where Tunica is the focus is an article he published on grammatical gender in Tunica, a rare feature among documented languages in the geographic area

(Gatschet 1889). Most of Gatschet's work on Tunica remained unpublished until John Swanton undertook to verify Gatschet's work and incorporated it into his own work on Tunica.

John Swanton

John Swanton was an American anthropologist and ethnologist. After graduating from Harvard, he studied linguistics under Franz Boas for two years at Columbia University before obtaining his doctorate in anthropology from Harvard in 1900 with a thesis describing Chinook verb morphology (Swanton 1900). Swanton also documented Dakota during this time. After earning his doctorate, he continued his fieldwork in the Pacific Northwest and joined the BAE. Swanton's work in the Pacific Northwest included linguistic descriptions of Haida and Tlingit.

For much of his career, his focus was on the languages and cultures of the American Southeast. He published extensively on the cultures of Southeastern peoples, most notably in *Indian Tribes of the Lower Mississippi Valley and Adjacent Coast of the Gulf of Mexico* (1911). He published BAE bulletins on several indigenous Southeastern languages, including Biloxi, Ofo, Choctaw, Atakapa, Chitimacha, and Tunica (Kroeber 1940:3).

Swanton traveled to Louisiana in 1907, 1908, and 1910 to verify and expand upon Gatschet's 1886 fieldwork. He obtained two stories in Tunica from chief Volsin Chiki, and also spoke to Gatschet's consultant, William Ely Johnson, to check information Gatschet had gathered in 1886. He typed Gatschet's elicited vocabulary on cards,

and used these cards to verify their content, writing corrections where he noticed phonetic or lexical differences from Gatschet's work.

Swanton (1911) discussed some of the cultural information he gathered during these visits. Later, Swanton (1921) compiled his grammatical analysis into a sketch grammar of Tunica, the first detailed analysis of the language other than Gatschet's article on grammatical gender (Gatschet 1889). Like Gatschet, Swanton was intrigued by Tunica's grammatical gender system: "The most remarkable feature of Tunica is the prominence given to the expression of sex", which he mentions as being one of the features separating Tunica from other Southeastern languages (Swanton 1921:6, 35–36).

In 1930 and 1931, Swanton returned to Louisiana to gather more information about Tunica, this time with two stories gathered from Sesostrie Youchigant (Swanton 1930c, 1931).

He sent much of what he'd gathered of Gatschet's material, along with his own, to Mary Haas in 1933, when she was doing fieldwork on Tunica (Swanton 1930a; McLendon 1997:528–529).

Mary Haas

After studying English as an undergraduate at Earlham College, Mary Haas enrolled in the University of Chicago's comparative philology graduate program in 1930, where she studied under Edward Sapir and Leonard Bloomfield, among others (McLendon 1997:525–526). She did fieldwork in 1931 with Morris Swadesh on

Nitinat (Ditidaht) on Vancouver Island, focusing on Nitinat music as well as the language itself.¹¹

She followed Sapir to Yale in 1931, where she served as his research assistant. Sapir worked with Franz Boas to send Sapir's students to the field to study American languages. The goal was to document at least one language from each indigenous North American language family. Boas assigned Haas to work on Tunica in 1933; she traveled to Louisiana and began her work with Sesostrie Youchigant. Unlike Gatschet, who had communicated with William Ely Johnson in French, Haas used English to communicate with Sesostrie Youchigant.¹² That winter, Swanton sent her some Tunica materials that he had compiled from Gatschet's and his own fieldwork. Haas made a subsequent visit to work with Sesostrie Youchigant in the Summer of 1934, while also working on Natchez, to verify this material (McLendon 1997:528–529). On this trip, she hired a Creole French interpreter to help her understand terms for which Youchigant did not know the English equivalent (529). Her dissertation, which she defended in 1935, was a description of Tunica grammar.

Haas made subsequent trips to Louisiana in 1938 and 1939 to gather and verify material for her compilation of Tunica texts (Haas 1950) and a revised and expanded version of her Tunica grammar (Haas 1940). In 1953, she published a Tunica dictionary (Haas 1953).

11. Haas's musical education and her musicianship would help her in this work, as well as in documenting Tunica songs and the prosody of the Tunica language.

12. Mary R. Haas, interview by Ernest C. Downs, August 7, 1975, transcript.

In the late 1930s and early 1940s, Haas worked on Creek, Choctaw, Seminole, Koasati and other Muskogean languages, as well as Yuchi (isolate) and Biloxi (Siouan) (McLendon 1997:530–533). She taught Thai at Berkeley in the early 1940s and became a faculty member there in 1946, where she would remain until her retirement in 1975 (524, 538).

In all, Haas's field notes on Tunica comprise over 1300 pages. Her grammar is by far the most detailed, being more than three times the length of Swanton's sketch grammar. Gatschet and Swanton together gathered fewer than 20 stories and narratives. Haas gathered more than 50 unique stories from Sesostrie Youchigant, sometimes gathering two or three tellings of each tale.

Haas attempted to place Tunica within her proposed Gulf language family, alongside the Muskogean languages, Natchez, Chitimacha, and Atakapa (Haas 1951:71). Later in her career she questioned whether similarities between these languages indicated a genetic relationship, or if they were areal features (Golla, Matisoff, and Munro 1997:831).

All three documenters were most interested in typologically categorizing the Tunica language, but their specific backgrounds and the time period of their research dictated how they conducted that research. Swanton seemed impressed with Gatschet's work on the language, and wrote that he "was able to add little to it except in the way of verification". He attributes this in large part to Gatschet being able to communicate with speakers in French (Swanton 1921:1). Haas, conversely, viewed both Gatschet's and Swanton's work as lacking in important re-

spects. Gatschet and Swanton, Haas argued, were working before the advent of phonemic theory (Haas 1953:175); she notes certain gaps in transcription, such as Gatschet’s failure to include glottal stops in his transcriptions (Haas 1940:9n1).¹³ Haas was also puzzled by Swanton’s sketch grammar of Tunica, saying in an interview that it “is the strangest thing I’ve heard: it’s just plain wierd [sic]”.¹⁴ Swanton, on his part, suspected that Haas’s work would supersede the information he had gathered on Tunica: on a copy of a set of Tunica texts from both his own and Gatschet’s fieldwork that he had sent to Haas, he wrote, “Material probably made unnecessary by work of Mary Haas” (Swanton 1930a).

While Gatschet and Swanton’s work predated the popularization of phonemic theory, and their transcriptions lacked distinctions between certain phones, their documentation is invaluable. They captured Tunica language earlier than Haas and—minor phonetic inadequacies aside—provided reliable data concerning morphology, syntax, and the lexicon. And though the terminology Haas used is sometimes outmoded or confusing,¹⁵ the concepts she lays out still serve as a guide for analysis.

13. In addition to not transcribing glottal stops, Gatschet also fails to differentiate between [o] and [ɔ]. Swanton similarly both omits glottal stops and does not differentiate between [o] and [ɔ].

14. Mary R. Haas, interview by Ernest C. Downs, August 7, 1975, transcript.

15. Haas’s use of the terms “mutable durative” and “immutable durative”, for instance, is puzzling. “Mutable” and “immutable” are not familiar terms in linguistics, and Haas does not clearly define them. She notes that “[i]n English translations it is sometimes difficult to render adequately the subtle distinction between the two types of duratives. Both may be rendered by the English present or past progressive... However, the immutable durative may also be rendered by the English simple present or past...” (Haas 1940:50).

2.3.2 The speakers

Documented Historical Tunica¹⁶ consists of only three speakers before the break in intergenerational transmission. Speaker variation is present in all languages, but in obsolescing languages, speakers may not speak the language often, or may not have spoken it at all for quite some time. Thus, speaker ability can vary more widely in an obsolescent language, and it is helpful to look at what we know of the speakers whose language was documented.

William Ely Johnson

William Ely Johnson, also known as Étienne Chiki, worked with Gatschet in 1886. Johnson spoke five languages; in his autobiographical statement to Gatschet in Tunica, Johnson noted that he also spoke French, English, Tunica, Choctaw, and Biloxi (Gatschet 1886b:54). Gatschet spoke effusively of him, calling him excellent and even suggesting the possibility of bringing him to Washington, D.C., presumably to continue working on the language (Gatschet 1886a).

Johnson was 29 in 1886, and his description of his upbringing provides an explanation for the languages he speaks. He was born to a Tunica father and Biloxi mother. His father died when Johnson was two. At the age of nine, Johnson's mother took him to live among white settlers. She died when he was 10, and he

16. This dissertation refers to Tunica spoken before the break in intergenerational transmission Historical Tunica.

afterwards grew up in “Tunica country”, likely the Tunica settlement in Avoyelles Parish (Gatschet 1886b:54–55; Haas, n.d.:5).

Johnson claimed to be the last Tunica speaker in the 30-person village in Le-compte, Louisiana, where he was living when he worked with Gatschet. Johnson told Gatschet that the village was full of young people, that all the old people had died, and that of his brothers, he was the only one who remembered how to speak Tunica (Gatschet 1886b:85–86; Haas, n.d.:9). Based on the situation described, Johnson would not have had the opportunity to speak Tunica with others daily at the time of Gatschet’s visit.

In addition to elicited words and phrases, Johnson related eight texts to Gatschet, from a short excerpt about warriors to a recitation of the story of the Orphan Boy. While Johnson may not have spoken Tunica daily, he remained familiar enough with it to speak in Tunica on a variety of topics.

In 1908, Johnson also worked with John Swanton, answering Swanton’s questions about the information Gatschet had obtained from him (Swanton 1908).

Sesostrie Youchigant, who worked with Mary Haas, said that Johnson’s first language was Biloxi and that his Tunica “left much to be desired” (Haas 1953:179). Based on Johnson’s own description of his childhood, it seems plausible that Tunica may not have been Johnson’s first language. However, Swanton’s consultants (including Johnson himself) checked information that Gatschet gathered from Johnson, and when Sesostrie Youchigant worked with Haas, she went over some of Gatschet’s material with him, making this material more vetted than the texts of other consultants.

Volsin Chiki

Volsin Chiki was chief of the Tunica from circa 1890 until 1911, when he stepped down due to “blindness and other infirmities” (United States Department of the Interior 1980:16). He worked with Swanton in 1910, as what Swanton referred to as his “principal informant” (Swanton 1911: plate 17a). He was a cousin to William Ely Johnson through his father’s brother. Swanton visited Marksville three times—in 1907, 1908, and 1910. In 1910, he gathered two texts from Volsin Chiki: *The Woman who Married a Rattlesnake*, and *the Flood Story*. Volsin Chiki also provided the longer *Thunder story* to Swanton, but only in English (319–322).

Sesostrie Youchigant

When Volsin Chiki stepped down as chief, the position went to his nephew, Sesostrie Youchigant, who was born circa 1870. His father moved away when he was young, and he was raised by his mother, who farmed cotton and corn, which he would grow up to farm as well (Haas 1950:150–151).

Sesostrie Youchigant did not converse in Tunica after his mother’s death in 1915. Even before her death, he was more likely to respond to her in French than in Tunica, though she always spoke to him in Tunica (Haas 1940:9). His English was somewhat limited; in a note to Swanton accompanying stories she had gathered from Sesostrie Youchigant in English, Caroline Dorman wrote, “I should like to see what these would turn out to be if taken in Tunica. His English is so inadequate”

(Dorman 1931). Haas noted that in addition to Tunica, he spoke “Louisiana French and English, the former more readily than the latter” (Haas 1940:9).

Swanton thought Sesostrie Youchigant was “one of the best informants [he] ever had, willing and able to answer questions of all kinds”, and mused about bringing him to Washington to work with him further (Swanton 1930b). Haas was impressed with the number of stories that he knew, but spoke of trouble eliciting isolated words from him, due to him being hard of hearing and not understanding her questions well (Haas 1933). Haas noted that he was “willing, patient, and helpful”, but that he expressed the difficulty involved in remembering a language he had not spoken in 20 years (Haas 1953:175).

Sesostrie Youchigant worked with Swanton in the Summer of 1930 and in November 1931. He worked with Haas in 1933, and again over four more visits between 1934 and 1939 (175). His speech constitutes the vast majority of the Tunica-language corpus. It formed the basis of Haas’s detailed grammatical analysis of Tunica, whose work, in turn, is the basis of revived Modern Tunica.

Looking at what is known of the biographies of the speakers, all were multilingual, though what effect this had on their production of Tunica is difficult to ascertain. As for their use of Tunica, it is likely that William Ely Johnson did not speak Tunica daily, as he was “the only Tunica far and wide” (Gatschet 1886a). It is known that Sesostrie Youchigant “had no occasion to converse in Tunica since the death of his mother” nearly twenty years before he spoke to Mary Haas (Haas 1940:9). Information about Volsin Chiki’s history, beyond his role as tribal chief, is

scarce. But he lived in proximity to Sesostrie Youchigant, and Swanton writes that “not more than half a dozen” spoke the language proficiently (Swanton 1921:1); in contrast to the other two speakers, Volsin Chiki may well have had a small speech community with whom to converse in Tunica, though it is unknown how frequently this occurred.

All three speakers may have suffered from deficits in linguistic performance due to their not having spoken Tunica regularly in the time just before they worked with the linguists who documented their speech.

2.4 Conclusion

The early historical record of the Tunica people is contained in three Tunica historical narratives—one from William Ely Johnson and two from Sesostrie Youchigant—along with brief mentions from several European sources. The path of their migration, the Europeans who wrote about encountering them, and the other languages spoken the three Tunica speakers suggest that the languages of the peoples with whom they were likely in close contact may have included Choctaw, Chickasaw, Natchez, Biloxi, Ofo, Avoyel, French, Spanish, and English.

The three documenters of Tunica all remain well-respected linguists. They each worked with the theoretical and technical tools at their disposal to document Tunica speech. Gatschet and Swanton failed to recognize the phonetic distinction between [o] and [ɔ], and both failed to notate the glottal stop ([ʔ]). Haas was more precise in phonetic notation, but her detailed grammar of the language still suffers from

certain puzzling or outmoded terms, and analyses that bear retesting with current tools.

The biographies of the Tunica speakers establish how small each speaker's speech community was and the infrequency with which each had opportunities to speak Tunica. These concerns, along with Sesostrie Youchigant's claim that Johnson's proficiency in Tunica was questionable, provide further motivation for the comparison of these speakers' texts with one another to assess their differences and to determine which, if any, might be ascribed to language attrition.

Chapter 3: Language obsolescence

Language obsolescence is a relatively new field of linguistic study. It can be broadly defined as the process of the decreasing use of a language by its speech community and the attendant reduction of the speech community itself as they cease speaking the language entirely. Since in-depth analysis of the linguistics of obsolescent languages began in the 1970s, work in this area has flourished.

I have used *language obsolescence* in this work rather than one of the several related terms that are common in the broader field of language endangerment and revitalization. Among the most common are *language death*, *language attrition*, and *language shift*. *Language death* implies a finality that is belied by projects of language reclamation, reawakening, and revival; the term also focuses on the endpoint of the process rather than on the process itself. *Language attrition* is oriented more to the speaker's command of the language than the state of the language itself. Because the documentation of Tunica contains only speculation about speaker competence—and because there is no vibrant speech community to which to compare a specific speaker's competence—using the term language attrition is inapt here. *Language shift* is an essential process in any non-catastrophic language endangerment situation, but there is scant documentation of each speaker's relative proficiency in the languages he spoke, and multilingualism is only briefly discussed in the present work, rather than being a central focus. *Language obsolescence* fits

best with the aim of this study, which is to examine the evolution of the language as its speaker population dwindled and the language gradually ceased to be used.¹

The first academic treatment of language obsolescence as a category of study was “Sociologic Notes on Obsolescent Languages” (Swadesh 1948). In the article, Swadesh analyzes nine languages, seven of which are native languages of North America. The “modest aims” of Swadesh’s analysis are to provide exemplars of obsolescent languages and the situations accompanying their obsolescence (226). While the linguistic effects of language obsolescence are beyond the scope of his work, his conclusion is prescient: he views the linguistic effects of language obsolescence as an area of legitimate study, and he lays out a framework for conducting research in the area, including taking autobiographies from speakers to determine language attitudes and linguistic ideology (which Swadesh terms the “sociology of language conflict”) and closely examining which aspects of the language change first and which are more resistant to change (234).

Current theories posit four main pathways to language obsolescence: *sudden language death*, in which all speakers of a language die either due to natural disaster or genocide; *radical language death*, where political oppression causes the entire speech community to stop speaking the language; *bottom-to-top language death*, where the language continues to be spoken in formal settings, but ceases to be spoken in day-to-day interactions; and *gradual language death*, in which speakers of

1. It is important to note that “obsolescence” is used in this dissertation in its basic sense of “becoming less and less used”, and not “becoming less and less useful”.

a language gradually shift to speaking a different language, first in certain contexts, and finally in all contexts (Campbell and Muntzel 1989). Most cases of language obsolescence—including Tunica’s—fall into this last category.

In gradual language obsolescence, the language loses prestige and begins to be used in progressively fewer domains. This often occurs due to political, economic, or social domination by another group speaking a different language. The swift, three-generation language shift of a non-English-speaking immigrant to the United States to that immigrant’s monolingual English-speaking grandchild is a stark example of the power of political, economic, and social hegemony to spur language shift.

The documented history of the Tunica speech community was one in which the community was not often in a position of power. Gatschet, Swanton, and Haas all note the small size of the Tunica community, and with the Muskogean-lexified Mobilian Trade Jargon being the language of wider communication in the area inhabited by the Tunica, the Tunica language was likely subject to linguistic pressures encouraging language shift even before sustained European contact.² The increasing political and economic power of French-, Spanish-, and English-speaking settlers in Louisiana doubtless accelerated language shift—first to French, then to English—as much of the economic and political life of the area was increasingly conducted on terms and in the language prescribed by settlers.

2. Haas posits that languages of wider communication—both Mobilian and French—were part of the reason why Tunica usage had been in “a decadent state” for some time by the 1930s (Haas 1940:9–10).

3.1 Structural consequences of language obsolescence

Linguists have commented on structural consequences of language obsolescence for as long as they have studied such languages. Scholars in the 19th century viewed languages as analogous to organisms, having a life cycle that inevitably ended in decay and death. Franz Bopp echoed the opinions of other prominent 19th-century linguists when he described languages as “organic natural bodies, which are formed according to fixed laws, develop as possessing an inner principle of life, and gradually die out because they do not understand themselves any longer, and therefore cast off or mutilate their members or forms, which were at first significant, but gradually have become more of an extrinsic mass” (1827; quoted in Jespersen 1922:55). Conceptualizing obsolescent languages as “mutilated” and comparing languages to biological organisms both became less prominent in the 20th century.

In the North American context, the study of obsolescing languages began in earnest with Franz Boas and Edward Sapir’s project of gathering data on as many indigenous languages as possible to determine where they fit within language families. Because of the language situation in the United States and Canada, many of these languages—including Tunica—were highly endangered and had dwindling speaker populations, if they had not already ceased to be spoken entirely. In the early 1900s, linguists would comment on issues like the varied linguistic abilities they saw between speakers of the same obsolescing language.³

3. See, for instance, Bloomfield (1927), for a discussion of the varied levels of simplification in Menominee speech between different speakers, which he attributes to processes of language contact and language shift.

The first in-depth study of an obsolescent language that looked at the linguistic ramifications of obsolescence was Nancy Dorian's work on East Sutherland Gaelic (Dorian 1973, 1977, 1978, 1981). Subsequent work has shown a variety of changes that occur as specific languages obsolesce. Early studies of language obsolescence examined obsolescing dialects or first language attrition among immigrant communities with little contact with their linguistic homelands. One notable exception to this early trend is *Young People's Dyrbal* (Schmidt 1985b), which examined the generational differences of how the Dyrbal language, an indigenous Australian language, was spoken.

3.1.1 Common structural changes in obsolescence

As the number of studies of obsolescing languages continues to grow, the changes that emerge vary widely and depend on the unique circumstances of the language. Nevertheless, certain changes tend to occur in languages whose speech community and domains of use are in decline.

Broadly speaking, obsolescent languages tend to undergo a reduction in forms and functions, as well as an increase in unpredictable variability due to inconsistent application of grammatical rules. These processes can affect all levels of linguistic structure and are affected by the dominant language to which speakers of the language are shifting. Marked features—those that occur in the obsolescing language but not the language that the speech community is shifting to—are more likely to be subject to reduction and increased variability (Andersen 1982; Dorian 1973). Dorian describes variability in the obsolescent dialect of Eastern Sutherland Gaelic

(ESG) as “a dismal patchwork of inconsistencies and (from the point of view of the standard language) mistakes, haphazardly distributed over villages, speakers, and occasions” (Dorian 1973:414).

Phonetically, sounds that are not common in the dominant language are more likely to change than sounds common to both languages. In Tunica, for instance, one would expect phonemes like preaspirated stops, which are not shared with the dominant language, to be more prone to reduction than a phoneme that is shared between both languages. Campbell and Muntzel (1989) note that some Pipil speakers have merged previously distinct long and short vowels into short vowels due to the lack of a phonemic vowel length distinction in Spanish. At the same time, as phonological rules become inconsistently applied throughout the speaker base, variants like the use of a short vowel where a long vowel is called for develop.

Morphemes are also subject to reduction and variation. Dorian (1977) cites, for instance, the complete loss of word-final case marking for the genitive in ESG. Regularization of irregular forms and a preference for analytic structures over synthetic ones also result in the loss of morphemes (Aikhenvald 2020; Dorian 1977). In Tunica, a morpheme like =*ani*, a quotative evidential, might be expected to be subject to loss, as it has no parallel structure in English or French and can often be inferred by context. Morphemic reduction can manifest as a loss or weakening of case or gender marking where the dominant language does not have similar systems (Aikhenvald 2020). Tunica’s system of grammatical gender contrasts with the lack of such a system in English, and a system with different rules in French.

Syntactic reduction takes the form of a reduction in sentence complexity. In addition to sentences choosing more analytic constructions, obsolescent languages tend to have fewer complex sentences (Voegelin and Voegelin 1977). For quantitative purposes, the decrease in dependent clauses can be roughly measured by comparing the number of verbs per sentence (Wolfram 2008). Syntactic change can also loosen word order rules for languages whose word order differs from the language to which the community is shifting (Aikhenvald 2020). Conversely, for those obsolescent languages with robust case structure and relatively free word order, the loss of case through morphemic simplification can result in stricter word order constraints, as case is determined through ordering rather than through case marking (Campbell and Muntzel 1989). Freer word order due to the influence of English and French would be a possible locus of change in Tunica.

Lexical changes are most apparent in the loss of less frequently used lexical items, especially words only relevant to domains in which the language is no longer used at all. An influx of loanwords and calques may fill in these lexical gaps, as well as provide terms for new concepts and objects (Wolfram 2008).

3.2 Rate of change in obsolescence

A common non-structural consequence of obsolescence is its effect on the rate of language change. Schmidt (1985b) writes that in Dyirbal “vast amounts of change are compressed into a short timespan of about 25 years” (213). Dorian (1973) describes a rate of change in ESG that “has been rapid enough to show clear differences between the usage of the oldest and the youngest fluent speakers available, a span of

just over 40 years” (Dorian 1973:413). The pressures on endangered languages due to intense contact with a dominant language, decreasing domains of use, and the possibility of rapid cultural and political shifts all work against linguistic stability.

Another factor contributing to an increased rate of change is not unique to obsolescing languages: the size of the speech community. Computer simulations of language change have suggested that speaker population is inversely correlated to rate of linguistic change (Nettle 1999).⁴ Unless the endangered language goes from vibrant to unspoken in a very brief timeframe, the smaller speech community should have an impact on the rate of change.

3.3 Conclusion

Every endangered language is endangered in its own way, depending on its circumstances. Language is not a biological organism, destined to obsolesce after a certain time as it decays with old age; language obsolescence is nearly always due to the political, economic, and cultural domination of a linguistic group. The intensity of language contact, the speed of language shift, the degree of domination, and the similarities and differences between the dominant and endangered language all affect the language change that occurs.

Certain changes are more common to obsolescing languages, including reduction of forms and functions, a move from synthetic to analytical structures, loss of case

4. Conversely, Nettle (2012) found that smaller speech communities were more likely to have more complex morphology. Perhaps the rate of linguistic change holds in obsolescent languages due to the size of the speaker population, while obsolescence confounds the general finding of morphological complexity among smaller speech communities.

and gender systems, and loss of lexical items. But the literature contains many counterexamples. For instance, Dorian (1978) says that ESG “might be said to be dying, at least with regard to noun plurals and gerunds, with its morphological boots on” (608).

Because language change in obsolescence is so dependent on circumstance, examining how obsolescence affected Tunica, given what is known of its linguistic environment as described in Chapter 2, will provide additional information about the effects of obsolescence in differing contexts.

Chapter 4: Tunica revival

Though the last known native speaker of Tunica, Sesostrie Youchigant, died in 1948, community interest in the language continued. In 1964, Chief Joseph Pierite, Jr. wrote to Mary Haas, saying that “there are only a few of my people, including myself which know how to speak our language a little”. He requested copies of her work, as “we would like the younger ones of our tribe to learn the language and songs” (Pierite 1964). Haas sent him copies of her work that same year. In the 1970s, Chief Joseph Pierite, Jr.’s daughter-in-law, Donna Pierite, and her family began working mainly with Haas’s material, but with Gatschet and Swanton’s work as well, in order to learn and teach the language (Tunica-Biloxi Indian Tribe 2017:45; Klopotek 2011). Donna Pierite ran Tunica language summer camps at her home in New Orleans in the 1990s, and contributed a Tunica language section to the Tunica-Biloxi tribe’s newspaper, *Tawaka*.

4.1 *Kuhpani Yoyani Luhchi Yoroni* (KYLY): The Tunica Language Working Group

In 2010 Tunica-Biloxi Tribal Council Member Brenda Lintinger approached Tulane University linguistics professor Judith Maxwell suggesting a collaborative relationship between the two institutions with the goal of reviving the Tunica language (Tunica-Biloxi Indian Tribe 2017:45). Maxwell agreed and enlisted the help of

Table 4.1: Development of Tunica orthography from Haas to Modern Tunica

Haas	Early revival	Current
č	ch	ch
š	sh	sh
ɛ	e	ɛ
e	ei	e

Tulane linguistics students. This Tunica Language Working Group, or Kuhpani Yoyani Luhchi Yoroni (KYLY) in Tunica, studied the structure of the language as laid out in Mary Haas’s Tunica grammar (Haas 1940). They developed a standardized orthography that they used in the first publication of the revitalization effort, a children’s book adapting two of the Tunica stories that Sesostrie Youchigant had told to Mary Haas (Youchigant and Haas 2011). Table 4.1 shows the characters that differ between the orthographies of Haas’s transcriptions; the early revival orthography that was used in the 2011 children’s book; and the current Modern Tunica orthography, which was gradually revised into its current form based on community feedback.

Over time, ideas for the project grew in both number and complexity. Additional publications, language summer camps, immersion workshops, and cultural activities would require organization and coordination. The Tunica-Biloxi tribe responded to these needs by creating the Language & Culture Revitalization Program (LCRP).

4.1.1 The Language & Culture Revitalization Program (LCRP)

The Tunica-Biloxi Language & Culture Revitalization Program (LCRP) was created in 2014 as a new department of the Tunica-Biloxi tribal government. The

LCRP's project manager, Brenda Lintinger, and program director, John Barbry, provided both administrative support and a formal channel to the tribal council. This provided more structure and funding for the continuing efforts to reintroduce and teach the language to the wider community, as well as funding for cultural education. It also allowed Donna and Elisabeth Pierite to continue their work as language and cultural lifeways instructors, now under the auspices of the LCRP.

In 2017, the LCRP was able to obtain funding to hire a linguist to work with them in Marksville full time. The linguist would be responsible for aiding other LCRP staff in interpreting historical documentation, creating new pedagogical materials, developing curricula, and teaching language classes. I became the tribe's first on-site linguist in 2017, working as a member of the LCRP.

Just before my departure in 2018, the LCRP drafted a grant proposal to the Administration for Native Americans (ANA) requesting funding for the training of Tunica language teachers. The three-year grant would fund the hiring of language apprentices for whom the LCRP would serve as mentors. The mentors would teach them the Tunica language and Tunica-Biloxi cultural practices. The goal of the grant was to have the apprentices complete the language and instructor training, gain intermediate mastery of the language, and become Tunica language instructors (Tunica-Biloxi Tribe of Louisiana 2018).

The ANA grant was successful, and Tunica-Biloxi tribal members were hired as language apprentices by the LCRP. Three of those language apprentices, having gone through three years of language and cultural training, have become assistant language instructors.

KYLY has continued as an umbrella group comprising the LCRP mentors, assistant language instructors, and on-site linguist; Dr. Maxwell; Tulane linguistics graduate and undergraduate students; Tulane linguistics alums; and other interested scholars.

4.1.2 Organizing a community-engaged scholarship (CES) project

Because the project consists of two main groups separated by a three-hour car ride—one in New Orleans and one in Marksville—frequent videoconferences have become the primary way to ensure the needs of the LCRP in Marksville and work of the linguists at Tulane (and elsewhere) remain aligned.

KYLY operates on the philosophy of community-engaged scholarship¹ (Anderson 2017:3). There are many definitions of CES’s central tenets. At its core, CES “must have an intentional *public purpose* and direct or indirect benefit to a community” (emphasis in original) (Stanton 2008:24). KYLY’s public purpose involves not just any community, but the tribal community whose members form much of the group itself; the language with which KYLY works is the heritage language of the Tunica-Biloxi tribe, so KYLY’s mission is to have its efforts—both pedagogical and academic—support the revitalization of Tunica in the tribal community.

The philosophy of CES that KYLY follows is more similar to the following definition for community-based participatory research (CBPR): “CBPR promotes trust and shared power and decision-making between researchers and community repre-

1. Community-engaged scholarship is known by many names, including community-engaged research, community-based participatory research, and public scholarship.

sentatives, 2-way capacity building, and mutually beneficial co-creation and dissemination of study findings” (Mikesell, Bromley, and Khodyakov 2013:e7). Baldwin et al. (2022), discussing the Myaamia revitalization effort, define community-engaged research as “a framework that seeks and nurtures community involvement, leverages community knowledge, and is led by community need” (176) and define it in the Myaamia language as *neepwaantiinki*, or “partners in learning” (169). KYLY builds trust through frequent communication, consensus-based decision-making, co-authored research, and knowledge-sharing. In short, KYLY forms a community of partners in learning.

While KYLY works hard to follow the principles of community-engaged scholarship, there are built-in power imbalances that the group must constantly guard against. Baldwin et al. (2022) caution that “[a]ll too often the [community-engaged] research enterprise is dominated by academics who seek minimal community input” (169). Whitaker (2017) describes the dynamics in the Tunica language revitalization program: “The linguists and those who volunteer from Tulane are afforded much by way of expertise, and as such are put into a position of gatekeeping linguistic knowledge, much like traditionalist elders of other revitalization programs” (2). Because Historical Tunica documentation is written by linguists for linguists, using terms that are sometimes outdated, linguists are seen as both translators and synthesizers of this knowledge into current and consistent terminology that will make sense for teaching—rather than just describing—the language. Whitaker (2017), citing Heaton (2013), notes the importance of linguists for “tak[ing] the language from a place of description to a place of creation” (Whitaker 2017:23).

4.1.3 KYLY goals

There are many decisions to be made within any language revival (or revitalization) project. If they are made without community involvement, the language workers may work toward goals that do not address community needs or desires, eroding community support. If the decisions are not agreed upon by consensus, the group may suffer from a lack of concerted effort toward the defined goals; instead, language workers may work at cross-purposes, prioritizing incompatible goals.²

Prior ideological clarification limits these issues by addressing language ideologies at the outset of language programs (Dauenhauer and Dauenhauer 1998). Ideological clarification “covers the conflicts of ‘beliefs, or feelings, about languages’ (Kroskrity 2004) that are the inevitable outcome of the interaction of indigenous, colonial, post-colonial, and professional academic perspectives. The differences between these points of view are displayed and even magnified by language renewal activities” (Kroskrity 2009). These perspectives inform a project’s goals and strategies, and conflicting points of view should be harmonized before a language revival project is undertaken. The goals and tools that emerge from this period of ideological clarification will determine not only how the language is revived but also what the revived language will look like.

When KYLY began, the advisory group determined that the goal would be the reintroduction of Tunica as a language of everyday conversation among mem-

2. Projects can, of course, deliberately organize themselves into subgroups that work on different aspects of the revival effort in furtherance of a unified goal.

bers of the tribe. The project has endeavored to create a revival structure that is geared toward an immersion classroom environment. Much of the structure of the Tunica language pedagogy is derived from the language programs Judith Maxwell helped design for *Oxlajuj Aj*, a Kaqchikel language program in Guatemala. Both the Kaqchikel and Tunica language programs aim to teach language in an immersion setting using a five-step methodology derived from communication-based instruction (CBI) (Supahan and Supahan 2001).³

KYLY took Haas’s Tunica grammar (Haas 1940) and *Tunica Texts* (Haas 1950) as the foundation of Modern Tunica. Despite the thoroughness of Haas’s examination of the language, gaps remain. In the introduction to her Tunica grammar, Haas wrote, “While it cannot be denied that there are certain lacunae in this material, most of these are to be interpreted as lacunae in [Sesostrie Youchigant’s] knowledge of the language” (Haas 1940:10). Whether these gaps come from a lack of speaker knowledge or lack of exhaustive elicitation, gaps do exist.

4.1.4 KYLY’s revitalization philosophy

KYLY has had to handle gaps and inconsistencies in the linguistic documentation. Because the Tunica language revival uses a classroom-based model, KYLY has been guided by principles of morphological completeness and linguistic regularity

3. KYLY’s implementation of the CBI methodology has the instructors go through the following five steps: modeling the content of the lesson through presentation of a conversation; prompting students to interact with the material via physical response (responding to commands to touch or point, for instance); prompting students to answer yes/no questions related to the topic covered; asking questions about the content for which students can use known vocabulary as their responses; and asking questions that require the target vocabulary in response. For a more complete description of the methodology that KYLY uses, see Kuhpani Yoyani Luhchi Yoroni (2023:vi–ix).

Table 4.2: Tunica vowel blending chart

	i	e	ɛ	a	ɔ	o	u
i	i	e	ɛ	ɛ			u
e		e		ɛ			
ɛ			ɛ	ɛ			
a	i		ɛ	a			u
ɔ			ɔ	ɔ			
o			ɔ	ɔ		o	
u	i	o	ɔ	ɔ	ɔ	o	u

The left column represents the first vowel in the sequence, the top row represents the second vowel in the sequence, and the resultant vowel is at their intersection. Note that not every possible vowel combination is attested in Tunica, resulting in the empty cells.

in making decisions about the language. These principles aim to resolve issues in the documentation in ways that facilitate the teaching of Tunica.

Tunica phonological rules are complex. Syncope and partial assimilation are common features of the language. The group has kept some of these features, such as “vowel blending”. When a vowel-final morpheme precedes a vowel-initial morpheme, a system of partial assimilation and elision rules reduces the VV sequence into a single vowel. Table 4.2 shows the attested vowel sequences and how they resolve. Example 1 shows some instances from the corpus of vowel blending.

- (1) a. “uwin, tika rikini,” tinikɔni.
 uwi tika rikini tihk-ni-ku=ani
 he big overmuch her-say-3ms.HAB=QUOT
 “He is too big,” he told her. (4A:k)⁴

4. Where possible, I have used examples from documented Tunica texts. The notation 4A:k and similar notations refer to the location of the example text in Haas (1950). For instance, 4A:k refers to the fourth story, first telling, eleventh paragraph in *Tunica Texts*. Examples from Gatschet and Swanton are followed by citations for Gatschet’s field notes or Swanton’s texts, respectively.

- b. hinyatihch, ukik uyash'ekeni.
 hinyatihch, u-ki=ku uhk-yashi-aki=ani
 then his-uncle=3ms he-to.be.angry-3fs.INCHO=QUOT
 His uncle became angry. (4A:w)
- c. hiⁿyati nisara ili lot-uniχki.
 hinyatihch nisara ili lota-**u**nihki
 then youths two run-EXIST.3md
 Then two youths ran off. (Swanton 1910)
- d. **ε**stayi saχku shkaⁿ yolukyana namu **ε**rusa.
i-astayi sahku=shkan yoluyana namu **i**-erusa
 my-body one=but words many I-know
 I am only one person, but I know many words. (Gatschet 1886b)

In other instances, the group has reinserted phonetic segments that are synco-
 pated in the historical documentation. In documented Historical Tunica, for exam-
 ple, stem-final or word-final /hku/ tends to partially or completely elide in certain
 environments. One such environment is found in Tunica numbers greater than—but
 not multiples of—ten. Example 2 gives examples of Tunica numerals whose forms
 differ between Historical Tunica and Modern Tunica. Since [hku] deletes in its un-
 stressed position before *-teya*, and because many Tunica numbers end in /hku/,
 elision is very common in larger Tunica numbers. To make the underlying morphol-
 ogy more explicit, KYLY decided that the [hku] sequences in numbers—and in the
 language at large—would not elide in Modern Tunica. In addition to morphological
 completeness, this decision would make adding consonant-initial affixes more regu-
 lar, regardless of the word to which they attach. Thus, the decision also simplifies
 a feature of the language that was previously phonologically conditioned.

(2) Differences in Tunica numerals

Number	Historical Tunica	Modern Tunica
11	michu sahteya sahku	michu sahkuteya sahku
37	michu enihteya tsihku	michu enihkuteya tsihku

4.2 Competing ideologies in language revival

Despite efforts at prior ideological clarification at the outset of the project, competing ideologies are inevitable: “ideological clarification is not a one-time achievement but rather an ongoing process in need of periodic fine-tuning” (Kroskirty 2009:78). The turnover of the university side of a university-community partnership is high, so most members who came to the project through the university did so well after the inception of KYLY and were not involved when the ideologies and goals were initially established. The assistant language instructors joined the LCRP and KYLY as language apprentices in 2018, eight years after KYLY’s inception. While members of KYLY who are not based in Marksville are volunteers who work with Tunica as time allows or as their advice is needed, the assistant language instructors interact with Tunica on a daily basis as part of their job. They have brought fresh perspectives and their own ideologies to the group.

While KYLY has adapted as it has grown and developed, the group would benefit from a restatement of the overarching goals defined early in the project and a reevaluation of how best to achieve these goals given major intervening events like the creation of the LCRP and the addition of three assistant language instructors.

The competing ideologies inherent in any language revitalization effort push the revitalized or revived language in different directions. The competing ideologies within KYLY illustrate areas of contention within the language. Whitaker describes one of the most prominent: “There is always the question of what will be easy, accessible, and transparent versus what is authentic, justified, and true to the source material” (Whitaker 2017:97). Does ease of instruction win the day, or is adherence to attested forms the more important consideration, regardless of whether those forms fit the grammar rules that the group has developed? The group has not answered this question definitively and takes up these grammar issues individually at monthly videoconference meetings. Decisions can favor the ease of learning principle for one issue, and adherence to documented Tunica for another. The idiosyncratic results of the decisions create inconsistencies in Modern Tunica that differ from the inconsistencies in documented Historical Tunica. The group often resolves these new inconsistencies at a later meeting either by reversing course on the decision or by changing the principle that applies to that particular facet of the language.

Developing more precise agendas, implementing methods and procedures for deciding the answers to grammatical issues, and regularly reevaluating goals with all stakeholders would streamline the revitalization effort.

Another issue that has a bearing on language creation is pedagogy and power. Questions often arise as to how to present the language to learners. As Whitaker (2017) makes clear, the Tunica project is “top-down, with most planners coming from outside of the tribe, and most learners (mainly children) left out of decisions” (2).

For example, the New Tunica Dictionary (*Yanatame Nisa Luhchi Yoroni*) did not until recently indicate whether an entry is a neologism. Originally, this information was omitted because of concerns regarding the perceived authenticity of neologisms:

When neologisms are displayed in the dictionary alongside other Tunica words, their standing as “legitimate” language grows. For this reason, the lexicographer chose to place neologisms in the dictionary with no explicit label informing dictionary users of the word’s newness. Such a label could lead to these words being viewed as less authentic.

Anderson (2017:153)

Because attitudes of linguistic purism can have negative effects on revitalization and revival efforts (Dorian 1994), a decision that promotes the equal validity of all words makes sense. However, feedback from members of KYLY in Marksville has been largely supportive of labeling neologisms in the dictionary, resulting in a recent decision to do so. This process shows the tension between academic theory and desires of the community. Recent updates to the online dictionary reflect this modified ideology.

A third ideological issue concerns Modern Tunica’s status in the language and corpus planning process. Is Modern Tunica still undergoing an intensive language planning and corpus planning effort that will continue to change KYLY’s understanding of important aspects of the language and the way it should be taught? If so, should KYLY continue to take a lead role in making decisions about the language to ensure that the language can be taught in a cohesive and coherent way? Or should Tunica be regarded as a language with an active speech community, undergoing language change like any other living language, to be left alone as the speech community innovates?

The answer depends largely on the stances one takes on the ideological issues above. Apart from ideological issues, the practical concern of continued examination of the documentary materials argues in favor of KYLY continuing to steward the language as the examination of the data continues.

4.2.1 Continued reanalysis of documented Tunica

Because of KYLY's dependence on documentation, and because of the aforementioned inconsistencies and gaps in the data, the group has at times drawn incorrect conclusions due to lack of data or misreading of data. The best example of this is just how elusive the full, correct greeting sequence was.

When the group was just getting underway, the first lesson they created was a lesson on greetings. Swanton's sketch grammar of Tunica contained a greeting sequence:

The native salutation was :
e'ti ma, la'pū, MY FRIEND, ARE YOU WELL?
And the reply :
la'pū, maha't, WELL, AND YOU?

Swanton (1921:31)

This greeting sequence was put into the modern orthography as:

A: Etima lapun?
B: Lapu, mahat?

This greeting sequence was initially used when speaking to anyone. Soon after, however, when reading through the Haas grammar, KYLY found that while *mahat* should be used when speaking to a man, it is replaced by *hemat* when speaking to a woman.

And it was nearly ten years later that an examination of Haas's notebooks and her now-digitized but still challenging-to-interpret wax cylinder recordings revealed that *eti ma* (or *etima*, as it had come to be spelled) should become *eti hema* when addressing a woman.

Modern Tunica is changing not only due to the addition of neologisms, but also due to further study of documentary materials. Haas's 13 notebooks, Gatschet's field notebook, Swanton's handwritten texts, and the Gatschet-Swanton vocabulary cards have not yet been exhaustively examined.⁵

4.3 Conclusion

KYLY, in its current form as a collaboration between the Tunica-Biloxi Tribe and Tulane University, has been working on reviving Tunica for over a decade. In that time, it has made great strides. The LCRP provides additional organizational capacity and tribe-based support to further the community's goals. KYLY and the LCRP have created an impressive pedagogical base for teaching Tunica in the Tunica-Biloxi community and beyond. However, the group must regularly ensure that the goals and philosophies underpinning the project serve the community's needs and desires and that issues of conflicting ideologies are addressed and resolved.

Modern Tunica will necessarily differ from documented Tunica. Its method of transmission is different, it exists in a different time and environment, and it must

5. There is currently a project underway, led by Patricia Anderson in collaboration with the American Philosophical Society (APS), that is transcribing the Haas notebooks. KYLY is also in the process of transcribing the Gatschet-Swanton vocabulary cards.

serve the needs of the community as it currently exists. Further study of documented Tunica will only help to ensure that Modern Tunica serves its current time, place, and environment while maintaining close ties to the language as it was spoken.

Chapter 5: Methodology

This chapter describes the principles used in determining what to include in the corpus, how the corpus was digitized, and the methods designed for automating frequency analysis of the corpus.

5.1 Corpus creation

To create the corpus in this work, all written and transcribed Tunica was evaluated to determine what texts were appropriate for inclusion.

5.1.1 Documented Historical Tunica

The entirety of documented Historical Tunica consists of the work of three speakers—William Ely Johnson, Volsin Chiki, and Sesostrie Youchigant—as documented by Albert Gatschet, John Swanton, and Mary Haas.¹ The sources of Historical Tunica are listed in Table 5.1.

For a text to be included in the corpus, it must meet two criteria. First, the text must be sufficiently analyzed by the linguist who transcribed it. Secondly, the text should be extemporaneous rather than scripted. These criteria ensure that

1. Section 2.3 contains brief biographies of each of these individuals.

Table 5.1: Sources of Historical Tunica

Corpus	Main sources	Years collected
Gatschet	Field notebook	1886
Swanton	Vocabulary cards Texts	1907–1910
Haas	Field notebooks Tunica Texts	1933–1939

the transcriptions of the texts are as error-free as possible and that the speech transcribed is natural speech.

The corpus in this work omits data gathered through translation exercises and word lists. This eliminates the elicitation portions of Gatschet and Haas’s field notebooks, as well as Swanton’s vocabulary cards. Eliminating these segments of the documentation avoids the differences in phonology that occur between elicited vocabulary and connected speech and focuses the corpus on “naturalistic” speech rather than elicitation; in elicitation contexts, words are often gathered via a common language and thus may be more susceptible to direct translation.²

Gatschet’s 1886 field notebook contains eight interlinear texts from his work with William Ely Johnson. Both Swanton and Haas would later have access to these materials, making their own copies and, in Haas’s case, annotating them with changes in transcription conventions and reanalysis.

Swanton’s work consists of two typed stories gathered from Volsin Chiki in 1910, as well as stories gathered in 1930 and 1931 from Sesostrie Youchigant. Swanton

2. We know that Gatschet and William Ely Johnson communicated in French, while Swanton most likely communicated in English with Volsin Chiki. Haas communicated mostly in English with Sesostrie Youchigant.

Table 5.2: All transcribed Historical Tunica texts

Linguist	Speaker(s)	Year(s)	Number of texts
Albert S. Gatschet	William Ely Johnson	1886	8
John Swanton	Volsin Chiki	1910	2
John Swanton	Sesostrie Youchigant	1930	~10
Mary Haas	Sesostrie Youchigant	1933–1939	76

typed all the texts Gatschet had collected—and the two stories he himself had obtained from Volsin Chiki—and forwarded them to Haas when she was working with Sesostrie Youchigant. As she had done with Gatschet’s texts, Haas annotated the Swanton texts as well. Swanton did not forward any of the material he had gathered from Sesostrie Youchigant in 1930 or 1931 to Haas.

Mary Haas and Sesostrie Youchigant’s work from 1933 to 1939 resulted in 76 unelicited texts, ranging from the epic Thunder Myth to the two-sentence Star People. During the time he worked with Haas, Sesostrie Youchigant told several of the stories multiple times, with both Haas and Sesostrie Youchigant saying that his facility with the language improved over time, making each telling of a story better than the previous attempt (Haas 1950:9).

Table 5.2 lists all the known sources of documentation of unelicited Historical Tunica texts. Some of the texts that Swanton transcribed are written in his unclear handwriting. These handwritten texts appear to be from fieldwork with Sesostrie Youchigant in 1930.

All documented Historical Tunica texts are included in the corpus except for the texts Swanton gathered from Sesostrie Youchigant. That data has been omitted

because most of it is in handwritten form and has not yet been digitized. The two works that were typed may have been works in progress; in the winter of 1933–1934, Swanton sent Mary Haas typed copies of Albert Gatschet’s work with William Ely Johnson, along with two texts he had gathered from Volsin Chiki in 1910. He did not send her any of the material that he had gathered in 1930 or 1931. Significantly, he placed a note on the texts he had collected in 1930: “Material probably made unnecessary by work of Mary Haas” (Swanton 1930a).

5.1.2 Modern Tunica texts

Published documentation of Modern Tunica, listed in Table 5.3, is growing alongside its emergent speaker base. Due to the nature of revival, much of the Tunica language that has been published is didactic; texts and dialogues created for the Tunica textbook and recordings of classroom instruction constitute the vast majority of it. There are also a number of instructional videos of Tunica material culture with Tunica narration and Tunica/English subtitles. Lastly, there is a 2017 recording of a conversation between Donna Pierite and Elisabeth Pierite, two Tunica-Biloxi singers and legend keepers, and at the time the only two Tunica language instructors (Wikitongues 2018).³

3. More analysis of Modern Tunica speech will soon be available. Harvey (2022) describes a methodology for gathering Modern Tunica speech remotely that “produces clear recordings that can be used for revitalization goals and linguistic analysis” (1). The recordings are of word lists, translation exercises, conversations, and teacher training lessons. Harvey (2023) discusses the implementation and results of these recordings of Modern Tunica. Harvey’s work will drastically increase the amount of recorded Modern Tunica.

Table 5.3: Sources of Modern Tunica texts

Type	Number of texts
Textbook material	10 texts ~130 dialogues
Classroom recordings	~100
Instructional cultural videos	~15
Conversations	1

Despite the large number of written texts in Modern Tunica, very little of the published material captures Tunica spoken extemporaneously. The texts and dialogues in the Tunica language textbook, *Rowinataworu Luhchi Yoroni* (2023), were all subject to multiple rounds of editing by Kuhpani Yoyani Luhchi Yoroni (KYLY) to ensure that they conformed to the group’s decisions about how the language should be taught and written. The recordings of Tunica classroom lessons are similarly didactic in nature. The instructional cultural videos in Tunica are less linguistically didactic; the primary purpose of these videos is to teach how to create cultural objects, e.g., a traditional Tunica dress or a pine needle basket. However, the narration is scripted beforehand. The only example of recorded and published extemporaneous Tunica speech is a 2017 conversation between Donna Pierite and Elisabeth Pierite, two Tunica-Biloxi singers and legend keepers, and at the time the only two Tunica language instructors (Wikitongues 2018). Though a portion of the video is in English, most of it is strictly in Tunica.

Because the textbook material is the result of several rounds of edits by KYLY and is not extemporaneous speech, it has been omitted from the corpus. The classroom recordings have also been omitted because switching between Tunica

and English results in a lack of continuous Tunica speech of appreciable length. The instructional videos are entirely in Tunica but are scripted beforehand and often recited in voiceover. Unfortunately, only the conversational video satisfies the criteria of extemporaneous Tunica speech of sufficient length, and thus is the only example of Modern Tunica included in the corpus.

5.1.3 The corpus

The decisions in the previous sections created a corpus that includes only examples of extemporaneous connected speech that the transcribing linguist has sufficiently analyzed. Table 5.4 contains the number of texts from each time frame that fit the criteria for for the corpus. The next step was the digitizing and parsing of these texts.

Table 5.4: Corpus texts

Date	Name	Texts
1886	Gatschet	8
1910	Swanton	2
1933–1939	Haas	76
2017	Modern Tunica	1
Total		87

5.1.4 Digitizing the corpus

KYLY chose Summer Institute in Linguistics’s (SIL) Fieldworks Language Explorer (FLEx) software to serve as the digital database for the revitalization project mostly due to community familiarity with the software. The benefits and drawbacks of

FLEx and the method through which the corpus was digitized are explained in the following sections.

FLEx

Beginning in 2013, the KYLY began using FLEx to digitize Mary Haas's *Tunica Dictionary* (1953).⁴ FLEx is an expansion of SIL's earlier Field Linguist's Toolbox program, which was designed to facilitate gathering and analyzing elicitation and texts in the field. FLEx contains various tools for linguistic analysis organized around the linguistic documentation trio of a dictionary, a grammar, and a compilation of texts. These three areas inform the three main sections of the FLEx interface: Lexicon, Grammar, and Texts.

Because FLEx data is stored in a database, the interlinear texts can reference lexicon entries for their morphemic glosses. This, in turn, allows for a concordance that deals elegantly with homonyms and can search for correctly tagged morphemes or words based not just on their orthographies, but their syntactic and semantic attributes as well. FLEx's built-in XAmple parser enables automatic parsing of wordforms that have already been analyzed manually once.

FLEx interlinear texts can include a number of tiers, and KYLY has used eight of them. The information in these tiers was not always complete when work on this dissertation started, necessitating the filling in of any information lacking from the tiers in use. The tiers used are listed on the following page and shown in Figure 5.1.

4. The making of the New Tunica dictionary and the principles that informed its creation are described in detail in Anderson (2020).

Figure 5.1: An interlinear sentence in FLEx

1.4 Word	Hinyatihch	, tawakaku			, kal'uhkeni		
Morphemes	hinyatihch	ta-	waka	-ku	kali	-uhki	-ani
Lex. Entries	hinyatihch	ta ₋₂	waka ₁	-ku ₁	kali ₁	-uhki ₂	-ani ₁
Lex. Gloss	now (conj.)	AGNZ	to command	3ms	to stand	3MS.completive.uhki	QUOT, "it is said"
Lex. Gram. Info.	conn	v>n	v.I.intr	n:(GNS)	v.I.intr	v:Subject	v:(Quotative)
Word Gloss	now	the commander			he was standing, it is said		
Word Cat.	conn	n			v.I.intr		

Free Now the commander was standing, it is said.

1. Word
2. Morpheme
3. Canonical form of the morpheme as defined in the lexicon
4. Morpheme gloss
5. Grammatical information (e.g., parts of speech and types of affix)
6. English word gloss
7. Word category (part of speech)
8. Free translation

For languages with complex sandhi rules, it can be difficult to set up FLEx's lexicon and grammar to account for the full set of phonological processes that take place at morpheme and word boundaries. The XAmple parser requires entering all possible allomorphs for forms that change due to sandhi. For more complex processes, like the vowel blending seen in Tunica, this requires a large set of allomorphs for both stems and affixes that may help the parser to parse but is laborious and obscures the systematic nature of the phenomenon.

FLEx also includes HermitCrab, an experimental parser that uses a set of user-defined phonological rules to determine valid wordforms and to parse words without the need for a large set of allomorphs where the phonological changes are systematic. However, this parser is much slower. The FLEx manual cautions that the experimental parser should only be used when necessary and that the default parser should be used wherever possible.

FLEx does not have a suprasegmental tier for information like stress and tone. For Tunica, this omission prevents the inclusion of two critical aspects of Tunica prosody. Tunica stress is mostly stem-initial. As it differs from English stress patterns, Tunica language learners often struggle with stress placement. Including stress information in FLEx would ensure that any language learner exploring the Tunica texts would clearly understand where the stress lies. Tunica also uses phrase-final tones. Haas marked these diligently but they have not been studied in depth. Without a tier for tone, analysis of its function and distribution becomes much more challenging.

Beyond the purely linguistic drawbacks of FLEx, there is also the issue of accessibility for the community. The learning curve for FLEx is steep, and while exporting the dictionary is trivial with SIL's Webonary service, providing access to the interlinear texts requires a third-party solution.⁵ FLEx enables collaboration by allowing storage of the database on their LanguageDepot platform, from which it can be synchronized to different computers. However, editing grammatical or textual information in FLEx requires running the program on a Windows or Linux computer; texts cannot be edited via the web,⁶ and FLEx can only be run on a Mac using Windows or Linux virtualization software.

5. As of writing, KYLY has not chosen a platform for accessible viewing of Tunica Texts, but have considered LingView (Pride, Tomlin, and AnderBois 2020) for online viewing. A reformatted side-by-side, sentence-by-sentence version of Haas's Tunica Texts (1950) has been created as part of this work that could serve as a model for print versions of all Tunica texts.

6. SIL does offer access to lexicon editing via the web using languageforge.org.

Despite these drawbacks, FLEx's integration of texts, grammar, and lexicon; its powerful parsing tools and software ecosystem; and the fact that much of the Tunica language lexical and textual information was already in a FLEx instance made it the best choice for the current study.

Configuring the digitized corpus

Starting in 2013, members of KYLY began transcribing the texts contained in Haas's *Tunica Texts* (1950). Over time, most of the Haas texts and one or two of the Gatschet and Swanton texts were transcribed, added to FLEx, and partially parsed and analyzed. While it seemed like the groundwork was set to perform any number of analyses on the language, the data were still far from ready.

Haas's published texts use the phrase rather than the word as the unit of analysis due to the phrasal nature of Tunica prosody. As such, the texts contain no indication of word boundaries. Because FLEx's parsing and other tools work best when words are the unit of analysis, transcribers not only created word breaks, which were necessary and could be created without losing any information contained in the original, but also made other changes to standardize the orthography and phonological rules adapted for Modern Tunica.

One of the main principles guiding the creation of Modern Tunica is morphological transparency: in order to make morphemes more easily recognizable to new language learners, some segments that were elided in the Haas texts due to Historical Tunica phonological rules were restored by the transcribers when the texts were first entered into FLEx. While this serves the pedagogical goals of Modern

Tunica, it can hinder the phonetic and phonological study of Historical Tunica or lose evidence of stylistic variation. As a result, the texts transcribed in FLE_x could be used as a starting point but they had to be rechecked against the originals to restore a more accurate representation of the texts as they had been spoken.

Another issue that needed correction in the FLE_x transcriptions involved the tagging of several high-frequency morphemes that are homonymous in Tunica. For example, the same set of prefixes is used to express a verbal object for active verbs, a verbal subject for stative verbs, and nominal possession. Many of these homonymous morphemes were tagged incorrectly due to mistakes in manual tagging or incorrect guesses by the automatic parser.

Orthographic standardization is an important part of corpus creation, as it allows for direct comparisons. Gatschet used different notational conventions than Haas, for instance, but most of these are one-to-one correspondences that can be trivially unified. Other orthographic differences are indicative not of transcription conventions but of actual differences in speech. Determining which orthographical differences are meaningful and which are mere changes of convention required an analysis of the primary documents.

One area where the discrepancies are not merely notational or easily unified is in intervocalic stops. Gatschet's texts are replete with voiced stops ([b], [d], [g]) in intervocalic environments, whereas Swanton's transcriptions of Volsin Chiki's speech contain the much more occasional intervocalic voiced stop. Haas's transcriptions of Sesostrie Youchigant contain no voiced stops outside of loanwords and one other lexical item (*-gachi*, meaning 'mother'). It is possible that each speaker voiced

his stops less than the one who came before him, or that this trend results from changing transcription conventions over time.

Unlike Gatschet's and Swanton's texts, Haas's texts use phonemic rather than phonetic transcription. Haas explains how her phonemic transcription differs from Gatschet's and Swanton's transcriptions in the introduction of the *Tunica Dictionary*:

Other changes consist in a phonemicization of the forms quoted by [Gatschet and Swanton]. Thus, <h> is written in place of <x>, <n> is written in place of nasalization, and <ʔ> is written wherever it is known to occur even though both Gatschet and Swanton usually failed to record it.⁷

Haas (1953:180)

As the forms in the dictionary correspond to the forms in *Tunica Texts*, and as there are no occurrences of <x> in either *Tunica Texts* or *Tunica Dictionary*, it is clear that both the dictionary and Haas's texts are strictly phonemic.

Though she describes the phonological rules that affect these morphemes in her *Tunica* grammar (Haas 1940:18), matching phonetic transcriptions to a phonemic transcription makes any phonetic analysis more difficult.

With these discrepancies and difficulties, it was essential to ensure that corpus data was consistent and accurate. Appropriate orthographic changes were made to texts that used different transcription conventions with one-to-one correspondences. Word breaks were added to Haas texts while attempting to keep as close to the original transcription as possible.

7. Swanton and Haas transcribe Gatschet's <χ> as <x>. Haas (1940:18) states that the phonetic production of /h/, which she uses in her published texts to the exclusion of <x>, surfaces as either [h] or [x].

Table 5.5: Transcription differences

IPA	[χ], [x], [h]	[̃]	[æ], [ɛ] ⁹	[tʃ]	[ʃ]	[ɔ]	[ʔ]
Gatschet	<χ>, <h>	< ⁿ >	<ä>	<tch>	<sh>	N/A	N/A
Swanton	<x>, <h>	< _ε >	<ä>	<tc>	<c>	N/A	N/A
Haas	<h>	<n>	<ɛ>	<č>	<š>	<ɔ>	<ʔ>
Modern Tunica	<h>	<n>	<ɛ>	<ch>	<sh>	<ɔ>	<'>

FLEx supports allomorphs and variants, and both functions were used to aid parsing, to preserve variations from the canonical forms of each morpheme or word. Table 5.5 shows the transcription changes between documenters, as well as the forms that Modern Tunica uses.

The corpus used in this dissertation follows Modern Tunica orthographic conventions⁸ where there is a one-to-one (or zero-to-one) correspondence between all historical transcriptions (<ɛ>, <ch>, <sh>, <ɔ>, and <'>). To preserve phonetic information in the Gatschet and Swanton texts, <χ> and <x> have been preserved where they appear, with these being normalized to <h> for the underlying lexical entry. Similarly, nasalization in Gatschet and Swanton (<ⁿ> and <_ε>, respectively) have been unified to <n> and preserved in Gatschet and Swanton texts, with the underlying /n/ noted as the canonical form.

8. Modern Tunica orthography differs from the texts in Haas (1950) in the removal of word-initial glottal stops, the removal of stress and tone marks, and the addition of spaces to mark word boundaries.

9. [æ] and [ɛ] refer to a phoneme that, according to Haas, falls somewhere between the two in the vowel space: “*ɛ* is a low vowel, slightly closer than the *a* of Eng. *mat* but not so close as the *e* of Eng. *met*” (Haas 1940:15).

5.2 Automating frequency analysis

As noted in Section 5.1.4, FLE_x and the ecosystem of auxiliary programs around it provide a wide range of analytical tools for linguistic analysis. However, when looking at diachronic and synchronic differences in language use through frequency analysis, some tasks are best handled by bespoke tools that work with a FLE_x database exported into a more standard format. To automate the analysis of a given morphological or syntactic phenomenon, the interlinear texts in the FLE_x's database were exported into a generic eXtensible Markup Language (XML) file. A Python script was then created to iterate through the file to examine verb aspect distribution.

For all other features examined, R (R Core Team 2022) was used with the `interlineaR` package (Loiseau 2018) to import and analyze the interlinear texts and lexicon that FLE_x can export into generic XML format and Lexicon Interchange FormaT (LIFT), respectively.

Having both the lexicon and the corpus of texts as data objects in R enabled the merging of richer lexical information like the semantic domain of content morphemes into the interlinear texts.

Automating this process required that the texts have a normalized orthography and morphological parsing that was consistent, accurate, and complete. Though the process has been painstaking and long, the benefits to both the tribal and academic communities in improving the accessibility and analysis of the documented language have made the effort worthwhile.

Chapter 6: Areas of study

To determine what aspects of the language to focus on, four criteria that might suggest an area of language change were developed:

1. Phenomena that speakers or documenters remark upon as having been different in the past
2. Phenomena that appear unstable within a single speaker's speech
3. Phenomena where change is evident between speakers or between time periods
4. Phenomena that have not been fully explained in any existing grammar

Several phenomena fit these criteria. The phenomena that were selected cover multiple levels of linguistic structure and are all phenomena that have been subject of discussion and questions in Kuhpani Yoyani Luhchi Yoroni (KYLY) meetings. Thus, whether these phenomena exhibit clear evidence of language change or not, the information about how they operated in Historical and Modern Tunica will help inform future decisions as the revitalization effort continues.

All phenomena that meet the first criterion are included in the analysis. The other phenomena chosen for study are the verb aspect system, the Tunica gender system, the quotative evidential, elision in possessive noun prefixes, and Tunica word order.

Each phenomenon introduced in this chapter will be analyzed in Chapter 7.

6.1 Phenomena noted by documenters and speakers

There are a few instances where either a speaker of Tunica or a documenter of Tunica notes a feature of the language that they believe has undergone a change. These instances are worth verifying.

Gatschet does not mention any specific area where language change was noted by a speaker or indicated by the data. Swanton's sketch grammar of Tunica (1921), however, proposes a verbal morpheme, /-pa/ or /-p/, that he thought might have previously been independent and productive but was by that time fully incorporated into the verb stem. Swanton posits that this archaic suffix is a causativizer.¹ His main reason for focusing on these particular morphemes is his observation that /p/ occurs much more frequently in verbs than would be expected (13–14). Though he cites few examples of the verbs he believes contain these morphemes, Swanton's observations warrant a brief look at whether the frequencies he observed in his own data hold in the Fieldworks Language Explorer (FLEX) lexicon that contains lexical entries from the work of all three linguists.

Haas also lists morphemes she suspects to have once been productive but which are now preserved only in certain words. Haas also mentions two specific instances of diachronic language change. In 1933, early in her documentation of Tunica, Haas sent a letter to Edward Sapir with her initial impressions of the language. The letter

1. Swanton also proposes /u/ as a former nominalizing suffix, but he is highly uncertain about its existence; it has been omitted from analysis here.

includes a note about the gender-number affixes (GNAs) that attach to nouns or noun phrases.

This is done by suffixing *-k* (or *-ku* or *-ky*, all from older *-kuhy*) to masculine nouns and *-c* or *-htc* (from older *-htci* ?).² (I give the older forms on the word of my informant, who says the old folks used to talk that way. He seems to be much more familiar with *-kuhy* than with *-htci*, hence the question in regard to the latter.)

Haas (1933)³

Haas would later revise her interpretation of these morphemes as “carry[ing] the implication of selectivity. Hence, *to'niku'hu* may mean ‘it is the man who’ as opposed to *to'niku* which means simply ‘the man’” (Haas 1940:65). In current terms, *=kuhu* and *=hchihhi* serve as Tunica analogs to English *it*-clefts. Because the speaker understood these forms to have been more common in the past, they are worth studying.

The second mention of diachronic change concerns the conjugation of non-volitional verbs. “[T]here is evidence that stems denoting involuntary action (e.g., ‘to breathe’; ‘to cough’) were formerly used as transimpersonals. The more usual procedure now is to treat such stems as intransitives” (59). Transimpersonal verbs have an unspecified agent represented syntactically by a third-person feminine singular subject marking. The change from transimpersonal to intransitive would represent a change from a marked structure to an unmarked structure with analogs in French and English. Texts will be compared to see if verbs of non-volitional

2. Haas (1946) lists the alternate form of *-htc* as *-hčihhi*

3. Haas’s transcription conventions are left in place here. The ogonek <_> marks nasality, and <c> corresponds to /f/. See Table 5.5 for non-IPA orthographic correspondences for all documenters.

action move from being conjugated as transimpersonals to intransitive verbs over time.

Several phenomena show variation within a single speaker's speech. Among these phenomena is the grammatical gender system that so interested all three documenters. Haas notes that Sesostrie Youchigant was inconsistent in his use of gender marking: "My informant is rather uncertain about the gender of some things, using the masculine suffix at times in the feminine at others" (Haas 1933).

Features like GNAs show differences in usage between speakers: William Ely Johnson frequently used shortened forms of GNAs (e.g., *-sin*), whereas Sesostrie Youchigant more often used the longer form (e.g., *-sinima*).

Other aspects of the language are not fully explained by any linguist, either those who documented the language or those involved with the revitalization effort. Among these are a host of distributional questions about when to use certain synonymous morphemes. For instance, Haas categorizes four morphemes as imperatives but writes, "[t]hough it seems likely that there is some affective difference between *-n* and the other three imperative postfixes (*[-ki, -hchan, and -tan]*, below) it has not been possible to ascertain just what the difference is" (Haas 1940:118).

Using the other three criteria, the following areas were chosen:

1. Verb aspect distribution
2. Elision of *[-hk]* in noun possession
3. Gender marking
4. Prevalence of quotative evidential =*ani* (it is said)
5. Word order distribution

Table 6.1: Tunica auxiliary verbs

Verb	Independent meaning	Inflectional meaning
uhki	to be, to exist	<i>perfect aspect</i> or priority in time
ura	to lie	<i>durative aspect</i>
una	to sit	<i>durative aspect</i>
usa	to come	venitive (to come while [verb]ing)
uwa	to go	andative (to go while [verb]ing)

6.2 Verb aspect

Tunica has a rich verbal morphology and aspectual system, making the Tunica verb complex a particularly appealing feature to explore. The linguists documenting Tunica noted the different inflections and described them to varying degrees. Haas's description is the most complete and detailed, but questions remain about when to use one aspect over another. Haas's translations and descriptions of Tunica aspect detail multiple possible English equivalents for some Tunica verbal inflections, which is the case in many translations but is especially true when translating from a primarily aspectual language (Tunica) to a primarily tense language (English). Context and the semantic properties of the stem verb will affect the interpretation of the inflected form in these cases.

Haas divides Tunica verbs into three categories: auxiliary verbs, active verbs, and static verbs (Haas 1940:40).⁴ The following sections cover each of these in turn, and the questions that remain about the distribution of verb aspects in documented Tunica.

4. Haas's terminology differs from current linguistic terms. Her terms are preserved in this dissertation, with two exceptions: what she calls semelfactive is termed the completive, and static verbs are referred to as stative verbs.

6.2.1 Auxiliary verbs

Tunica has five auxiliary verbs.⁵ These verbs can appear as independent verbs or be used to inflect another verb (Haas 1940:41). In inflectional use, three auxiliaries lose much or all of their semantic content and instead serve to indicate aspect. Table 6.1 lists the auxiliary verbs with their meanings as independent verbs and in their inflectional forms.

The meanings and usage of the inflectional forms are particularly nuanced and at times unclear; Haas refers to the “ideological complexities” that arise when auxiliaries are used inflectionally (41).

The following examples show the differences in meaning between the five auxiliary verbs’ independent and inflectional forms.⁶

5. Haas categorized the causative suffix as a sixth auxiliary verb. However, one of her main criteria for auxiliary verbs is their ability to operate as independent verbs, whereas the causative suffix is always bound. Thus, causative verbs are classified here as a second class of non-auxiliary active verb inflection, Class II verbs. These verbs and their use are described further in Section 6.2.2.

6. The glossing conventions used are as follows: 1 = first person, 2 = second person, 3 = third person, AGNZ = agentive nominalizer, AL = alienable, CAUS = causative, CLASSII = class II verb inflection, COMP = complementizer, COMPL = completive, CONJ = conjunction, d = dual, EXIST = existential, f = feminine, FNL = phrase-final suffix, FUT = future, HAB = habitual, IMP = imperative, INAL = inalienable, INCHO = inchoative, m = masculine, NEG = negative, p = plural, POSS = possessive, QUOT = quotative, RECP = reciprocal, REDUP = reduplication, REL = relative, s = singular, SBJV = subjunctive, SUBR = subordinator.

- (3) uhki
- a. tawakaku u'uhkəni
 ta-waka=ku u-**uhki**=ani
 AGNZ-command=3ms REDUP-**EXIST.3ms**=QUOT
- There used to be a commander. (4A:a)⁷
- b. yak'ahkini
 yaka-**ahkini**
 come-**be.1s**
- I have come. (7A:e)

In 3a, *uhki* denotes existence, with the reduplication indicating the remote past. When juxtaposed against another verb, as in 3b, Haas translates the verb with a perfect aspect. Inflectional *uhki* need not necessarily be perfect, but for Haas, all verbs translated into English as perfect have inflectional *uhki* in Tunica.

The auxiliaries *una* and *ura* are even more ideologically complex, to use Haas's term.

- (4) una
- a. uwin, hɔwash unani.
 uwi hɔwashi **una**=ani.
 3ms outside **sit.3ms**=QUOT
- He was sitting outside. (4A:d)
- b. pohkunani.
 po-hk-**una**=ani
 see-HAB-**sit.3ms**=QUOT
- He was watching. (4A:d)

7. As mentioned in footnote 4 of Chapter 4 on page 44, I have used examples from documented Tunica texts where possible. The notation 4A:a and similar notations refer to the location of the example text in Haas (1950). For instance, 4A:a refers to the fourth story, first telling, first paragraph in *Tunica Texts*.

- c. sehiyuru pitahkunani.
 sehi-yuru pita-hk-**una**=ani
 morning-long walk-HAB-**sit.3ms**=QUOT

He walked all day long. (4A:b)

(5) ura

- a. hinyatihch, nehtalinah pokatēni, katan on ilin,
 hinyatihch nehtali-nahku po-kati=ani kata oni ili
 then bed-like look-3fs=QUOT where person two

uranahch

urana=hch

lie.3md=SUBR

Then, (it) looked like a bed where two people were lying. (10:h)

- b. awachihk'uranani.
 a-wachi-hk-**urana**=ani
 RECP-fight-HAB-**lie.3md**=QUOT

They (kept) fighting each other. (4A:h)

In 4a, where *una* is independent, it simply means to sit. 4b shows a juxtaposed *una* whose translation implies that the semantics of sitting is completely absent. However, 4a, which explicitly places the agent in a seated position, immediately precedes 4b in the text, so a more precise reading of 4b might be “He was sitting watching”. In 4c, on the other hand, *una* does indeed lose its semantic meaning, as sitting is inherently incompatible with walking. Haas interprets this use of *una* as a durative.

The same pattern is found with *ura*, where the independent form (5a) denotes lying, while Haas translates 5b as a durative. In the case of 5b, the presence of the habitual affix *-hk* makes the verb a mutable durative (see also 4c). Haas considers verbs of the same form without the *-hk* morpheme immutable duratives.

Haas notes that the difference between mutable and immutable duratives is hard to express in English:

Both may be rendered by the English present or past progressive. For example, a sentence like “the man is standing” requires the Tunica immutable durative while a sentence like “the man is running” requires the Tunica mutable durative. However, the immutable durative may also be rendered by the English simple present or past, as in “the man stands, stood,” or by such an expression as “remains (or stays), remained (or stayed) doing ...” The mutable durative, on the other hand, is preferably rendered by the present or past progressive.

Haas (1940:50)

If the immutable durative can—and often does—express the same semantic idea as the mutable durative, another variable might be dispositive or at least highly correlated with choosing one form over the other. The meaning of the verb itself, and the nature of the action it signifies (e.g., volitional or non-volitional), could be at play.

The distribution of *una* as opposed to *ura* in forming duratives is also something that Haas cannot define precisely: “It should be noted that whether one uses [ura] or [una] in expressing the durative is largely a matter of idiom” (50). She posits that where animates are concerned, the choice may depend both on positionality (semantic content of the auxiliaries) and gender (*una* for masculine animates in a vertical position, and *ura* for feminine animates in a vertical position).

- (6) hipuhkurani.
 hipu-hk-**ura**=ani
 dance-HAB-**lie.3ms**=QUOT
 (The alligator) was dancing, it is said. (14A:f)

In 6, the gender of the male alligator would, according to Haas, privilege *una* over *ura*, but his inherently horizontal position requires that *ura* be used. Positionality

trumps gender for animates with inherent position, or when position is essential to the story.

The auxiliaries *usa* and *uwa* retain all of their semantic meaning when serving as an auxiliary, simply specifying the direction of the agent at the time of the action.

(7) *usa*

- a. *teheyak* *usak'ahcha,* *he'esh*
ti-ahaya=ku ***usa-k'ahcha*** *he'esh*
 her-brother=3ms **come.3ms-FUT** today
 Her brother will come (back) today. (4B:br)

- b. *tasiwaku* *harahkusani.*
tasiwa=ku *hara-hk-**usa**=ani*
 owl.mammoth=3ms sing-HAB-**come.3ms**=QUOT
 The owl mammoth was coming singing, it is said. (10:h)

(8) *uwa*

- a. *hinyatihch,* *uwanani.*
hinyatihch ***uwana*=ani.**
 then **go.3md**=QUOT
 Then they went on. (10:f)
- b. *hinyatihch,* *uhchuwihch,* *uhker'uwani.*
hinyatihch *uhk-chu-wi=hch* *uhk-eru-**uwa**=ani*
 then *him-take-3ms=SUBR* *him-carry-**go.3ms**=QUOT*
 Then, when he took him, he went carrying him, it is said. (14A:d)

Because of the similarities between internal and external sandhi in Tunica, it is hard to determine at first glance whether *usa* and *uwa* are serving as auxiliaries or simply as part of a serial verb construction in 7b and 8b. However, Haas's diagnostics for auxiliaries—including suppletion, ability to stand alone, and ability to take reduplication—led her to analyze them as auxiliaries. Additionally, *usa* and *uwa* never immediately precede another verb in the same clause, nor do non-

Table 6.2: Class I completive

Person	SG	DU	PL
1	-ni	-ina	-iti
2m	-i	-wina	-witi
2f	-a	-hina	-hiti
3m	-wi	-una	-ta
3f	-ti	-sina	-siti

Table 6.3: Class II completive

Person	SG	DU	PL
1	-htani	-inta	-inta
2m	-wita	-winta	-winta
2f	-heta	-henta	-henta
3m	-uta	-unta	-anta
3f	-ata	-sinta	-sinta

subordinated, fully-inflected verbs appear immediately before them. These lend support for their classification as auxiliaries.

6.2.2 Completive aspect

The completive aspect is indicated through one of two sets of verbal inflection, depending on the class of the verb. The far more common form, which Haas calls “non-causative”, is referred to here as Class I. Table 6.2 shows the Class I completive verb paradigm.

Class II, which Haas calls the causative auxiliary,⁸ uses the set of completive suffixes in Table 6.3. Most Class II verbs have a causative reading and many have a non-causative Class I counterpart, as seen in 9. However, some verbs are strictly Class II with no apparent causative reading, as in 10: *uru*, ‘to whoop; to shout’, is intransitive and has no Class I equivalent (Haas 1940:46).

Haas notes that the completive is punctual, and is most often translated into English as the simple past (as in 9a) or sometimes as the past perfect (as in 9b) (39).

8. See footnote 5 on page 71 for an explanation of why this work does not categorize this aspect as an auxiliary but rather as a non-auxiliary active verb inflection.

Table 6.4: Class I habitual⁹

Person	SG	DU	PL
1	-kani	-hk'ina	-hk'iti
2m	-ki	-wina	-witi
2f	-ka	-hina	-hiti
3m	-ku	-hk'una	-kata
3f	-kati	-sina	-siti

Table 6.5: Class II habitual⁹

Person	SG	DU	PL
1	-hk'atani	-hk'intā	-hk'intā
2m	-wita	-wintā	-wintā
2f	-heta	-henta	-henta
3m	-hk'uta	-hk'unta	-hk'anta
3f	-hk'ata	-sinta	-sinta

- (9) a. hinyatihch, hatikan, haraweni.
 hinyatihch hatika hara-**wi**=ani
 then again sing-**3ms.COMPL=QUOT**
 Then he sang again. (4B:bh)
- b. nisalatihch, uhtaharanch,
 nisa-la-ti=hch uh-taharani=hchi
 new-set-3fs.COMPL=SUBR his-fiddle=3fs
 har'utahch, toniseman, hipuna'arani.
 hara-**uta**=hch ta-oni=sema hipu-na'ara=ani
 sing-**CLASSII.CAUS=SUBR** the-person=3mp dance-LIE.3mp=QUOT
 When it was nearly night, he played his fiddle (while) the people were dancing, it is said. (14A:i)
- Lit:* When the sun had just set, when he caused his fiddle to sing, the people were dancing, it is said.
- (10) hinahkuhch, sahu^{ət}, ur'utani.
 hinahkuhch sahu=ku=hat uru-**uta**=ani
 then other=3ms=on.one's.part shout-**CLASSII.3ms=QUOT**
 Now another, on his part, shouted, it is said. (2:b)

6.2.3 Habitual aspect

The habitual aspect is largely based on the completive endings in the last section.

Class I and Class II habitual forms are in Tables 6.4 and 6.5.

9. Haas posits that the dual and plural forms of Class I habituals—and all Class II habituals—start with /hk/, but that phonological rules delete [hk] before continuants, and keep them before glottal stops (Haas 1940:47–48). In Modern Tunica, the glottal stops have been dropped and [hk] prepended

Haas refers to the habitual aspect as “either a seriated punctual or a linear aspect, i.e., it implies either that the action is habitually performed or that it is in a state of progress”. Depending on context, Haas translates habituals as the simple present (11a), simple past (*nikateni* in 11b), present progressive (*eruku* in 11b), and past progressive (11c and 12) (Haas 1940:39).

(11) Class I

- a. hinyatihch, taharaku, harawihch, harakata.
 hinyatihch ta-hara=ku hara-wi=hch, hara-**kata**
 then AGNZ-sing=3ms sing-3ms.COMPL=SUBR sing-**3mp.HAB**

Then, as the singer sings, everyone sings. (41A:b)

- b. “po'ɔn. eruku,” nikatēni.
 po-a-n eru-**ku** ni-**kati**=ani
 look-2fs.COMPL-IMP laugh-**3ms.HAB** say-**3fs.HAB**=QUOT

“Look! He is laughing,” she said, it is said. (4A:az)

- c. hinyatihch, taru kɔsakɔni.
 hinyatihch ta-ru kɔsa-**ku**=ani
 now the-hickory scrape-**3ms.HAB**=QUOT

Now he was making hickory sticks, it is said. (4A:a)

(12) Class II

- hinyatihch, samahk'atani.
 hinyatihch sama-**hk-ata**=ani
 now cook-**HAB-3fs.CLASSII**=QUOT

Now she was cooking, it is said. (4A:e)

to all dual and plural forms and to singular Class II forms. Section 4.1.4 explains why KYLY decided to prioritize morphological completeness in Modern Tunica, resulting in the restoration [hk] and [hku] where they would have been elided in documented Historical Tunica.

6.2.4 Stative verbs

Stative verbs (what Haas calls static verbs) consist of a set of approximately 30 verbs, most having to do with emotional states ('angry', 'scared') or involuntary physical states ('sick', 'thirsty'). They take pronominal patient prefixes and usually take no agent inflection (as in 13a). The exceptions to this are when the verb is in the inchoative or in the future, where such verbs invariably take third-person feminine singular marking, either in the habitual (13b), completive (13c), or using *aki*, the 3fs form of the auxiliary *uhki* (13d).¹⁰

- (13) a. hinyatihch, sehitihch, tiyahpani.
 hinyatihch sehi-ti-hch, tihk-**yahpa**=ani
 now dawn-3fs.COMP=SUBR 3fs-**be.hungry**=QUOT
 Now when it dawned, she was hungry, it is said. (3B:b)
- b. ushpitukatēni
 uhk-**shpitu-kati**=ani
 3ms-**be.forgetful-3fs.HAB**=QUOT
 He forgets (you), it is said. (13:ao)
- c. “iman, wishpitutik'ahcha,” unikatēni.
 ima wihk-**shpitu-ti-k'ahcha** uhk-ni-kati=ani
 1s 2ms-**be.forgetful-3fs.COMP-FUT** him-tell-3fs.HAB=QUOT
 “You will forget me,” she said to him, it is said. (13:am)
- d. hinyatihch, ukik uyash'ekeni.
 hinyatihch u-ki=ku uhk-**yashi-aki**=ani
 now his-uncle=3ms 3ms-**be.angry-BE.3fs**=QUOT
 Then his uncle became angry, it is said. (4A:w)

10. See example 3 and the subsequent paragraph on page 71 for an introduction to *uhki*.

Though the limited number of stative verbs makes the data on them scant, the distribution of agent inflection on statives is of interest.

Theories of obsolescence and revival would expect aspects with no English parallel to become less prevalent. The previous section has shown evidence that this may be the case at least in the area of the interpretation of nonvolitional verbs as intransitives rather than the more marked transimpersonal.

The completive verbal inflections appear to be the most basic form of inflection; the habitual, future, and imperative all use the completive forms as their base. One would expect inflectional forms like auxiliary verb inflection that don't use the completive as a base to become disfavored during obsolescence. This should undoubtedly be the case in Modern Tunica, not due to natural processes of simplification but rather through KYLY's language planning decisions: auxiliaries are not included in the Tunica Language Textbook (Kuhpani Yoyani Luhchi Yoroni 2023). Because the habitual and completive aspects were deemed simpler and more regular forms, they are taught earlier for ease of pedagogy.¹¹

6.3 Gender

As mentioned in Section 2.3, much of the early linguistic interest around Tunica focused on its use of grammatical gender. Gatschet (1889), Swanton (1921), and Haas (1940) all prominently note Tunica's use of masculine and feminine nouns.

11. KYLY is in the process of writing a second textbook that is expected to contain chapters on use of both independent and inflectional auxiliary verbs.

Gatschet saw Tunica as unique in this regard: “The Tonika language is the more remarkable on this account, because it is the only language heretofore discovered *in America* which divides all objects of creation into two great sex-categories” (Gatschet 1889:167, emphasis in original).¹²

Swanton thought that elements of the Tunica gender system were undergoing grammaticalization and leveling:

It is perhaps for this reason that the plural forms sà'ma [sema] and si'nima [sinima] are sometimes used like plural suffixes to nouns. When si'nima is employed in this manner it is usually contracted to sin, and this contracted form appears occasionally after nouns usually considered masculine. It is possible, therefore, that sin was in process of becoming a true plural suffix to nouns without regard to gender.

Swanton (1921:8)

Swanton characterizes his attempted gender classification of nouns to be “incomplete” and comments that Gatschet’s work also has gaps in elicitation and confirmation of gender for some nouns (22). Some nouns were given as masculine and feminine at different times (23).

Haas’s *Tunica Dictionary* (1953) is replete with nouns for which the gender is unknown. This is because they were used with both masculine or feminine marking on different occasions and because inanimate nouns that do not appear as agents need not be marked for gender.

Heaton and Anderson (2017) analyze the Tunica gender system based on Haas’s documentation of Sesostrie Youchigant’s language. They note that grammatical gender is not common in North American languages and is especially uncommon in

¹² Gatschet uses “sex” here to denote categorization that is based on a gender paradigm rather than an animacy paradigm.

Southeastern languages, Yuchi being the only other exemplar (Heaton and Anderson 2017:341).

Heaton and Anderson agree with the grammar rules in Haas (1940): human animates abide by natural gender (mixed groups take masculine gender markings); inanimates have arbitrary gender in the singular and dual but are invariably feminine in the plural (as Swanton also noted); and inanimate collectives are feminine (103–105). Heaton and Anderson find that the feminine was likely the less marked gender (sometimes referred to as “default gender”) for non-human animates and inanimates (Heaton and Anderson 2017:342). They note that in addition to plural being marked as feminine, impersonal and transimpersonal verbs (an impersonal verb with the experiencer in object position) are conjugated solely as third-person feminine singular. Inanimate collectives are also marked as feminine, yet another example of indefinite, inanimate nouns tending toward the feminine. As a final piece of evidence, they find that loanwords from Spanish and French take feminine markings in Tunica (362).¹³

Tunica gender is marked in the following ways:

1. Subject inflection (example 14)
2. Pronominal object inflection (example 15)
3. Gender-number affixes on third-person nouns or noun phrases (example 16)
4. Possessive prefixes (example 17)

13. The same is not true for calques, which retain the gender of the Tunica noun implicated in the calque.

- (14) Inashiyakawe.
ihk-nashi-yaka-**wi**
me-lead-come-**3ms.COMPL**
He led me. (Swanton 1910)
- (15) Hi^ahiati ukunisema hotu **siwaliwima^a**
Hinyatihch uhk-oni-sema hotu **sihk-wali-wi-man**
Then his-man-3mp all **them-call-3ms-CONJ**
laputa.
lahpa-uta
land-3ms.CLASSII.COMP
So he called all of his people and he landed. (Gatschet 1886b:109)
- (16) hinyatihch, tingrasanakachohaku, yakan, ulaspihch, hot
hinyatihch ta-ingrasa-naka-choha-**ku** yaka uhk-laspi-**hchi** hotu
then the-American-war-chief=**3ms** come his-money=**3fs** all
uhan'uhkeni, polunt'ε tihayin.
uhk-ha-n-uhki=ani polun-t'ε tihk-hayi.
him-borrow-CLASSII-3ms.EXIST=QUOT hundred-big her-oldness.
Then the American general came and borrowed all of his money—a million dollars. (24:c)
- (17) hinyatihch, tihika hot'orahch, uroptin,
hinyatihch tihika hotu-ara-hch **uhk-roptini**
then summer finish-3fs.LIE=SUBR **his-cotton**
mar'uhkeni.
mari-uhki=ani.
gather-EXIST.3ms=QUOT
At the end of the summer, he picked his cotton, they say. (29:c)

Subject inflection is the only obligatory verbal inflection. It is limited almost entirely to animates, except for a few examples of the inanimate patients of unaccusative verbs, e.g., 'sink' (example 19), 'break' (example 18). Object inflection is optional and is far more likely to appear when the object is animate. Suffixation on the noun or noun phrase rarely occurs where the noun lacks the definite prefix *ta-*.

Table 6.6: Tunica gender-number affixes

Person	SG	DU	PL
3m	=ku	=unima	=sema
3f	=hchi	=sinima	=sinima

Nouns with gender-number suffixation are more common with animate and plural nouns, the latter of which, as Swanton noted, seemed to be the locus of leveling for inanimates. As for possessive prefixes, inanimates are not usually possessors unless reference is being made to their possession of part of the whole inanimate object. Thus, the gender of inanimates can remain unclear even if the noun appears multiple times in the corpus. The inanimate gender system is of great interest because it is arbitrary rather than predictable based on natural gender.

- (18) hinyatihch, hon'usan, tayunkahch,
 hinyatihch honu-usa-n ta-yunka=hchi
 then descend-go.3ms-FNL the-rope-3fs

kahch'okeni.
 kahchu-**aki**=ani
 break-**EXIST.3fs**=QUOT

They went (on) down until the rope broke. (22A:b)

- (19) hipuna'arahch, tahalihch, lamihtan, yahk'arahch,
 hipu-na'ara-hch ta-hali=hchi lamihta-n ya-hk-ara-hch
 dance-lie.3p the-land=3fs soft-FNL become-HAB-lie.3fs=SUBR

tahal ak'atani.
 ta-hali aka-**ata**=ani
 the-land sink-go.**3fs**=QUOT

While they danced the ground was getting mushy and was sinking. (5A:d)

When examining the speech of speakers of an obsolescing language, questions of attrition and speaker competence come into play. Because Tunica speakers shifted

to French and then to English, Tunica’s gender mismatch with French may have caused interference in inanimate gender assignment. Similarly, English’s lack of a gender system for inanimates may have motivated attrition in a speaker’s knowledge of gender, particularly for low-frequency words (Andersen 1982).

Modern Tunica has leveled what remained of the gender system in Tunica due to the uncertainty about the gender of a significant number of nouns. Certain phenomena that early documenters noticed have been kept: the tendency of non-animate plural, mass, and collective nouns to be feminine informed the decision to make all such nouns feminine in Modern Tunica. Similarly, nouns relating to bodies of water tend to be feminine,¹⁴ and have remained feminine in Modern Tunica. Modern Tunica has extended existing classifications to neologisms: because the sun and the moon are both feminine in Tunica, the names created for the planets of the solar system are also feminine.

By examining which inanimate nouns get gender markings, whether there are discrepancies in gender assignment over time, and how gender assignment compares to that of French, it can be established if—and to what extent—gender in Tunica changed or adapted over time.

6.4 Evidential =*ani*

Most of the Historical Tunica corpus consists of storytelling. Sentences in Tunica storytelling often end with =*ani*, meaning “they say” or “it is said”. The =*ani* mor-

14. This may be due to water bodies being conceived of as mass or collective nouns in Tunica.

pheme seems to serve as a quotative¹⁵ evidential that marks sentences describing events that the speaker did not personally witness but that have been reported or told to them.

The following two sentences appear consecutively in a story Sesostrie Youchigant told about his childhood. He saw the events described in example 20a, so the sentence does not contain a quotative. In 20b, however, he begins to speak of things he did not himself see or verify, necessitating the quotative.

- (20) a. hinyatihch, esiku, uwin, tekshash am'uhki.
 hinyatihch i-esi=ku uwi tekshash ami-uhki
 then my-father=3ms he Pinewood go-EXIST.3ms
 Then my father went to Pinewood. (29:b)

- b. hiyuk'uk'uhkəni.
 hi-yuka-uki-uhki=**ani**
 there-arrive-stay-EXIST.3ms=**QUOT**
 When he got there, he stayed, it is said. (29:b)

This is the only evidential morpheme in Tunica, and it is non-obligatory. The different frequency of use of these evidentials between speakers is obvious. The evidential appears to be completely absent from William Ely Johnson's storytelling; in Volsin Chiki's stories they appear frequently but not in every sentence where a reportative would be called for. Sesostrie Youchigant uses the evidential very often.

Both Swanton and Haas often omitted "it is said" from their English translations of the text. Haas explained her omission as due to its frequency and lack of direct translation into English (Haas 1953:8). The translated "it is said" has been included

15. It might be said to be a reportative evidential (Willett 1988), but Haas calls it a "quotative postfix", so quotative is used in this work.

in English translations of Tunica throughout this dissertation for accuracy and completeness.

Examining the exact frequency of use of this particle in the speech of the three consultants may imply something of the speaker's familiarity with Tunica storytelling conventions.

6.5 Elision in possessive prefixes

Possession in Tunica is accomplished through noun prefixation. Gatschet (1886b), Swanton (1921), and Haas (1940) all lay out the possessive prefix paradigm. Swanton shows that each possessive prefix has two forms but does not comment on any difference between them or why they appear in pairs.¹⁶ The most plausible reason for the lack of further comment is that he saw each pair as environmentally-conditioned allomorphs.

Haas (1940), however, saw these two forms as belonging to two sets of related possessors, one set marking alienably-possessed nouns and one marking inalienably-possessed nouns, as shown in Table 6.7. The two sets are distinguished by adding an [hk] or [k] to what Haas deemed the inalienable forms. Throughout Tunica, however, *hk* tends to reduce to *h* or syncope completely in certain environments.

Haas gives the following rules concerning the consonantal syncope of *k* and *hk*:

2.252 A *k* which would come to stand between an *h* and a voiceless stop is dropped. Examples:

16. Swanton lists morpheme pairs for every person and number other than the third-person feminine singular, for which he lists only one morpheme, *tihk-*.

ti'tiht[?]ε river < *ti'tihk(i)* bayou + *-t[?]ε*, aug. suff.
ya'ruhto'hku hatchet < *ya'ruhk(i)* ax + *-to'hku*, dim. suff.
yu'sahka'yi yellow catfish < *yu'sahk(i)* catfish + *ka'yi* yellow
[?]*ihti'ra* my cloth < [?]*ihk-*, 1S pref., + *ti'ra* cloth

2.253 An *hk* group which would come to stand before a continuant other than *h* is dropped. Examples:

[?]*a'sume'li* Easter day < [?]*a'suhk(i)* day + *me'li* black
ta'kosusi'nima the crawfish < *ta'-* + *ko'suhk(i)* crawfish + *-si'nima*,
 FD-P suff.
[?]*usa't[?]ε* his horse < [?]*uhk-*, 3MS pref., + *sa't[?]ε* horse

2.254 In the case of an *hk* group which would come to stand before *h...* [w]hen a prefix ending in *hk* comes to stand before a stem beginning in *h*, the *hk* is dropped as before all other continuants...

Haas (1940:25)

These rules would eliminate the *hk* in possessive prefixes in a wide variety of environments, making the alienable prefixes identical to the inalienable prefixes, as shown in examples 21 and 22.

- (21) hinyatihch, **urish** atani.
 hinyatihch, **uhk-ri-sh** ata=ani.
 then **his.AL-house-to** GO.3fs=QUOT
 Then she went to his house, it is said. (4A:ae)

- (22) hinyatihch, **tisasiniman**, tihtap[?]onasiteni.
 hinyatihch, **tihk-sa=sinima** tihk-tapi-onasiti=ani
 then **her.AL-dog=3fp** her-grab-BE.3fp=QUOT
 Then her dogs seized her, it is said. (4A:as)

Haas herself seems to have changed her view of how the possessive prefixes worked. In her 1946 sketch grammar, Haas makes no mention of alienable and inalienable prefixes being marked by different forms of the prefix, instead using the following set of rules to condition the infixation of *h* or *hk*:

An *h* is regularly inserted between a pronominal prefix (with the exception of those ending in *n*) and a stem beginning in a voiceless stop

Table 6.7: Tunica possessive prefixes (Haas 1940:37)

	singular		dual		plural	
	al.	inal.	al.	inal.	al.	inal.
1	ihk-	i-	ink-	in-	ink-	in-
2m	wihk-	wi-	wink-	win-	wink-	win-
2f	hihk-	hi-	hink-	hin-	hink-	hin-
3m	uhk-	u-	unk-	un-	sihk-	si-
3f	tihk-	ti-	sink-	sin-	sink-	sin-

provided the stem has a stress on its first syllable... Between a pronominal prefix and a stem beginning in ? [glottal stop] a consonant group -hk- is regularly inserted except after prefixes ending in n where only -k- is permitted... Moreover, when the pronominal prefixes are appended to a stem beginning in a vowel the vowel of the prefix contracts with this initial stem vowel...

Haas (1946:347–348)

While Haas’s explanation changed, the outcome of her rules is the same. Most Tunica roots are CVCV, so Haas posits that all independent Tunica roots are consonant-initial. Independent roots that might appear to start with a vowel actually start with a glottal stop. Those roots that Haas says are bound roots—due to inalienable possession—are deemed to start with a vowel and not a glottal stop, as shown in example 23a. Under Haas’s rules, the possessive prefixes will not get a *k* or an *hk* appended because there is no glottal stop. Those roots that are not inalienably possessed are deemed to have a glottal stop, and therefore the *hk* is appended, as in example 23b.

- (23) a. -asa f. (occ. m.) poss. n. tail (Haas 1953)

nis[?]ahan tayak [?]osahch, yurɔni.
 nisa-aha ta-ya=ku u-asa=hch yuru=ani
 new-NEG the-deer=3ms **his.INAL**-tail=SUBR long=QUOT

In olden times Deer's tail was long, it is said. (16:a)

b. [?]ala f. (occ. m.) arrow (Haas 1953)

hinyatihch, uhahkamuchi hotutihch, **uhk[?]ala**
 hinyatihch uhk-hahkamuchi hotu-ti-hch **uhk-ala**
 then his.AL-bread finish-3ms.COMPL=SUBR his.AL-arrow

shuchikɔni.
 shuchi-ku=ani
 shoot-3ms=QUOT

When his bread was gone, he shot his arrows, it is said. (4A:bi)

Interestingly, Gatschet's consultant, William Ely Johnson, uses the definite article *ta-* on an ostensibly inalienable noun without any possessive prefix whatsoever.

(24) tanisharaⁿ wanan
 ta-nishara wanani
 the-spleen fever
 inflammation of the spleen (Gatschet 1886b)

In example 24, the spleen, a body part that Haas labels as inalienably possessed, is used here with articular *ta-* rather than a possessive pronoun. Johnson's lack of possession on nouns that Haas considers inalienably possessed justifies further investigation into when and where possession occurs and what phonetic form the possessive prefix takes.

The decision not to level possession into a single form in Modern Tunica goes against the principle of ease of pedagogy discussed in Section 4.2. English has no parallel alienable/inalienable system of possession and it is certainly easier to learn a single possessive form than two. The principle of morphological completeness

Table 6.8: Haas (1946) word order chart

	SC	PM	IS	IO	PM	PC	PW	PM
1							pi'tahk [?] una'nì. He was walking.	
2						lɔ'ta To run	wiwa'nān. do you want?	
3					hɔ'waš, Outside		howe's [?] uhke'nì. he jumped out.	[hɔ'wāš.]
4				ya' [?] ān A deer			[?] uwe'n [?] uhke'nì. he found him.	
5			ta'ruštāku, Rabbit				[?] u'š [?] epa'nì. he was glad.	
6			to'nīku, The man	ta'yak the deer			[?] ura'p [?] uhke'nì. he killed him.	
7	h-hč Then		[?] u'wīn, he		ma'hon just		[?] una'nì, he was sitting	[?] u'rīs. at home.
8	h-hč Then	ha'tikān, again	ta'ruštāku Rabbit	tɔ'mahkak Alligator		[?] uhpe'kawan to hit him	yako'nì. he was doing.	

SC = sentence connective, PM = predicate modifier, IS = independent subject, IO = independent object, PC = predicate complement, PW = predicate word

solves the issue of students having to learn in which environments alienable prefixes reduce; in Modern Tunica, alienable prefixes invariably use their full *hk* forms.

It is worth exploring whether this elision of alienable forms is part of a morphological reduction, a merger in progress between inalienable and alienable possession, or simply an environmentally-conditioned allomorph.

6.6 Word order

Section 2.1 discusses the Tunica people's known interactions with other indigenous peoples and languages. SOV languages are common in the American Southeast and most of the indigenous languages with which we know the Tunica people had close contact were also SOV (Martin 2004:85). The existence of Tunica words for tribe names is another indication of contact, though the degree of that contact cannot be

ascertained. Of the tribes for which Sesostrie Youchigant knew names in Tunica, the Alabama, Biloxi, Chickasaw, Chitimacha, Choctaw, Koasati, Natchez, and Ofo all have SOV word order.¹⁷ The other tribes he knew names for were the Avoyel (undocumented), Houma (Muskogean; very likely SOV), Karankawa (no dominant word order), Koroa (undocumented; thought to be a Tunican language mutually intelligible with Yazoo), Opelousas (undocumented; possibly Atakapan, which is SOV), and Pascagoula (undocumented).

Mobilian Jargon may have had a different word order than Tunica: Drechsel (1996) proposes that Mobilian Jargon used OSV word order rather than the SOV that is common among the Muskogean languages that serve as Mobilian's lexifier (250).

As contact with French, Spanish, and English became more and more intense in the 18th and 19th centuries, Tunica speakers were confronted with a succession of SVO languages with considerable political, economic, and military power in the area. These dominant SVO languages may have caused a loosening of or change in Tunica word order as Tunica speakers began shifting first to French, then English.

6.7 Conclusion

A variety of features have been selected in an attempt to examine language change across levels of linguistic structure. Phonology (*hk* elision), morphology (gender

17. There is a word that appears in one of Sesostrie Youchigant's texts, *washashe*, that is glossed as "the shout of a captive". Haas notes that the word probably means Osage, as it is similar to Natchez and Creek exonyms for Osage. Osage is SOV.

marking, verb aspect, the quotative evidential), and syntax (word order distribution), along with features identified by documenters and speakers of Historical Tunica.

The next chapter will interrogate whether change was occurring over the span of three generations of Historical Tunica and examine the differences between Historical Tunica and Modern Tunica. It will also explore aspects of Tunica that have not yet been analyzed across a relatively complete corpus.

Chapter 7: Data

Having outlined the phenomena to be studied in Chapter 6, we now describe of the data in the corpus.

Table 7.1 contains basic information about the size of the corpus, shown divided into subcorpora from each time period. Though not exhaustive, and heavily weighted toward Mary Haas and Sesostrie Youchigant’s collaboration, the standardized input of these texts allows for comparison.

When comparing texts, the text’s genre is pertinent. Studies have shown that linguistic variation is stark enough between different types of discourse (e.g., fiction vs. non-fiction) that analyzing a few basic linguistic variables can determine genre with some accuracy (Karlgren and Cutting 1994; Kessler, Nunberg, and Schutze 1997; Malrieu and Rastier 2000). Dividing the texts into genres enables the comparison of texts of the same genre where possible. Haas (1950) divides the texts into seven categories. Table 7.2 shows these categories, along with the number of

Table 7.1: Corpus information

Date	Name	Texts	Sentences	Words	Morphemes
1886	Gatschet	8	343	1,542	3,468
1910	Swanton	2	173	486	1,162
1933–1939	Haas	76	3,005	13,242	32,933
2017	Modern Tunica	1	217	587	884
Total		87	3,738	15,857	38,447

Table 7.2: Tunica text categories in Haas (1950)

Category	Unique texts	Total texts
Myths	7	14
Tales	6	8
Animal Tales	6	7
Historical and Pseudohistorical Narratives	9	11
Personal Narratives	6	6
Miscellaneous Ethnological Data	14	14
Fragments and Miscellanea	15	15
Fragments: Strange Beings	9	9
Miscellanea	6	6

unique texts and the total number of texts in the category. Haas’s categories were used as a starting point.

Haas describes Myths as having a “sacred character” while Tales, Animal Tales, and some of the Fragments and Miscellanea constitute “secular lore” (1950:1). For this corpus, the texts from Fragments and Miscellanea were reassigned to the category that matched the genre of the fragment. Six were assigned to Tales, six to Animal Tales, two to Ethnographic Data, and one to Historical Narratives.

The texts from the other corpora were integrated into these categories. Information about the corpus, separated by both genre and subcorpus, is shown in Table 7.3.

Unfortunately, no genre category contains texts from all four subcorpora, but Myths provides exemplars from all three eras of documented historical Tunica. Personal Narratives contains the Modern Tunica text alongside texts from two historical periods, allowing for a comparison of obsolescent and post-revival language.

Of particular interest are two myths told by two different speakers. William Ely Johnson in 1886 and Sesostrie Youchigant in the 1930s both told the story of

Table 7.3: Corpus by text genre

Genre	Texts	Sentences	Words	Morphemes
Myths	17	1,571	6,485	16,292
Gatschet	1	206	760	1,799
Swanton	2	173	486	1,162
Haas	14	1,192	5,239	13,331
Modern Tunica	0	0	0	0
Tales	14	646	2,861	6,845
Gatschet	0	0	0	0
Swanton	0	0	0	0
Haas	14	646	2,861	6,845
Modern Tunica	0	0	0	0
Animal Tales	14	472	2,057	5,072
Gatschet	0	0	0	0
Swanton	0	0	0	0
Haas	14	472	2,057	5,072
Modern Tunica	0	0	0	0
Historical	13	489	2,137	5,226
Gatschet	1	27	111	233
Swanton	0	0	0	0
Haas	12	462	2,026	4,993
Modern Tunica	0	0	0	0
Personal Narratives	10	318	1,091	1,975
Gatschet	3	24	187	373
Swanton	0	0	0	0
Haas	6	77	317	718
Modern Tunica	1	217	587	884
Ethnological	19	242	1,226	3,038
Gatschet	3	86	484	1,064
Swanton	0	0	0	0
Haas	16	156	742	1,974
Modern Tunica	0	0	0	0

the orphan who becomes Thunder. And both Volsin Chiki in 1910 and Sesostrie Youchigant in the 1930s related the Tunica Flood Myth story. Sesostrie Youchigant told both the Thunder Myth and the Flood Myth twice, making it possible to analyze the variation within a single speaker's retellings and the variation between speakers.

7.1 Historical phenomena noted by documenters and speakers

Section 6.1 presented the phenomena that the linguists and consultants noted as being likely areas of diachronic change at the time of documentation. This section analyzes the data in the current corpus related to these phenomena.

7.1.1 Swanton's incorporated suffixes

Swanton's evidence for his hypothesis that *-p* was formerly a causativizing suffix and that *-u* was a nominalizing suffix stem from his observed difference in the frequency with which each of these phonemes appears in verbs and nouns. He claims /p/ is much more frequent as the final consonant in verbs than in nouns and that this discrepancy might be explained if *-p* were a verbal suffix that had been incorporated into the verb root. Swanton's choices surrounding his analysis are puzzling. First, he classes adjective stems with verbs. If he is trying to determine whether *-p* is a verbal causativizer, including adjectives with verbs seems to be counter to his objectives. The table of his results contains close but slightly differing totals for nouns and verbs. There is no discussion of whether this is the total number of nouns and verbs (and adjectives) or a subset. Swanton's table is reproduced in Table 7.4.

Table 7.4: Last consonant in verbal and nominal stems
from Swanton (1921)

Sound	Verb	Noun
r	59	42
k	55	83
p	53	11
t	28	20
tc	22	20
h	20	11
l	19	30
n	16	35
s	15	14
y	15	13
m	12	16
c	10	30
w	9	4
Total no. cases	333	

Table 7.5 analyzes the non-compound nouns and verbs in the lexicon of the current Tunica dictionary. With half as many nouns as verbs, a random phonetic distribution would result in about half as many nouns as verbs with each final consonant. There are, however, outliers on both sides.

Swanton's data and the current corpus show that instance of /p/ as the last vowel have the smallest noun to verb ratio but the ratio for /h/ is also unexpectedly low. Additionally, /ʃ/ shows a significant deviation from the mean ratio, with more nouns than verbs having last consonant /ʃ/ despite twice as many verbs as nouns in the lexicon. Swanton's hypothesis deserves further study, but frequency distribution is not dispositive.

Table 7.5: Last consonants in the current Tunica lexicon

Phoneme	Verbs	Nouns	Total	Noun/Verb ratio
h	47	12	59	25.5%
k	101	63	164	62.4%
l	44	26	70	59.1%
m	33	25	58	75.8%
n	40	31	71	77.5%
p	64	13	77	20.3%
r	82	32	114	39.0%
s	29	13	42	44.8%
ʃ	21	25	46	119.0%
t	62	20	82	32.3%
tʃ	34	17	51	50.0%
w	22	11	33	50.0%
	579	288	867	49.7%

7.1.2 Haas and Sesostrie Youchigant's gender-number affix (GNA) analysis

Haas probed Sesostrie Youchigant about the difference between the two different forms of each singular gender-number affix (GNA). Sesostrie Youchigant nearly always used the main =*ku* and =*hchi* forms for third-person masculine and feminine gender-number affixes. Because the alternative forms are so seldom used, it is difficult to determine the distribution of their usage.

Table 7.6: Forms of third-person singular GNAs

Gender	Main form	Alternate form
M	=ku	=kuhu
F	=hch(i)	=hchihi

The following examples represent every occurrence of the alternate forms in the texts.

- (25) “ikikuhun, ihal'uhki,” nikōni.
 i-ki=**kuhu** ihk-hali-uhki ni-ku=ani
 my-uncle=**3ms** me-send-BE.3ms say-3ms.HAB=QUOT
 “My uncle has sent me,” he said. (4A:ad) and (4B:y)
- (26) “wihkəkashi ushtamar'ehēkuhun,
 wihk-əka-shi uhk-shtamari=aha=**kuhu**
 2ms.POSS-child-male he-to.be.polite=NEG-**3ms**
 hopiwik'aha, uniwihki.”
 ho-pi-wi-k=aha uhk-ni-wihki
 out-emerge-3ms-SBJV=NEG him-tell-BE.2ms
 “You had told your impudent boy not to come out.” (4A:v)
- (27) uhayishikuhun, tarushtak uyan[?]uhkeni.
 u-hayishi=**kuhu** ta-rushta=ku uhk-yana-uhki=ani
 he-to.be.above=**3ms** the-rabbit=3ms him-speak.to-BE.3ms=QUOT
 The One Above spoke to Rabbit. (14A:a)
- (28) hinahkuhch, tachahtakuhun, kakun, olu ya'uhkik'ihch,
 hinahkuhch, ta-chahta=**kuhu** kaku olu ya-uhki-k'i=hch
 thus the-Choctaw=**3ms** someone sin do-BE.3ms-if=SUBR
 hihch uwi ya'uhki.
 hihchi uwi ya-uhki
 there he do-3ms
 If anyone had committed a crime, it was the Choctaw (who) had done
 (so). (24:g)

Despite Sesostrie Youchigant’s statement, quoted by Haas, that “the old folks used to talk that way” (Haas 1933), there were no instances of these alternative GNAs in either Gatschet’s or Swanton’s material. The limited data available make it hard to come to any firm conclusion about the precise meaning of =*kuhu* or =*hchih*. Haas’s suggestion that they are analogs of English *it*-clefts is plausible for some of the phrases (especially example 28), but awkward for others. For instance, example 27 is the first line of the story, where an *it*-cleft seems an unlikely construction as there

is no set of people from which one can be singled out with an *it*-cleft. In example 25, Eagle has just asked the speaker what he is doing. Had Eagle asked who sent him, an *it*-cleft would be wholly appropriate, but less so with the more generic question at hand. As Kuhpani Yoyani Luhchi Yoroni (KYLY) works to transcribe the rest of Gatschet, Swanton, and Haas's unpublished work and field notebooks, perhaps other instances of these alternate forms will be found in elicited phrases, providing more information about the usage of the forms.

7.1.3 Reanalysis of verbs of nonvolitional action

Haas noted that Sesostrie Youchigant indicated that nonvolitional action verbs could be conjugated as transimpersonals—impersonal verbs with the experiencer of the action in the object position—or as intransitives, with the experiencer in the subject position, with the latter being the more common conjugation. Sesostrie Youchigant lists *heha* 'breathe', *uhu* 'cough', *owi* 'sweat', and *lehu* 'pant' as able to be conjugated either way, while one nonvolitional verb, *hiya* 'awaken', is only conjugated transimpersonally (Haas 1940:65). Haas posits that such verbs may have undergone a change from transimpersonal conjugation to intransitive conjugation.¹

The corpus data for these verbs is sparse, but it does provide some support to the hypothesis. The only verbs in the list that appear in the texts are *uhu* 'cough'

1. A similar morphosyntactic change is noted in Campbell and Muntzel (1989), which describes American Finnish allowing the inclusion of an agent in passive constructions (likely due to the influence of English), whereas passives in Finnish prohibit agents in these forms. Dorian (1973) discusses a similar case of analogical addition of a subject in a passive construction that previously did not allow subjects. In the case of Tunica, non-volitional verbs seem to have been remade as intransitive verbs under the influence of French.

and *owi* ‘sweat’; *uhu* appears as a verb only in a text Haas gathered from Sesostrie Youchigant in the 1930s (example 29), while *owi* appears twice in a text Gatschet gathered from William Ely Johnson in 1886 (examples 30 and 31).

- (29) uhukoni.
 uhu=ku=ani
 cough=3ms.HAB=QUOT
 He was coughing, it is said. (19:b)
- (30) hiⁿatish ukowikati.
 hinyatihch uhk-owi-kati
 then him-sweat-3fs
 Then he sweats. (Gatschet 1886b)
- (31) tuni yaxshiku ukowiχkinda.
 ta-oni yashi=ku uhk-owi-hk-inta
 the-person sick=REL him-sweat-HAB-CLASSII.CAUS.1p
 We make the sick person sweat. (Gatschet 1886b)

Sesostrie Youchigant’s sole use of a nonvolitional verb is inflected as an intransitive verb with the experiencer as the subject. In the two instances from William Ely Johnson, gathered nearly 50 years beforehand, the experiencer (the one who is sweating, in examples 30 and 31) is in the object position. In example 30, the subject marking is the third-person feminine singular, used in Tunica impersonal verbs as the expletive subject. In example 31 the experiencer remains in the object position, but the non-causative expletive subject has been replaced by the causative first-person plural; deliberate action by others is precipitating the sweating, so those others replace the third-person feminine singular subject marking as the agents.

7.2 Verb inflection

Section 6.2 describes Tunica’s verbal inflection system. One would expect a decrease in inflectional variety due to the stylistic shrinkage associated with obsolescing languages (Mougeon and Beniak 1989). Similar to the reanalysis of certain non-volitional verbs as intransitive rather than transimpersonal, one would expect obsolescent Tunica to lose less regular forms. Since most inflections use the completive suffixes as their base, one would expect verbs that do not use the completive as a base to be less used in obsolescence and revival. This would implicate auxiliary verbs as an inflectional category that would see less use. Table 7.7 shows the percentages for the different inflectional categories by era.

Table 7.7: Verb inflection percentages by era

era	completive	habitual	<i>-uhki</i>	<i>-usa/-uwa</i>	<i>-ura/-una</i>	stative	indep. aux.
Gatschet	57.82%	9.05%	10.89%	2.15%	3.37%	8.13%	8.59%
Swanton	36.40%	6.69%	39.33%	0.00%	1.67%	7.53%	8.37%
Haas	30.95%	23.34%	19.84%	4.50%	10.32%	4.56%	6.49%
Modern Tunica	47.73%	35.23%	1.14%	0.00%	0.00%	13.64%	2.27%

Instead of a decrease in inflectional variety over time among the Historical Tunica subcorpora, there is a wide variation in the relative percentages of different verb inflections, and the last Historical Tunica subcorpus shows that Sesostrie Youchigant’s speech made the most use of the different forms of inflection. He has by far the highest percentage of the positional or “durative” inflectional suffixes, *-una* and *-ura*, as well as the directional inflections *-uwa* and *-usa*.

It is possible that the three speakers of Tunica had widely varying speaking styles that employed different inflections and aspects in their storytelling. This,

however, seems unlikely, as Swanton's consultant, Volsin Chiki was the maternal uncle of Haas's consultant, Sesostrie Youchigant. Tunica family organization was matrilineal, and maternal uncles were very close relations indeed. A more likely scenario is that the speakers of Historical Tunica had differing facility with the language; perhaps Sesostrie Youchigant had a better command of the Tunica's rich aspectual system and thus felt more at ease creating sentences that made use of all of them.

The Modern Tunica corpus adheres to what would be expected in a revival setting whose teaching methods prioritize the completive and habitual forms over auxiliary inflection. This is reflected in the very low rate of *-uhki* usage to reflect the perfect and the lack of any directional or positional inflection. It will be interesting to see if these numbers change with the release and use of the second Tunica language textbook, which will teach auxiliary verbs.

The greater use of stative verbs is also indicative of strategies to express oneself in Tunica without the need for conjugation. Tunica stative verbs do not require any inflectional suffixation at all.² Additionally, stative verbs are especially useful in conversation, as they convey one's emotional state, which is particularly important in discourse compared to narrative storytelling.

2. Stative verbs can take third-person feminine singular suffixes to make them inchoative.

Table 7.8: Gender and number distribution

	Singular	Dual	Plural	Total
F	600	–	116 ³	716
M	1011	15	193	1219
	1611	15	309	1935

7.3 Gender

Section 6.3 describes Tunica’s gender system and Swanton and Haas’s hypothesis that the Tunica gender system might be undergoing—or had recently undergone—leveling, at least in making the plural invariably feminine for inanimates. Haas also notes that Sesostrie Youchigant would use both genders with the same noun at varying times.

7.3.1 Leveling of gender in the plural

The texts support the observation that non-animate plurals are invariably feminine, with one possible exception. Table 7.8 shows data on the patterning of gender across singular, dual, and plural nouns with gender marking on the noun itself.

The most significant discrepancy is between the genders in the singular: 62% of the gender-marked singular nouns are masculine. Surprisingly, there are more gender-marked masculine nouns than feminine nouns in the plural as well. Dividing the data by subcorpus, as in Table 7.9, reveals similar patterns.

3. Since there is no specific dual GNA for the feminine, this number represents both dual and plural feminine GNAs.

Table 7.9: Noun-marked gender by person, number, and subcorpus

	MS	FS	MD	MP	FP
Gatschet	73	41	4	12	9
Swanton	6	6	0	6	0
Haas	922	547	11	163	107
Modern Tunica	10	6	0	12	0

Table 7.10: Gender-marked plural nouns by animacy, gender, and subcorpus

Gender/ Number	Type	Gatschet	Swanton	Haas	Modern Tunica	Totals
FP	People	2	0	3	0	5
	Animals	2	0	11	0	13
	Inanimates	2	0	2	0	4
FP Total		6	0	16	0	22
MP	People	8	2	14	3	27
	Animals	0	0	0	0	0
	Inanimates	0	0	0	0	0
MP Total		8	2	14	3	27
Grand Total		14	2	30	3	49

The notion that gender-marked noun stems are more likely to be animate is borne out here. Because the corpus has not yet been semantically tagged for animacy, the animacy of the noun roots with plural markings was undertaken manually. Table 7.10 provides the number of unique gender-marked plural noun roots in each subcorpus and divides them into animacy categories.

The table shows that although there are more masculine-marked plurals, no masculine-marked plural refers to an inanimate in any subcorpus. In fact, the masculine-marked plural nouns refer solely to people. On the other hand, feminine plurals are often used to refer to groups of animals, and four plural inanimate noun roots are marked with feminine gender: in the Gatschet subcorpus, *tayashisinima*

‘the diseases’ and *taniriwekasinima* ‘the graves’ both take feminine gender-marking, and in the Haas subcorpus, *talasinima* ‘the cane’ or ‘the arrows’ appears, as does *taherisinima* ‘the boats’.

Based on the limited data in the texts, animate nouns are more likely to be gender-marked and all non-human animates and inanimates take the feminine in the plural. Whether inanimate plural nouns were always invariably feminine in Tunica or whether this invariability represents a regularization of the Tunica gender system is impossible to determine from the data in the corpus.

7.3.2 Variable gender

Another argument for gender being a site of language change is Haas’s claim that Sesostrie Youchigant genders some nouns inconsistently; several appear in the same subcorpus as both masculine and feminine. All but two of the 28 nouns that are variably gender-marked are in the Haas subcorpus, with the other two coming from the Gatschet corpus. Most of the nouns, including both of the nouns in the Gatschet corpus, refer to animates and are explainable by the differing natural gender of the referent itself. Differing gender in roots like *-ahaya* ‘sibling of the opposite sex’, *ɔka* ‘child’, or *hichut’ε* ‘eagle’ would be variably marked depending on the natural gender of the siblings, the child, or the eagle.

A few of the nouns that appear in both masculine and feminine forms do not do so because of natural gender; these may indeed be examples of speaker uncertainty. A prime example is Sesostrie Youchigant’s use of *-asa* ‘tail’. The term appears six

times with a masculine GNA and 14 times with a feminine GNA, sometimes being marked as masculine and feminine within the same text.

Haas had this to say about Sesostrie Youchigant's use of *-asa* in a text about Deer and Turtle:

Note that *-a'sa* is sometimes used with the FS suff. and sometimes with the MS suff. It is possible that the informant was not sure of the exact gender classification of this [noun]... In this text *-a'sa* occurs three times with the FS suff. and twice with the MS suff.

Haas (1950:113n3)

Speaker uncertainty is especially plausible since the tail in this story belongs to the same entity. Further examples support this argument. Examples 32 and 33 both refer to the same event, once as it happens and once just after it has occurred, yet the gender marking differs.

- (32) hinyatihch, tanarak ɔsak usaman'uhkeni.
 hinyatihch ta-nara=ku u-asa=**ku** uhk-sama-n-uhki=ani
 then the-snake=3ms his-tail=**3ms** it-cook-CAUS-BE.3ms=QUOT
 Then he cooked the snake's tail. (11A:c)

- (33) hinyatihch, tanarak ɔsah**ch**, saman'uhkihch,
 hinyatihch ta-nara=ku u-asa=**hchi** sama-n-uhki=hch
 then the-snake=3ms his-tail=**3fs** cook-CAUS-BE.3ms=SUBR
 nesha walikɔni.
 nesha wali-ku=ani
 mudfish call-3ms=QUOT
 The snake's tail (that) he had cooked he called mudfish. (11A:d)

Unfortunately, *-asa* appears only once in another subcorpus (Swanton), and in that instance, it is not marked for gender and number. This rules out a diachronic analysis of *-asa*.

The terms that vary in marking do not appear to belong to any discernible class or category of noun. They include terms like *tasehapo* ‘mirror’, *wishkatahiyunka* ‘bowstring’, and *rihkunaruhki* ‘pestle’.

It is important to emphasize that unexpected variations in gender should not be immediately attributed to speaker uncertainty. In addition to plurals seeming to be invariably feminine, Haas proposes that the use of feminine singular on an otherwise masculine noun indicates that the noun is in an uncountable or collective state. The entry for *hahka* ‘corn’ in Haas’s *Tunica Dictionary* (1950) is illustrative:

hahka (art. *tahka*) m. corn, maize; f. col. (a mess of) corn

Again we must rely entirely on the Haas subcorpus, which contains the only instances of gender-marked *hahka*. There are 11 instances of *hahka*, eight masculine-marked and three feminine-marked. Of the three feminine-marked instances, two pertain to pounding corn and one to planting it. The masculine-marked forms refer to creating the corn plant, breaking off the cob from the plant, and eating the corn.

It is plausible that the masculine marker is used to refer to the corn when it is on the cob, and that feminine-marked *hahka* refers to corn in the form of a mass of corn kernels (either for planting or pounding), which Haas calls “(a mess of) corn”.

Cultural considerations can also account for gender variation. A glance at the data shows that the term for a fire is feminine, except for one instance in which it is masculine. Looking more closely at that instance reveals that the sentence describes a fire’s transformation into an old man, prompting the change in gender marking.

Table 7.11: Prefixed noun gender marking

Era	Non-marked	Total	% unmarked
Gatschet	143	253	56.5
Swanton	48	59	81.4
Haas	1134	2619	43.3
Modern Tunica	41	51	80.4

7.3.3 Frequency of GNAs over time

Languages often tend to drop morphemes as they obsolesce (Schmidt 1985a; Campbell and Muntzel 1989). One area where this theory can be tested is in an analysis of how often the GNA is absent when a noun prefix (either articular *ta-* or a possessive prefix) is present. GNAs are much more likely to appear on nouns when a prefix is also present.

Table 7.11 shows the number and percentage of prefixed nouns that are not marked for gender. The results do not point to any clear pattern over time. The Swanton subcorpus has a much higher percentage of prefixed nouns unmarked for gender.

Genre differences do not seem to account for the discrepancies, as Table 7.12 illustrates. The difference in the Haas subcorpus between Tales and Animal Tales is particularly puzzling and may indicate that these frequencies are very particularized to the nouns in the individual texts.

Two stories, the Orphan Myth and the Flood Myth, appear as texts in two separate subcorpora. A comparison of different tellings of the same general story by two different speakers in two different periods limits the effect attributable to differences between the narratives themselves. Comparing these texts suggests that

Table 7.12: Percentage of non-gender-marked prefixed nouns by genre and era

Genre	Gatschet	Swanton	Haas	Modern Tunica
Animal Tales			32.4	
Ethnological	80.6		47.3	
Historical	50.0		36.5	
Myths	45.6	81.4	44.3	
Personal Narratives	66.7		45.3	80.4
Tales			55.7	

Swanton's consultant, Volsin Chiki, may be an outlier regarding how often he marks gender on nouns. The Haas corpus has two tellings of both the Orphan Myth and the Flood Myth. Haas's transcription of Sesostrie Youchigant's speech shows that his marking of nouns remains fairly consistent between his two tellings of each story. William Ely Johnson's telling of the Orphan Myth to Gatschet and Sesostrie Youchigant's two tellings of the same story to Haas have very similar frequencies of gender-marked nouns (Table 7.13). Volsin Chiki's frequency of marking gender on prefixed nouns of the Flood Myth he told to Swanton, on the other hand, differs greatly from Sesostrie Youchigant's telling of the same story to Haas. The stories vary somewhat but have similar elements and themes. Haas had this to say about the differences: "Youchigant's version of the flood myth differs slightly from the version given by Swanton's informant, Volsine [sic] Chiki. Since Chiki was Youchigant's maternal uncle and since Youchigant learned his version from his mother, it seems strange that the two versions should differ. However, it seems reasonable to assume that the two versions may have sprung from a slightly longer version containing the essential elements of both recorded versions" (Haas 1950:2).

Table 7.13: Frequency of gender marking of prefixed nouns
in tellings of the Orphan Myth

Text	unmarked	total	% unmarked
Gatschet	68	149	45.6
Haas 1	180	373	48.3
Haas 2	187	429	43.6

Table 7.14: Frequency of gender marking of prefixed nouns
in tellings of the Flood Myth

Text	unmarked	total	% unmarked
Swanton	37	43	86.0
Haas 1	29	70	41.4
Haas 2	22	51	43.1

The Modern Tunica subcorpus marks gender on prefixed nouns at about the same rate as the Swanton corpus but much less often than the Gatschet and Haas subcorpora. KYLY has almost completely leveled gender in the singular, which is now invariably masculine, except for collectives, bodies of water, celestial objects, and animates with natural gender. Due to this leveling and the lack of grammatical gender in English, less diligence in marking gender in Modern Tunica is understandable.

The study of gender marking on nouns during Tunica obsolescence and revival has shown that all inanimate plurals are indeed marked as feminine throughout the corpus. Thus, beyond Swanton's speculation, there is no evidence that invariably feminine gender markings in the plural represent linguistic change; if gender ever underwent a process of regularizing to the feminine in plural nouns, it would have happened before the language was first documented.

All evidence of variable gender where natural gender is not at issue is in the Haas corpus, the latest of the three Historical Tunica periods. While the number of tokens is small, it is certainly possible that, at least for Sesostrie Youchigant, the Tunica gender system was in unstable. The further leveling of the gender system in Modern Tunica resolves this instability.

Frequency of gender marking on nouns varied between the subcorpora, and not in any clear direction. It is uncertain whether such marking was the subject of stylistic variation between speakers or representative of a difference in command of the language.

Gender remains fertile ground for future study: once the corpus is tagged for semantic role, data on noun gender from verb marking may reveal more information about the Tunica gender system.

7.4 Evidential =*ani*

Section 6.4 describes the Tunica quotative evidential =*ani*. Table 7.15 shows the percentage of sentences that make use of it, divided into speaker and era.

There are no instances of =*ani* in the texts gathered from William Ely Johnson, even though the longest text that Johnson relayed was a myth, a genre where one would expect =*ani* to be prevalent. Volsin Chiki's texts consist of one myth and one tale. Both genres deal with events of the distant past. One would expect the frequency of =*ani* to be high in both. But Volsin Chiki's frequency of use pales in comparison to Sesostrie Youchigant's. His texts feature not only myths and tales

Table 7.15: Frequency of =*ani* by era

Years	Linguist	Speaker	% of sentences with = <i>ani</i>
1886	Albert Gatschet	William Ely Johnson	0
1910	John Swanton	Volsin Chiki	57.8
1933–1939	Mary Haas	Sesostrie Youchigant	95.8
2017	N/A	Donna & Elisabeth Pierite	0.9

but genres like personal narratives, which one assumes would have lower rates of =*ani*, yet =*ani* occurs in 95.8% of the sentences in Sesostrie Youchigant’s texts.

The only instances of =*ani* in Modern Tunica are three occasions where it is used when the speakers recite a sentence from one of Sesostrie Youchigant’s texts. It is not surprising that there would be a relatively low frequency of =*ani* in a conversation in the personal narrative genre (with some historical narrative). However, many sentences in the Modern Tunica subcorpus seem to call for a reportative or quotative evidential where it is absent.

The lack of =*ani* in Modern Tunica is understandable: it has no analog in any of the European languages to which Tunica speakers shifted and is not taught in the Tunica language textbook 2023. However, the increase in usage between speakers of Historical Tunica over time is puzzling. If Tunica followed the dissipation model of obsolescence, it would be expected that a morpheme that occurs as frequently as =*ani* does in Sesostrie Youchigant’s speech would be resistant to falling out of use. And though the concentration model would support the increased use of =*ani*, as it is a marked feature, a concentration model would not explain how the use of

an evidential sprung from nothing sometime between 1886 and 1910.⁴ The more likely scenario is that William Ely Johnson was unfamiliar with *=ani*. Sesostrie Youchigant called Johnson's Tunica fluency, and perhaps this is one instance where Johnson lacked knowledge.

Volsin Chiki's use of *=ani* is less frequent than Sesostrie Youchigant's; he seems to use it sometimes every sentence and sometimes every few sentences. There are two possible explanations for Volsin Chiki's sporadic use of *=ani*.

The first explanation is that *=ani* may be required only periodically during storytelling, not necessarily in every sentence and that Sesostrie Youchigant overgeneralized the use of *=ani* in his storytelling.

The other possibility is that Volsin Chiki omits *=ani* from time to time simply because he forgets. On at least two occasions in the texts, he repeats a phrase, adding an *=ani* in the second utterance, as in examples 34 and 35.

- (34) hiⁿyati niχsara ili lot-unixki.
 hinyatihch nisara ili lɔta-unihki
 then young.person two run-BE.3md
 The two youths ran off.

- (35) hiⁿyati niχsara ili lot-unixkeni.
 hinyatihch nisara ili lɔta-unihki=*ani*
 then young.person two run-BE.3md=*QUOT*
 The two youths ran off.

There is also a significant difference in the frequency of *=ani* between the two texts he recited, as seen in Table 7.16.

4. There is no indication that *=ani* is derived from a structure present in the Gatschet/William Ely Johnson texts; it is simply absent.

Table 7.16: Frequency of =*ani* per text from Volsin Chiki

Title of text	Sentences with = <i>ani</i>	Total sentences	Percentage
Rattlesnake myth	36	74	48.65
Flood myth	64	99	64.65

This discrepancy is due to the larger number of direct quotations in the Rattlesnake Myth compared to the Flood Myth. Since characters' speech describes events in which they are participating, =*ani* would not be appropriate in these cases. If a character in a story speaks several sentences, there is no need to append an =*ani* until the *nikɔni* 'he said' or *nikatɛni* 'she said' that often comes at the end of a character speaking. It is not obligatory to include a 'he said' or 'she said' after each direct quote, and sometimes several turns of conversation can take place with no =*ani* at their end. In the Rattlesnake Myth, for instance, Volsin Chiki quotes the protagonist directly for 13 sentences with no *nikɔni* 'he said' or *nikatɛni* 'she said' at the end.

The hypothesis for =*ani* was that myths and tales would have the highest frequency of use, followed by historical narratives. Ethnological narratives would have significantly fewer instances, and personal narratives the fewest. Table 7.17 shows the prevalence of =*ani* by genre for each speaker.⁵

The stark disparity between speakers in their use of =*ani* within the same genre is evident. Myths provide a genre to which all three speakers of Historical Tunica contribute texts. In this genre, =*ani* ranges from no occurrences at all in Gatschet's

5. Dividing the data only by genre without splitting it by speaker as well would do a poor job of showing the relative frequencies by genre because of the wide discrepancy in usage between speakers and the fact that not all eras contribute data from all genres.

Table 7.17: Percentage of sentences using =*ani* by genre and era

Genre	Gatschet	Swanton	Haas	Modern Tunica
Historical	0	–	98.5	–
Myths	0	57.8	97.7	–
Animal Tales	–	–	97.7	–
Tales	–	–	96.6	–
Ethnological	0	–	81.5	–
Personal Narratives	0	–	61.0	0.9

texts gathered from William Ely Johnson, to 57.8% of sentences in texts Swanton gathered from Volsin Chiki, to nearly every sentence containing an =*ani* in the texts Haas gathered from Sesostrie Youchigant, there could hardly be a wider disparity between the three.

The Personal Narratives genre shows a wide discrepancy between Haas’s data on the one hand and Gatschet and Modern Tunica data on the other; neither Gatschet’s consultant nor the speakers of Modern Tunica in the corpus use =*ani* except, in the latter case, in a recitation of a sentence from a Sesostrie Youchigant text.

Historical Narratives have the highest frequency of use, followed by Myths, Animal Tales and Tales,⁶ Ethnological information, and finally Personal Narratives.

Historical Narratives have a higher than expected use of =*ani*. As with the difference in use of =*ani* in Volsin Chiki’s two texts, the differentiator among the genres that do not pertain to the present circumstances—namely Myths, Tales, Animal Tales, and Historical Narratives—is the use of direct quotations. The hypothesis was that Historical Narratives would have a lower occurrence of =*ani* because some

6. The Tales and Animal Tales genres are kept separate only to preserve the categories that Haas used. Animals in Animal Tales speak—and often take human form at night—so there is little narrative distinction between the two genres.

of the events discussed would relate to the current state of affairs, on which the speaker would comment. This turns out to be exceedingly rare in the documentation. Moreover, historical narratives focus more on events affecting groups of people—tribes and European nationalities—than individuals. As such, there is no direct quotation, leading to a higher use of *=ani*.

Ethnological information, as expected, is lower than the first four categories, as speakers may have participated in or seen these cultural practices firsthand. Personal narratives also show a considerable drop-off in sentences containing *=ani*. This stands to reason; personal narratives focus mostly on events and situations in which the speaker participated.

The data on *=ani* reveals variation in Tunica speech in obsolescence. One speaker doesn't use it at all, and the two other speakers use it at very different rates. The quotative evidential may have been particularly susceptible to wildly varying use because of its lack of a true analog in any of the European languages in the area. Its current lack of use in Modern Tunica is attributable to both its omission in the Tunica Textbook and its lack of an English analog.

7.5 Elision in alienable possession

As described in Section 6.5, Haas (1940) distinguished between alienable and inalienable possessors in Tunica, with the difference being the presence or absence of a morpheme-final *hk*. However, in all of the Historical Tunica subcorpora, there are instances where one would expect the inalienable prefix but find a reduced form in-

Table 7.18: Elision of alienable possessors

Era	Elided	Total	% elided
Gatschet	7	120	5.8
Swanton	6	20	30.0
Haas	191	639	29.9
Modern Tunica	0	0	0

stead, without the *hk*. The frequency of the omission of the *hk* varies by subcorpus, as seen in Table 7.18.

The Gatschet subcorpus has a low rate of possessive prefix elision on alienable roots, whereas the Swanton and Haas subcorpora have almost identical elision rates of 30%. The Modern Tunica subcorpus lacks any alienable nouns with possessive prefixes; because the conversation revolves around family, the incidence of inalienable possession is high, but alienable prefixes are absent.

While the Swanton and Haas subcorpora show an increase in elided forms compared to the Gatschet subcorpus, this is probably not an indication of a merger between the two forms. In such a merger, one would expect possessive prefixes on vowel-initial (or glottal stop-initial) roots to all begin patterning one way or the other. However, only once in the corpus does a possessive prefix reduce before a vowel-initial alienable root (example 36), and the alienable form of the possessive is never found in the corpus before a root that Haas deems inalienable.

Table 7.19: Verb location by era

Era	Verb-final	No verb	Postposition-final	Adverb-final	Noun-final	Other
Gatschet	86.30%	4.08%	1.46%	1.75%	3.79%	2.62%
Swanton	90.17%	2.89%	2.31%	0.00%	3.47%	1.16%
Haas	87.72%	3.29%	3.83%	3.03%	1.56%	0.57%
Modern Tunica	37.79%	51.15%	0.46%	0.46%	3.23%	6.91%
Average	84.80%	6.13%	3.34%	2.62%	1.95%	1.15%

- (36) chohak yatich uonisema hotu
 choha=ku hinyatihch **uhk**-oni=sema hotu
 chief=3ms then **his.AL**-person=3mp all

sikuχpanu.
 sihk-kuhpa-n-uhki
 them-gather-CLASSII.CAUS-BE.3ms

At once the chief's people came together.

7.6 Word order

Tunica word order, described in Section 6.6, is well-illustrated by Mary Haas's examples of Tunica word order in Table 6.8. To assess a possible shift in word order, the corpus was divided up into sentence types by era: verb-final sentences, sentences with no verb (usually equatives without copula), noun-final (usually specifying a verb-marked object or subject explicitly), postposition-final phrases, adverb-final phrases.

Table 7.19 shows that the distribution of sentences by word order remains stable. English and French SOV word order does not seem to have precipitated a significant change in Tunica word order during the period when the language was documented.

It is only in the Modern Tunica corpus where there is a shift, not of ordering, per se, but omission. The number of sentences without verbs is drastically higher

in the Modern Tunica subcorpus. The explanations for this lie partly in the fact that the text in the subcorpus is a dialogue rather than the monologues of the other subcorpora. But it is also due to a greater reliance on zero-copula sentences. Tunica can optionally omit the copula *uhki*, and taking advantage of this allows for the easy creation of sentences without the need to conjugate a verb. Predicative sentences like those in examples 37 and 38 are common in the Modern Tunica subcorpus text.

(37) entiku uwi tawilaku lapusht'ε.
 in-eti=ku uwi ta-wila=ku lapu-sht'ε
 our.INAL-friend=3ms he AGNZ-weave=3ms good-very
 Our friend, he is a very good weaver.

(38) hon imapa taworuni luhchi yoroni
 hon ima-pa ta-woru-ni luhchi yoroni
 yes I-too AGNZ-study-CLASSII.CAUS language Tunica
 Yes, I am a Tunica language teacher also.

The content of the examples above could be expressed as “Our friend weaves very well” and “I teach the Tunica language as well”, but both interlocutors in the text availed themselves of the zero-copula forms.

This fits with the theory of simplification that exists both in obsolescence and revival; simpler structures with fewer morphemes are more frequent than their more complex alternatives.

7.7 Conclusion

This chapter analyzed of various aspects of Tunica that theories on obsolescence and revival suggest are areas where change might be expected. The data gathered did reveal some evidence of change in line with existing theory: there is some

evidence that verbs of non-volitional action could once only be inflected with the marked transimpersonal, but that over time they were reanalyzed to inflect with the unmarked intransitive.

There is also evidence that Sesostrie Youchigant was more unsure about the gender of certain words than either of the two earlier speakers; the Haas subcorpus was the only one that used different genders for the same inanimate object. Looking at the possible leveling of gender over time, however, revealed only that any leveling of non-animate plural gender to be exclusively feminine had already taken place by the time Gatschet documented Tunica.

Investigating the quotative evidential =*ani* contradicted the prediction informed by existing theories of obsolescence. The evidential's contribution to utterance (that the speaker did not have direct knowledge of the utterance) can often be inferred from context, and it has no direct parallel in either of the dominant languages that Tunica shifted to. However, the earliest subcorpus contains no evidence of =*ani* being used at all. Usage gradually increases through the Swanton and Haas subcorpora, where it is used very liberally. In revival, however, the theory correctly predicts that the usage would drop off.

Based on the data analyzed, it is challenging to fit Tunica into any coherent and consistent model of language obsolescence. Obsolescing languages are unique in their circumstances; these circumstances affect how languages change during obsolescence. What is borne out in the Tunica context is idiosyncratic variability between speakers that does not point to a uniform direction of language change.

Dorian's words about East Sutherland Gaelic are apt for obsolescent Tunica as well: "a...patchwork of inconsistencies...haphazardly distributed over villages, speakers, and occasions" (Dorian 1973:414). That said, the particular types of inconsistencies between speakers may reveal some evidence of the speakers' relative command of the language. Order lurks even in the chaotic processes of obsolescence.

In revival, the language has undergone many of the processes that theories of obsolescence predict: a decrease in forms used, a move toward more regular and less marked forms, an increase in the use of analytic structures (zero-copula constructions rather than inflected verb constructions), and use of forms with fewer morphemes (greater use of stative verbs).

As more speakers begin to speak increasingly fluent Tunica in ever-expanding domains, the language should again increase in variety of form and function, reversing the stylistic and lexical contraction during obsolescence.

7.7.1 Future study

The corpus of documented Modern Tunica will continue to grow as revival effort progresses. Furthermore, the release of the second Tunica Textbook that is now in progress may impact how Modern Tunica is spoken. A subsequent study examining diachronic change in Modern Tunica itself that makes use of new data at a time of growing pedagogical materials would serve as an interesting complement to the present work.

Transcribing the handwritten interlinear texts that Swanton gathered from Sesostrie Youchigant in 1930 and 1931 would expand (and force a renaming of) the Haas

subcorpus. With the addition of the new data to the existing corpus, the areas of study featured in this dissertation could be re-examined.

Tagging thematic roles and incorporating them into the interlinear text data would enable a more detailed study of Tunica gender marking, as well as any other aspects of the language dependent on semantic information.

This examination of language change in Tunica aimed for more than simply determining whether Tunica language change in obsolescence and revival fit neatly within a theoretical framework. The code developed for this study can be adapted by KYLY's members to examine any number of features of Tunica. This undertaking has generated methods and procedures that will allow quicker and more complete investigations into the corpus of documented Tunica. These methods and procedures can be developed further to facilitate the search for information about the language—and by extension, the culture—of a community that is actively reclaiming both.

Appendix A: Python script

```

"""
This script imports a FLEx-exported interlinear file, parses the XML, and
analyzes Tunica verb inflection.
"""

import csv                # for exporting to CSV
from statistics import mean # for calculating averages of lists
from lxml import etree as ET # for parsing XML

# set the path for the Tunica texts
tunttexts = 'tunttexts.xml'

tree = ET.parse(tunttexts) # parse the XML file
root = tree.getroot()      # get the root node

# to select a single story to analyze
storyfilter = ""

# create an empty list that will be used to create a CSV at the final step
csv_list = []

# loop over each text
for interlinear in root.iter('interlinear-text'):
    # initialize a list that will contain the number of morphemes in each
    # sentence in the current text
    morphlen = []
    # initialize a list that will contain the number of words in each sentence
    # in the current text
    wordlen = []
    # store the name of the text
    story_title = interlinear.find('item[@type="title"]').text

    # check if the story passes the filter
    if storyfilter in story_title:
        # set all counters to zero at the beginning of each story
        uhkicount = 0
        verbcount = 0
        mutableuracount = 0
        immutableuracount = 0
        mutableunacount = 0
        immutableunacount = 0
        compusacount = 0
        habusacount = 0

```

```

compuwacount = 0
habuwacount = 0
habicount = 0
habiicount = 0
compicount = 0
compicount = 0
stativecount = 0
# independent auxiliary verbs
auxcount = 0

# create an iterative that finds every instance of a sentence in the
# text and loop over each
for words in interlinear.iter('words'):

    # get individual morphemes for the current sentence
    morphs = words.findall('./morph')
    # get the words for the current sentence
    word = words.findall('./word/item[@type="txt"]')
    # for each word in the sentence...
    for haasword in words.iter('word'):
        # create a list of all morphemes in the word
        morphlist = []
        # set the habitual theme boolean to false
        hab_boolean = False

        # if the item isn't punctuation
        if haasword.find("item[@type='txt']") is not None:
            morphstuff = \
                haasword.findall('./morphemes/morph/item[@type="txt"]')
            for morphaloo in morphstuff:
                # append the text of the morpheme to a list
                morphlist.append(morphaloo.text)
            # look at the part of speech of each morpheme. two verb
            # types can only be reliably categorized by their part of
            # speech
            for part_of_speech in \
                haasword.findall(
                    './morphemes/morph/item[@type="msa"]'):
                # if the word contains a stative verb stem, increment
                # the stative verb counter
                if part_of_speech.text.startswith('v.st'):
                    stativecount += 1
                # if the word contains a "chameleon verb" part of
                # speech, increment the independent auxiliary counter
                elif part_of_speech.text == 'v.chameleon':
                    auxcount += 1

            # if the part of speech is a primary verbal affix, it
            # will start with a v. morphemes matching 'v:Any',
            # however, may sometimes appear on non-verbs due to
            # tagging inconsistencies. For each word that contains
            # a true verbal inflection morpheme, increment the
            # verb counter
            if (part_of_speech.text.startswith('v') and

```

```

        part_of_speech.text != 'v:Any'):
    verbcount += 1
    # because the verb may have several affixes that
    # are each marked as verbal affixes, break after
    # the first one so a verb doesn't get
    # double-counted
    break

# look at canonical form of the morpheme for uhki to deal
# with a FLEx tagging issue. if the canonical form is
# 'uhki', increment the independent auxiliary counter
for cf in \
    haasword.findall(
        './morphemes/morph/item[@type="cf"]'):
    if cf.text == 'uhki':
        auxcount += 1
# look at the English gloss of each morpheme
for gls in \
    haasword.findall(
        './morphemes/morph/item[@type="gls"]'):

    # if the verb contains an -hk, it will be the habitual
    # form of whatever the main verbal inflection is, so we
    # set the habitual variable true for the word so that
    # the main inflection gets correctly marked as habitual
    if 'habitual.theme' in gls.text:
        hab_boolean = True
# completive class I
    if '.completive' in gls.text:
        compicount += 1
# inflectional -una
    if '.durative.una' in gls.text:
        if hab_boolean:
            mutableunacount += 1
        else:
            immutableunacount += 1
# inflectional -ura
    if '.durative.ura' in gls.text:
        if hab_boolean:
            mutableuracount += 1
        else:
            immutableuracount += 1
# class II
    if 'classII' in gls.text:
        if hab_boolean:
            habiicount += 1
        # special case for class II inflection with
        # built-in habitual
        elif gls.text == 'classII.hk.1s':
            habiicount += 1
        else:
            compiicount += 1
# inflectional -usa
    if 'toward_speaker' in gls.text:

```

```

        if hab_boolean:
            habusacount += 1
        else:
            compusacount += 1
    # inflectional -uwa
    if 'away_from_speaker' in gls.text:
        if hab_boolean:
            habuwacount += 1
        else:
            compuwacount += 1
    # habitual
    if '.habitual' in gls.text:
        habicount += 1
    # inflectional -uhki
    if '.uhki' in gls.text:
        uhkicount += 1

    # put the number of morphemes in the current sentence in a list
    morphlen.append(len(morphs))
    # put the number of words in the current sentence in a list
    wordlen.append(len(word))

if morphlen and wordlen:
    # append relevant data about the story to a CSV list
    csv_list.append([story_title,
                    sum(morphlen),
                    sum(wordlen),
                    len(morphlen),
                    round(mean(morphlen), 2),
                    round(mean(wordlen), 2),
                    round(mean(morphlen)/mean(wordlen), 2),
                    compicount,
                    compicount,
                    habicount,
                    habiicount,
                    uhkicount,
                    compusacount,
                    habusacount,
                    compuwacount,
                    habuwacount,
                    immutableunacount,
                    mutableunacount,
                    immutableuracount,
                    mutableuracount,
                    stativecount,
                    auxcount])

# open a file for writing
with open('story_file.csv', mode='w') as story_file:
    # make a csv writer
    story_writer = csv.writer(story_file,
                              delimiter=',',
                              quotechar='"',

```

```
        quoting=csv.QUOTE_MINIMAL)
# write the header row
story_writer.writerow(["story title",
                       "total morphemes",
                       "total words",
                       "number of sentences",
                       "morphemes/sentence",
                       "words/sentence",
                       "morphemes/word",
                       "completive class I",
                       "completive class II",
                       "hab class I",
                       "hab class II",
                       "-uhki",
                       "comp -usa",
                       "hab -usa",
                       "comp -uwa",
                       "hab -uwa",
                       "comp -una",
                       "hab -una",
                       "comp -ura",
                       "hab -ura",
                       "stative",
                       "aux"])

# assign each story's data to a CSV row
for row in csv_list:
    story_writer.writerow(row)
```

Appendix B: R script

```

# Tunica text analysis

# load necessary packages
library(interlinearR) # turns FLEx LIFT dictionaries and XML
                      # interlinear texts into dataframes

library(dplyr)       # for easier dataframe manipulation
library(purrr)       # for one instance of concatenating and sorting
library(stringr)     # for easier find and replace
library(tidyr)       # for drop_na

# variables storing the text_id for different types of texts
new_tunica_texts <- c(1, 2, 89)
gatschet_texts <- 3:10
swanton_texts <- 87:88
haas_texts <- 11:86

# create a list of the texts separated by era
author_text_list <- list(gatschet_texts,
                         swanton_texts,
                         haas_texts,
                         new_tunica_texts)

# variables for different genres of texts
myths <- c(8, 11:24, 87, 88)
tales <- c(25:32, 80, 84, 85, 72:74)
animal <- c(33:40, 75:79, 81)
histories <- c(7, 41:51, 86)
personal <- c(52:57, 3, 4, 6, 1, 2)
ethnological <- c(58:71, 5, 9, 10, 82, 83, 89)

# creates a table of stats about the corpus
create_corpus_stats_table <- function() {
  gatschet_df <- morpheme_df[grep("^G\\d\\.\"", morpheme_df$title.en),
    names(morpheme_df)]
  swanton_df <- morpheme_df[grep("^S\\d\\.\"", morpheme_df$title.en),
    names(morpheme_df)]
  haas_df <- morpheme_df[grep("^\\d\\d\"", morpheme_df$title.en),
    names(morpheme_df)]
  newnica_df <- morpheme_df[grep("^Y\\d\"", morpheme_df$title.en),
    names(morpheme_df)]
  stats_table <- tibble(Name = character(), Texts = integer(),
    Sentences = integer(), Words = integer(), Morphemes = integer())
}

```

```

linguist_list <- list("Gatschet", "Swanton", "Haas", "New Tunica")
ling_df_list <- list(gatschet_df, swanton_df, haas_df, newnica_df)
j <- 0
for (i in ling_df_list) {
  j <- j + 1
  stats_table <- stats_table %>%
    add_row(Name = linguist_list[[j]],
           Texts = length(unique(i$text_id)),
           Sentences = length(unique(i$sentence_id)),
           Words = length(unique(i$word_id)),
           Morphemes = length(i$text_id))
}
stats_table <- stats_table %>% add_row(Name = "Total",
  Texts = sum(stats_table$Texts),
  Sentences = sum(stats_table$Sentences),
  Words = sum(stats_table$Words),
  Morphemes = sum(stats_table$Morphemes))
}

# creates a table of corpus stats pertaining to genre
create_genre_stats_table <- function() {
  myths_df <- subset(morpheme_df, (morpheme_df$text_id %in% myths))
  tales_df <- subset(morpheme_df, (morpheme_df$text_id %in% tales))
  animal_df <- subset(morpheme_df, (morpheme_df$text_id %in% animal))
  histories_df <- subset(morpheme_df, (morpheme_df$text_id %in% histories))
  personal_df <- subset(morpheme_df, (morpheme_df$text_id %in% personal))
  ethnological_df <- subset(morpheme_df, (morpheme_df$text_id %in%
    ethnological))
  genre_stats_table <- tibble(
    Genre = character(),
    Texts = integer(),
    Sentences = integer(),
    Words = integer(),
    Morphemes = integer())
  genre_list <- list("Myths",
                    "Tales",
                    "Animal Tales",
                    "Historical",
                    "Personal Narratives",
                    "Ethnological")
  genre_df_list <- list(myths_df,
                       tales_df,
                       animal_df,
                       histories_df,
                       personal_df,
                       ethnological_df)

  j <- 0
  for (i in genre_df_list) {
    k <- 0
    j <- j + 1
    genre_stats_table <- genre_stats_table %>% add_row(
      Genre = genre_list[[j]],
      Texts = length(unique(i$text_id)),
      Sentences = length(unique(i$sentence_id)),

```

```

Words = length(unique(i$word_id)),
Morphemes = length(i$text_id)
)
for (z in author_text_list) {
  k <- k + 1
  documenter <- linguist_list[[k]]
  author_genre <- subset(i, i$text_id %in% z)
  genre_stats_table <- genre_stats_table %>% add_row(
    Genre = documenter,
    Texts = length(unique(author_genre$text_id)),
    Sentences = length(unique(author_genre$sentence_id)),
    Words = length(unique(author_genre$word_id)),
    Morphemes = length(author_genre$text_id)
  )
}
}

# add a row with totals
add_total_row <- function(df_title) {
  tosum = seq(1, nrow(df_title), by = 5)
  df_title <- add_row(df_title,
    Genre = "Total",
    Texts = sum(df_title$Texts[tosum]),
    Sentences = sum(df_title$Sentences[tosum]),
    Words = sum(df_title$Words[tosum]),
    Morphemes = sum(df_title$Morphemes[tosum]))
  return(df_title)
}

# set paths
corpuspath <- "tuntxts.xml" # path to the interlinear texts file
dictpath <- "tundict.lift" # path to the LIFT dictionary file

# Import corpus

# Use interlineaR to convert the interlinear texts to a dataframe, choosing what
# to import.
# Here we specify Tunica as the vernacular language, and specify which Tunica
# morpheme information to import. We import:
# 1. how the morpheme appears in the text (txt)
# 2. the canonical form of the morpheme (cf)
# 3. which sense of a polysemic morpheme it is (hn)
#
# For the analysis language (English), we import:
# 1. the English gloss (gls)
# 2. the part of speech (msa)
corpus <- read.emeld(corpuspath,
  vernacular.languages = "tun",
  morphemes.vernacular.fields = c("txt", "cf", "hn"),
  morphemes.analysis.fields = c("gls", "msa"))

#Import dictionary

# Use interlineaR to import the Tunica dictionary

```



```

dictionary <- read.lift(dictpath,
  vernacular.languages="tun",
  simplify=TRUE)

# Merge the dictionary entries and the senses dataframe, the latter of which
# provides homonymous senses as well as semantic domains for each sense.
full_dict_entries <- merge(
  dictionary$entries[, -c(1)],
  dictionary$senses[, -c(1, 2)],
  by.x = "id",
  by.y = "lexem_id"
)

# Remove unnecessary columns from dictionary.
# The FLEx database for Tunica is not yet complete. For instance, it lacks
# semantic tagging, so including a semantic tagging column would result in a
# column of NA entries, making the dataframe harder to read
dict_entries <- full_dict_entries[, c(
  "id",
  "lexical-unit.tun",
  "gloss.en",
  "semantic-domain-ddp4",
  "morph-type",
  "grammatical-info.value")]

# Extract the morpheme dataframe from the corpus object (which is a dataframe of
# three dataframes)
corpmorph <- corpus$morphemes

corpwords <- corpus$words[corpus$words$text_id %in% 1:88, ]

# Merge the dictionary entries with the morphemes from the text if the text's
# canonical form matches the dictionary's lexical unit (form) AND the gloss is
# the same in the dictionary and the interlinear text (meaning). Merging on both
# prevents incorrect merging of entries
morpheme_df <- merge(
  corpmorph,
  dict_entries,
  by.x = c("cf.tun", "gls.en"),
  by.y = c("lexical-unit.tun", "gloss.en"),
  all.x = TRUE) %>% arrange(morphem_id)

# Merge with the word datatable so it can be seen at a glance what word a given
# morpheme belongs to.
morpheme_df <- merge(
  morpheme_df,
  corpus$words[, c("word_id", "txt.tun")],
  by = "word_id",
  all.x = TRUE,
  suffixes = c(".morph", ".full_word")) %>% arrange(morphem_id)

# normalize all words and morphemes lowercase
morpheme_df$txt.tun.full_word <- tolower(morpheme_df$txt.tun.full_word)
morpheme_df$txt.tun.morph <- tolower(morpheme_df$txt.tun.morph)

```

```

# for each word, create a column with all that word's morphemes and glosses
wordmerge <- morpheme_df %>%
  group_by(word_id) %>%
  reframe(cf.tun = toString(cf.tun),
          gls.en = toString(gls.en)) %>%
  distinct(word_id, .keep_all = TRUE)

# merge these columns back into the main database
morpheme_df <- merge(morpheme_df,
                    wordmerge,
                    by = "word_id",
                    all = TRUE,
                    suffixes = c(".morph", ".word"))

# create a list of all the stems and roots
just_stems_roots <- subset(morpheme_df,
                          (morpheme_df$type == "stem" |
                           morpheme_df$type == "root" |
                           morpheme_df$type == "bound stem"))

# list all stem morphemes for each word
word_stems <- just_stems_roots %>%
  group_by(word_id) %>%
  reframe(stems=toString(cf.tun.morph))

# merge those word stems back into the morpheme dataframe
morpheme_df <- merge(morpheme_df, word_stems, by = "word_id", all = TRUE)

# merge in a column containing the text's title
morpheme_df <- merge(morpheme_df,
                    corpus$texts[, c("text_id", "title.en")],
                    by = "text_id")

# merge in column containing the sentence the morpheme appears in
morpheme_df <- merge(morpheme_df,
                    corpus$sentences[, c("sentence_id", "gls.en")],
                    by = "sentence_id",
                    all.x = TRUE,
                    suffixes = c(".morph", ".sentence")) %>%
  rename(gls.en.sentence = gls.en) %>%
  dplyr::filter(text_id != 89)

# add columns with author and genre information
morpheme_df <- morpheme_df %>%
  mutate(era = case_when(
    text_id %in% gatschet_texts ~ "Gatschet",
    text_id %in% swanton_texts ~ "Swanton",
    text_id %in% haas_texts ~ "Haas",
    text_id %in% new_tunica_texts ~ "New Tunica"
  )) %>%
  mutate(era_num = case_when(
    text_id %in% gatschet_texts ~ 1L,
    text_id %in% swanton_texts ~ 2L,

```

```

text_id %in% haas_texts ~ 3L,
text_id %in% new_tunica_texts ~ 4L
)) %>%
mutate(genre = case_when(
  text_id %in% myths ~ "Myths",
  text_id %in% tales ~ "Tales",
  text_id %in% animal ~ "Animal Tales",
  text_id %in% histories ~ "Historical",
  text_id %in% personal ~ "Personal Narratives",
  text_id %in% ethnological ~ "Ethnological"
))

# get the number of sentences in each text
sentences_per_text <- morpheme_df %>%
  group_by(text_id, title.en, era, genre, era_num) %>%
  summarise(sentences_in_text = length(unique(sentence_id)))

# creates a dataframe with just words with the articular ta- prefix for
# examination of ta-'s pairing with gender-number affixes
ta_dataframe <- morpheme_df[grep("^1704$", morpheme_df$id),
  c("word_id", "cf.tun.morph", "cf.tun.word", "txt.tun.full_word",
    "stems", "era", "genre")]

# create a dataframe of all noun prefixes (articular ta-, but also possessive
# prefixes, which do not cooccur with ta-)
nounprefix_dataframe <- morpheme_df[grep("NounPrefix", morpheme_df$msa.en),
  c("word_id", "cf.tun.morph", "cf.tun.word",
    "txt.tun.full_word", "era", "genre", "text_id")]

# create a dataframe of agentive ta-, which nominalizes verbs, and also does not
# cooccur with articular ta-
agentive_dataframe <- morpheme_df[grep("^2254$", morpheme_df$id),
  c("word_id", "cf.tun.morph", "cf.tun.word",
    "txt.tun.full_word", "era", "genre", "text_id")]

# dataframe with all noun prefixes
prefixes_dataframe <- rbind(nounprefix_dataframe, agentive_dataframe)

# create a dataframe of all gender-number affixes
gns_dataframe <- morpheme_df[grep("GNS", morpheme_df$msa.en),
  c("word_id", "cf.tun.morph", "txt.tun.full_word",
    "cf.tun.word", "stems", "era", "era_num", "text_id")]

# add a gender column to categorize each gender-number affix by gender
gns_dataframe$gender <- with(gns_dataframe,
  ifelse((cf.tun.morph == "-sinima" |
    cf.tun.morph == "-hchi"), "F", "M"))

# add a number column to categorize each gender-number affix by number
gns_dataframe$number <- with(gns_dataframe,
  ifelse(grepl("-sinima|-sema", cf.tun.morph), "P",
    ifelse(grepl("-hchi|-ku", cf.tun.morph), "S", "D")))

```

```

# create a list of all unique words with gender-number suffixes
gns_by_gender_number_era <- gns_dataframe %>%
  reframe(.by = c(gender, number, stems, era),
    paste(unique(cf.tun.word), collapse = "; ")) %>%
  count(stems, gender, number, era, sort = TRUE)

# take the list of unique words with gender-number marking and
gns_gender_number_era <- gns_by_gender_number_era[, c("gender", "number",
  "stems", "era", "n")] %>%
  group_by(stems, era, number) %>%
  reframe(gender=toString(gender), n, stems,
    era, number) %>%
  mutate(across(c(gender), ~ map_chr(strsplit(., ", "),
    ~ toString(sort(.x))))) %>%
  group_by(stems, gender, era, number) %>%
  summarise(n = sum(n))

# find stems that appear in multiple subcorpora
gns_gen_num_era_dupes <- gns_gender_number_era %>%
  group_by(stems) %>%
  dplyr::filter(n(>1))

# find the number of times each word appears with each gender and number
gns_counts <- gns_dataframe %>%
  count(gender, number, stems, cf.tun.morph, sort=TRUE)

# find the number of unique stems that take each gender-number combination
gns_stem_number_counts <- gns_by_gender_number_era %>% count(gender, number)

# find the number of times each gender-number combination appears over the
# entire corpus
gns_gen_num_counts <- gns_dataframe %>% count(gender, number)

# find the number of times each gender-number combination appears over each
# subcorpus
gns_gen_num_era_counts <- gns_dataframe %>% count(gender, number, era)

# dataframe for cooccurring ta- and GNA
ta_gns_dataframe <- merge(ta_dataframe, gns_dataframe,
  by = c("word_id", "txt.tun.full_word",
    "stems", "cf.tun.word"),
  suffixes = c(".prefix", ".suffix"))

# a dataframe of all words that start with articular ta- but do not have a
# gender-number ending
ta_only <- merge(ta_dataframe,
  ta_gns_dataframe[, c("cf.tun.morph.suffix", "word_id")],
  by = "word_id", suffixes = c(".ta", ".gns"), all = TRUE) %>%
  dplyr::filter(is.na(cf.tun.morph.suffix))

# a dataframe of agentive nouns without a gender-number ending
agentive_only <- merge(agentive_dataframe,
  ta_gns_dataframe[, c("cf.tun.morph.suffix", "word_id")],

```

```

    by = "word_id", suffixes = c(".ta", ".gns"),
    all = TRUE) %>%
    dplyr::filter(is.na(cf.tun.morph.suffix))

# agentives with a gender-number ending
agentive_gns <- merge(agentive_dataframe, gns_dataframe,
  by = c("word_id", "txt.tun.full_word", "cf.tun.word"),
  suffixes = c(".prefix", ".suffix"))

# words beginning with "ta" where "ta" isn't marked as a separate agentive or
# articular morpheme
extra_tas <- morpheme_df[ grep("^ta[A-Za-z]", morpheme_df$cf.tun.word),
  c("word_id", "cf.tun.morph", "cf.tun.word",
    "txt.tun.full_word", "stems")]

# words beginning with non-discrete "ta" that have a gender-number suffix
extra_ta_gna <- merge(extra_tas, gns_dataframe[, c("cf.tun.morph", "word_id")],
  by = "word_id", suffixes = c(".ta", ".gns"), all.x = TRUE) %>%
  dplyr::filter(!is.na(cf.tun.morph.gns))

# nouns that take a gender-number suffix without a corresponding //ta-//
gns_only <- merge(ta_dataframe[, c("word_id", "cf.tun.morph")],
  gns_dataframe, by = c("word_id"),
  suffixes = c(".prefix", ".suffix"), all = TRUE) %>%
  dplyr::filter(is.na(cf.tun.morph.prefix)) %>%
  select(-cf.tun.morph.prefix)

# nouns containing any prefix (articular, agentive, or possessive) but no
# gender-number suffix
prefix_only <- merge(nounprefix_dataframe,
  gns_dataframe[, c("cf.tun.morph", "word_id",
    "era", "text_id")],
  by = c("word_id", "era", "text_id"),
  suffixes = c(".prefix", ".gns"), all = TRUE) %>%
  dplyr::filter(is.na(cf.tun.morph.gns))

# number of nouns with prefixes from each subcorpus
all_prefix_by_era <- nounprefix_dataframe %>% count(era)

# number of nouns with prefixes from each subcorpus-genre pairing
all_prefix_by_genre_era <- nounprefix_dataframe %>% count(era, genre)

# number of nouns with a prefix but no gender-number suffix by subcorpus
prefix_only_by_era <- prefix_only %>% count(era)

# number of nouns with a prefix but no gender-number suffix by subcorpus and
# genre
prefix_only_genre_era <- prefix_only %>% count(era, genre)

# ratio of prefixed nouns without gender-number suffixes to total prefixed nouns
# by genre and subcorpus
prefix_only_genre_era_ratio <- merge(prefix_only_genre_era,
  all_prefix_by_genre_era,

```

```

                                by = c("era", "genre"))
# add percentage column
prefix_only_genre_era_ratio$percentage <- with(prefix_only_genre_era_ratio,
  n.x / n.y)

# ratio of prefixed nouns without gender-number suffixes to total prefixed nouns
# by subcorpus
prefix_ratio <- merge(prefix_only_by_era, all_prefix_by_era, by = "era")
# add percentage column
prefix_ratio$percentage <- with(prefix_ratio, n.x / n.y)

# a dataframe of all words that have a gender-number ending but no articular ta-
# or possessive prefix
gns_noprefixes <- merge(nounprefix_dataframe, gns_dataframe,
  by = c("word_id", "txt.tun.full_word"),
  suffixes = c(".prefix", ".suffix"), all = TRUE) %>%
  dplyr::filter(is.na(cf.tun.morph.prefix)) %>%
  select(-cf.tun.morph.prefix)

# count occurrences of a number of noun prefix - noun suffix
# combinations

# count of occurrences of each stem with article ta- but no gender number affix
ta_only_counts <- ta_only %>% count(cf.tun.morph, stems, sort=TRUE)

# count of occurrences of each stem with article ta- but no gender number affix
# by subcorpus
ta_only_era_counts <- ta_only %>% count(cf.tun.morph, stems, era, sort=TRUE)

# count of occurrences article ta- but no gender number affix by subcorpus
ta_only_by_era <- ta_only %>% count(era, sort = TRUE)

# count of occurrences article ta- but no gender number affix by genre
ta_only_by_era_genre <- ta_only %>% count(era, genre)

# count of all nouns with articular ta- by subcorpus
ta_by_era <- ta_dataframe %>% count(era)

# count of all nouns with articular ta- by subcorpus and genre
ta_by_genre_era <- ta_dataframe %>% count(era, genre)

# ratio of unmarked definite nouns to all definite nouns by subcorpus
ta_ratio <- merge(ta_only_by_era, ta_by_era, by = "era")
# add percentage column
ta_ratio$percentage <- with(ta_ratio, n.x / n.y)

# ratio of unmarked definite nouns to all definite nouns by subcorpus and genre
ta_genre_ratio <- merge(ta_only_by_era_genre, ta_by_genre_era,
  by = c("genre", "era"))
# add percentage column
ta_genre_ratio$percentage <- with(ta_genre_ratio, n.x / n.y)

# count of occurrences of gender-marked definite nouns by stem
ta_gns_counts <- ta_gns_dataframe %>%

```

```

count(cf.tun.morph.prefix, stems, cf.tun.morph.suffix,
      sort=TRUE)

# count of occurrences of gender-marked definite nouns by gender
ta_gns_gender_counts <- ta_gns_dataframe %>%
  count(cf.tun.morph.prefix, stems,gender, sort = TRUE)

# count of occurrences of all gender-marked nouns by gender
gns_gender_counts <- gns_dataframe %>%
  count(stems, gender, sort = TRUE)

# count of occurrences of all gender-marked nouns by gender, split by subcorpus
gns_gender_counts_era <- gns_dataframe %>%
  count(stems, gender, era, sort = TRUE)

# combine words that show up as both genders within a subcorpus into one row for
# easy identification
gns_gender_counts_era_combined <- gns_gender_counts_era %>%
  group_by(stems, era) %>%
  reframe(gender=toString(gender), n,
          stems, era) %>%
  mutate(across(c(gender),
                ~ map_chr(strsplit(., ", "),
                ~ toString(sort(.x)))))) %>%
  group_by(stems, gender, era) %>%
  summarise(n = sum(n))

# create a table of just the stems that appear as both genders within a
# subcorpus
multi_gender_by_era <- gns_gender_counts_era_combined[ grepl(
  "F", gns_gender_counts_era_combined$gender) &
  grepl("M", gns_gender_counts_era_combined$gender), ]

# combine words that show up as both genders, irrespective of subcorpus
gns_gender_counts_combined <- gns_gender_counts %>%
  group_by(stems) %>%
  reframe(gender=toString(gender), n, stems) %>%
  mutate(across(c(gender),
                ~ map_chr(strsplit(., ", "),
                ~ toString(sort(.x)))))) %>%
  group_by(stems, gender) %>%
  summarise(n = sum(n))

# table with the number of occurrences for each stem that is gender-marked as
# both genders, across subcorpora, still separated by gender
all_gns_both_gender_markings <- gns_gender_counts %>%
  group_by(stems) %>%
  dplyr::filter(n()>1) %>% arrange(stems)

# number of occurrences for for each stem for definite nouns that are
# gender-marked as both genders
both_gender_markings <- ta_gns_gender_counts %>%
  group_by(stems) %>%
  dplyr::filter(n()>1) %>%

```

```

arrange(stems)

# combines stems that appear as both masculine and feminine into single rows
ta_gns_gender_counts_combined <- ta_gns_gender_counts %>%
  group_by(stems) %>%
  reframe(gender=toString(gender), n, stems) %>%
  mutate(across(c(gender),
    ~ map_chr(strsplit(., ", "),
    ~ toString(sort(.x)))) %>%
  group_by(stems, gender) %>%
  summarise(n = sum(n))

# number of occurrences across subcorpora for definite nouns with each
# gender-number affix
ta_anygns_counts <- ta_gns_dataframe %>%
  count(cf.tun.morph.prefix, cf.tun.morph.suffix, sort=TRUE)

# -ani analysis
# create a dataframe of all occurrences of -ani
ani_dataframe <- morpheme_df[ grep("^2995$", morpheme_df$id), ]

# datatable with the number of sentences in each text containing an -ani, and
# the number of sentences in each text
ani_text_count <- merge(ani_dataframe %>%
  summarise(ani_sentences = n_distinct(sentence_id),
    .by = text_id),
  sentences_per_text, all = TRUE) %>%
  replace(is.na(.), 0)

# add a column with the percentage of sentences containing an -ani
ani_text_count$percentage <- with(ani_text_count, ani_sentences /
  sentences_in_text*100)

# percentage of ani per sentence across entirety of an era
ani_era_percentages <- ani_text_count %>%
  summarise("number_ani" = sum(ani_sentences),
    "number_sentences" = sum(sentences_in_text),
    .by = era) %>%
  mutate(percentage = number_ani / number_sentences*100)

# percentage of ani per sentence across genres
ani_genre_percentages <- ani_text_count %>%
  summarise("number_ani" = sum(ani_sentences),
    "number_sentences" = sum(sentences_in_text),
    .by = genre) %>%
  mutate(percentage = number_ani / number_sentences*100)

# percentage of ani per sentence across genres and eras
ani_era_genre_percentages <- ani_text_count %>%
  summarise("number_ani" = sum(ani_sentences),
    "number_sentences" = sum(sentences_in_text),

```



```

        .by = c("genre", "era", era_num)) %>%
mutate(percentage = number_ani / number_sentences*100) %>%
arrange(genre, era_num)

# list of all inalienably possessed nouns (marked by an asterisk in FLEx)
inal_roots <- morpheme_df[ grepl("\\*", morpheme_df$cf.tun.morph), ]

# get all alienable possessive prefixes
possessive_prefix_dataframe <- morpheme_df[ grepl("\\.alienable",
                                                morpheme_df$gls.en.morph), ]

# get all inalienable possessive prefixes
inal_pos_prefixes <- morpheme_df[ grepl("\\.inalienable",
                                                morpheme_df$gls.en.morph), ]

# find occurrence where ta- is used before an inalienably possessed noun (not
# expected)
ta_inal <- merge(inal_roots, ta_dataframe, by = "word_id")

# find occurrences where the canonical form of the inalienable prefix does not
↪ match the surface form of the inalienable prefix
inal_pos_elision <- inal_pos_prefixes[(inal_pos_prefixes$cf.tun.morph !=
                                        inal_pos_prefixes$txt.tun.morph), ]

# get all possessive alienable prefixes that drop either k or an hk from the
# expected canonical form (g is included in addition to k due to intervocalic
# voicing of stops, which is reflected in the two phonetic subcorpora [Gatschet
# and Swanton])
possessive_prefix_elision <- possessive_prefix_dataframe[(
  possessive_prefix_dataframe$cf.tun.morph !=
  possessive_prefix_dataframe$txt.tun.morph) & !grepl("g|k",
  possessive_prefix_dataframe$txt.tun.morph) & grepl("\\.alienable",
  possessive_prefix_dataframe$gls.en.morph), ]

# number of alienable possessive prefixes by subcorpus
pos_era_count <- possessive_prefix_dataframe %>% count(era)
# number of elided alienable possessive prefixes by subcorpus
pos_era_elide_count <- possessive_prefix_elision %>% count(era)
# merge the two tables to compare number elided vs. total number
pos_era_elision_table <- merge(pos_era_elide_count, pos_era_count, by = "era")

# number of alienable possessive prefixes by genre
pos_genre_count <- possessive_prefix_dataframe %>% count(genre)
# number of elided alienable possessive prefixes by genre
pos_genre_elide_count <- possessive_prefix_elision %>% count(genre)
# merge the two tables to compare number elided vs. total number
pos_genre_elision_table <- merge(pos_genre_elide_count, pos_genre_count,
                                by = "genre")

```

```

# number of alienable possessive prefixes by genre and subcorpus
pos_eg_count <- possessive_prefix_dataframe %>% count(era, genre)
# number of elided alienable possessive prefixes by genre and subcorpus
pos_eg_elide_count <- possessive_prefix_elision %>% count(era, genre)
# merge the two tables to compare number elided vs. total number
pos_eg_elision_table <- merge(pos_eg_elide_count, pos_eg_count,
                             by = c("era", "genre"))

possessive_prefix_elision %>% summarise(n(), .by = era) %>% replace(is.na(.), 0)

possessive_prefix_elision %>%
summarise(n(), .by = genre) %>%
replace(is.na(.), 0)

# list of stems by the number of times they have elided alienable prefixes
elided_stems <- possessive_prefix_elision %>%
  summarise("occurrences" = n(), .by = stems) %>%
  arrange(desc(occurrences))

# list of stems by the number of times they have elided alienable prefixes, by
# subcorpus
elided_stems_by_corpus <- possessive_prefix_elision %>%
  summarise("occurrences" = n(), .by = c(stems, era)) %>%
  arrange(desc(occurrences))

# pull out nouns and verbs from each sentence and create a new column that marks
# each word that is a type of noun or verb as simply NOUN or VERB
broad_pos <- corpwords %>%
  mutate(noun_verb = case_when(
    startsWith(pos.en, "v") ~ "VERB",
    grepl(pos.en, pattern = "^n$|^pronoun$|^nprop$") ~ "NOUN"
  ))

# add the full sentence to the datatable for easier manual analysis
broad_pos <- merge(broad_pos, corpus$sentences[, c("gls.en", "sentence_id")],
                  by = "sentence_id", all = TRUE,
                  suffixes = c(".word", ".sentence"))

# for each sentence, show the complete sequence of parts of speech
word_order <- corpwords %>%
  reframe(.by = sentence_id, pos.en = toString(na.omit(pos.en)))

# take the word dataframe, drop any empty words (these are mostly punctuation,
# which get their own entries in FLEx's word dataframe), and simplify the noun
# and verb designations to NOUN and VERB

```

```

nv_word_order <- corpwords %>% drop_na(txt.tun) %>% mutate(across(pos.en,
  str_replace, "^n$|^pronoun$|^nprop$", "NOUN")) %>%
  mutate(across(pos.en, str_replace, "^v.*$", "VERB"))

# add the text of the entire sentence to be able to put each word in context
nv_narrow <- merge(nv_word_order, corpus$sentences[, c("gls.en",
  "sentence_id")], by = "sentence_id", suffixes = c(".word",
  ".sentence")) %>% drop_na(gls.en.word)

# add information about text title for even more contextual information
nv_narrow <- merge(nv_narrow, corpus$texts[, c("text_id", "title.en")],
  by = "text_id")

# create new columns to categorize texts into subcorpora and genre
nv_narrow <- nv_narrow %>%
  mutate(era = case_when(
    text_id %in% gatschet_texts ~ "Gatschet",
    text_id %in% swanton_texts ~ "Swanton",
    text_id %in% haas_texts ~ "Haas",
    text_id %in% new_tunica_texts ~ "New Tunica"
  )) %>%
  mutate(era_num = case_when(
    text_id %in% gatschet_texts ~ 1L,
    text_id %in% swanton_texts ~ 2L,
    text_id %in% haas_texts ~ 3L,
    text_id %in% new_tunica_texts ~ 4L
  )) %>%
  mutate(genre = case_when(
    text_id %in% myths ~ "Myths",
    text_id %in% tales ~ "Tales",
    text_id %in% animal ~ "Animal Tales",
    text_id %in% histories ~ "Historical",
    text_id %in% personal ~ "Personal Narratives",
    text_id %in% ethnological ~ "Ethnological"
  ))

# for each sentence, collapse into a single row with the parts of speech of all
# ↪ words in a list
nv_narrow <- nv_narrow %>% reframe(.by = c(sentence_id, gls.en.sentence,
  text_id, era, genre), w_order = toString(na.omit(pos.en)))

# create a new column, "verbfinal". if the part of speech list for the sentence
# ↪ has "VERB" as its last constituent, then fill verbfinal with T[true],
# ↪ otherwise, fill verbfinal with F[false]
nv_narrow$verbfinal <- with(nv_narrow, ifelse(grepl("VERB$", nv_narrow$w_order),
  "T", "F"))

# create another column that will divide sentences into three categories:
# verb-final (VF), verb non-final (VNF), and no verb (NV)
nv_narrow$three_way_distinction <- with(nv_narrow,
  ifelse(grepl("VERB$", nv_narrow$w_order), "VF",
  ifelse(grepl("VERB", nv_narrow$w_order), "VNF", "NV")))

# show counts for each of the three verb placement categories by subcorpus

```

```

nv_three_way_distinction_count <- nv_narrow %>%
  count(three_way_distinction, era, name = "occurrences")

# counts by subcorpus of the number of verb-final sentences and non-verb-final
↪ sentences
nv_narrow_verbfinal_count <- nv_narrow %>%
  count(verbfinal, era, name = "occurrences")

# column separating sentences with a verb from sentences without a verb
nv_narrow$anyverb <- with(nv_narrow,
  ifelse(grepl("VERB", nv_narrow$w_order), "T", "F"))

# counts by subcorpus of sentences with verbs and sentences without verbs
nv_narrow_anyverb_count <- nv_narrow %>%
  count(anyverb, era, name = "occurrences")

# new column where sentences are split into finer categories
nv_narrow <- nv_narrow %>% mutate(noun_final = case_when(
  grepl("VERB", w_order) & grepl("NOUN$", w_order) ~ "noun-final",
  grepl("VERB", w_order) & grepl("post$", w_order) ~ "postposition-final",
  grepl("VERB", w_order) & grepl("adv$", w_order) ~ "adverb-final",
  !grepl("VERB", w_order) ~ "no verb",
  grepl("VERB$", w_order) ~ "verb-final",
  .default = "other"
))

# number of noun-final sentences
nv_noun_after_verb_count <- nv_narrow %>%
  count(noun_final, era, name = "occurrences")

# for each sentence, show just the sequence of nouns and verbs
word_order_broad <- broad_pos %>%
  reframe(.by = c(sentence_id, gls.en.sentence),
  noun_verb = toString(na.omit(noun_verb)))

# get counts of sentences with each sequence of nouns and verbs
word_order_broad_count <- word_order_broad %>%
  count(noun_verb, name = "occurrences")

# get counts of sentences with every sequence of parts of speech
word_order_count <- word_order %>% count(pos.en, name = "occurrences")

# Get all habitual endings (marked by .habitual).
habitualwords <- morpheme_df[ grep(".habitual", morpheme_df$gls.en.morph), ]

# get the verb roots and stems for all words with habitual endings
# class I habituals

```

```

classIhabitualroots <- subset(morpheme_df,
  (morpheme_df$word_id %in% habitualwords$word_id) &
  (morpheme_df$type == "stem" | morpheme_df$type == "root" &
    morpheme_df$type == "bound stem") & (startsWith(morpheme_df$msa.en, "v.I.") |
    is.na(morpheme_df$msa.en)))

# class II habituals
classIIhabitualroots <- subset(morpheme_df,
  (morpheme_df$word_id %in% habitualwords$word_id) &
  (morpheme_df$type == "stem" | morpheme_df$type == "root" &
    morpheme_df$type == "bound stem") & (startsWith(morpheme_df$msa.en, "v.II") |
    is.na(morpheme_df$msa.en)))

# Make a table with the stem, the verb type (to distinguish transitive from
# intransitive; i may also at some point look at semantic differences as well),
# the count of each verb stem over the entire corpus, and in which texts the
# stem appears in in habitual form.
classIhabdplyr <- classIhabitualroots %>% group_by(cf.tun.morph, msa.en) %>%
  summarise(occurrences=n(), number_of_stories=n_distinct(text_id),
    words = toString(unique(txt.tun.full_word)),
    morphemes= paste(unique(cf.tun.word), collapse = "; "),
    stories=toString(unique(text_id))) %>%
  arrange(desc(occurrences))

# serial verbs only
habserialverbs <- classIhabitualroots %>%
  group_by(word_id) %>%
  dplyr::filter(n()>1) %>%
  arrange(morphem_id)

# replacement pairs for digraphs to IPA equivalentents for proper phoneme analysis
rep_str = c("sh"="ʃ", "ch"="tʃ", "\\*"="", "-"="")

# create a copy of the dictionary dataframe to which to apply changes
ipa_dict <- dict_entries

# replace digraphs with IPA equivalentents
ipa_dict$`lexical-unit.tun` <- str_replace_all(ipa_dict$`lexical-unit.tun`,
  rep_str)

# create simplified entries that put all types of nouns, verbs, and adjectives
# into groups
ipa_dict <- ipa_dict %>% mutate(noun_verb = case_when(
  grepl(tolower(`grammatical-info.value`), pattern = "\\sverb") ~ "VERB",
  grepl(tolower(`grammatical-info.value`), pattern = "adjective") ~ "ADJECTIVE",
  grepl(tolower(`grammatical-info.value`), pattern = "noun") ~ "NOUN"
))

# take all nouns, verbs, and adjectives between two and five phonemes long and
# see how many of each type there are across the corpus

```

```

# NOTE: canonical root lengths are from 2 to 4 phonemes, but catching
# five-phoneme roots allows the capture of words with preaspirated consonants
# with preaspirated consonants to be captured
noun_verb_counts <- ipa_dict[between(str_length(ipa_dict$`lexical-unit.tun`),
  2, 5), ] %>% summarise(.by = noun_verb, "occurrences" = n())

# a list of all native Tunica consonants
consonant_inventory = c("h", "k", "l", "m", "n", "p", "r", "s", "ʃ", "tʃ",
  "t", "w")

# create a new copy of the IPA dictionary to modify further, and add an
# "ultimate consonant" column to it.
ipa_dict2 <- ipa_dict %>% mutate(ultimate_consonant=NA)

# get the column names of the IPA dictionary and initialize a final
# consonant dataframe with those column names
final_consonant_df <- ipa_dict2[0, ]

# loop through each consonant in the consonant inventory
for(consonant in consonant_inventory){
  # add rows to final_consonant_df from ipa_dict2 where the root is from 2 to 5
  # phonemes long and contains a consonant, then one vowel (or more, though two
  # vowels should never occur) at the end of the root
  final_consonant_df <- rbind(final_consonant_df,
    ipa_dict2[grep(paste(consonant, "[aeiouAEIOU]*?$", sep = ""),
      ipa_dict2$`lexical-unit.tun`) &
      between(str_length(ipa_dict$`lexical-unit.tun`), 2, 5), ] %>%
    # filter out anything that isn't a noun, verb, or adjective
    dplyr::filter(!is.na(noun_verb)) %>%
    # set the ultimate_consonant column to equal the root-final consonant
    mutate(ultimate_consonant = consonant))
}

# modify the dataframe
final_consonant_df <- final_consonant_df %>%
# get rid of any duplicates (the same dictionary entry but multiple related
# meanings)
reframe(.by = c(id, `lexical-unit.tun`, noun_verb, ultimate_consonant)) %>%
# count the number of roots that have the same ultimate consonant and separate
# by part of speech
reframe("occurrences" = n(), .by = c(noun_verb, ultimate_consonant))

# For the first telling of the flood myth in the haas subcorpus,
# get all the nouns that have a noun prefix but no gender-number suffix
flood_haas_prefix_only <- subset(prefix_only, text_id == 18)
# get all the prefixed nouns, with or without gender-number marking
flood_haas_prefixednouns <- subset(nounprefix_dataframe, text_id == 18)

# do the same for the second telling of the flood myth in the haas subcorpus
flood2_haas_prefix_only <- subset(prefix_only, text_id == 19)

```

```

flood2_haas_prefixednouns <- subset(nounprefix_dataframe, text_id == 19)

# do the same for the flood story in the swanton subcorpus
flood_swanton_prefix_only <- subset(prefix_only, text_id == 88)
flood_swanton_prefixednouns <- subset(nounprefix_dataframe, text_id == 88)

# create a function that takes a dataframe, arbitrary name of a text, and number
# of the text
marking_count_per_text <- function(df, name, text_number)
{
  # add a row to the specified dataframe (which must have the columns "Text",
  # "unmarked", and "total"). Set the Text column to the name of the text, set
  # "unmarked" to the number of nouns in that text that are prefixed but have no
  # gender-number suffix, and set the total to the total number of prefixed nouns
  # in that text
  df %>% add_row(Text = name, unmarked = nrow(subset(prefix_only,
    text_id == text_number)),
    total = nrow(subset(nounprefix_dataframe,
    text_id == text_number)))
}

# create a flood_gender dataframe to compare frequency of gender-marking between
# the three instances of flood myths in the corpus using the
# marking_count_per_text function above
flood_gender <- tribble(~"Text", ~"unmarked", ~"total")

# add rows to the dataframe for each text under analysis
flood_gender <- marking_count_per_text(flood_gender, "Swanton", 88)
flood_gender <- marking_count_per_text(flood_gender, "Haas 1", 18)
flood_gender <- marking_count_per_text(flood_gender, "Haas 2", 19)

# create an orphan_gender dataframe to compare gender-marking frequencies using
# marking_count_per_text function
orphan_gender <- tribble(~"Text", ~"unmarked", ~"total")

# add the rows for the relevant texts
orphan_gender <- marking_count_per_text(orphan_gender, "Gatschet", 8)
orphan_gender <- marking_count_per_text(orphan_gender, "Haas 1", 16)
orphan_gender <- marking_count_per_text(orphan_gender, "Haas 2", 17)

```

Abbreviations

ANA Administration for Native Americans 39

APS American Philosophical Society 50

BAE Bureau of American Ethnology 15–17

CBI communication-based instruction 43

CES community-engaged scholarship v, 40

ESG Eastern Sutherland Gaelic 32–34, 36, 123

FLEx Fieldworks Language Explorer vii, 57–62, 64, 65, 67

GNA gender-number affix vi, viii, 68, 69, 99, 100, 105, 108, 110

KYLY Kuhpani Yoyani Luhchi Yoroni (Tunica Language Working Group) 3, 5, 6, 38, 40–43, 45–50, 56–58, 60, 61, 66, 78, 80, 101, 112, 124

LCRP Tunica-Biloxi Language & Culture Revitalization Program 4, 38–40, 46, 50

LIFT Lexicon Interchange FormaT 65

SIL Summer Institute in Linguistics 57, 58, 60

XML eXtensible Markup Language 65

Glossing abbreviations

1 first person	HAB habitual
2 second person	IMP imperative
3 third person	INAL inalienable
AGNZ agentive nominalizer	INCHO inchoative
AL alienable	m masculine
CAUS causative	NEG negative
CLASSII class II verb inflection	p plural
COMP complementizer	POSS possessive
COMPL completive	QUOT quotative
CONJ conjunction	RECP reciprocal
d dual	REDUP reduplication
EXIST existential	REL relative
f feminine	s singular
FNL phrase-final suffix	SBJV subjunctive
FUT future	SUBR subordinator

Works Cited

- Aikhenvald, Alexandra Y. 2020. "Language Change in Language Obsolescence." Chap. 21 in *The Handbook of Historical Linguistics*, 447–467. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118732168.ch21>.
- "Albert Samuel Gatschet—1832–1907." 1907. *American Anthropologist* 9 (3): 561–570. <https://doi.org/10.1525/aa.1907.9.3.02a00120>. JSTOR: 659541.
- Andersen, Roger W. 1982. "Determining the Linguistic Attributes of Language Attrition." In *The Loss of Language Skills*, 10:83–118.
- Anderson, Patricia M. 2017. "Yanatame Nisa Luhchi Yoroni: Lexicography, Language Revitalization, and the New Tunica Dictionary." PhD diss., Tulane University.
- Anderson, Patricia M. 2020. *Revitalization Lexicography: The Making of the New Tunica Dictionary*. University of Arizona Press.
- Baldwin, Daryl, G. Susan Mosley-Howard, George Ironstack, and Haley Shea. 2022. "Community-Engaged Scholarship as a Restorative Action." In *Replanting Cultures: Community-Engaged Scholarship in Indian Country*, edited by Chief Benjamin J. Barnes and Stephen Warren. Tribal Worlds : Critical Studies in American Indian Nation Building. Albany, New York: State University of New York Press.
- Bloomfield, Leonard. 1927. "Literate and Illiterate Speech." *American Speech* 2 (10): 432–439. <https://doi.org/10.2307/451863>. JSTOR: 451863.
- Boas, Franz. 1942. "Language and Culture." In *Studies in the History of Culture: The Disciplines of the Humanities*, 178–184. Banta.
- Brain, Jeffrey P. 1988. *Tunica Archaeology*. Papers of the Peabody Museum of Archaeology and Ethnology 78. Cambridge, MA: Harvard University Press.
- Brain, Jeffrey P. 2017. "Archaeology of the Tunicas." In *The Tunica-Biloxi Tribe: Its Culture and People*, 1–7. Tunica-Biloxi Tribe of Louisiana.
- Campbell, Lyle, and Martha C. Muntzel. 1989. "The Structural Consequences of Language Death." In *Investigating Obsolescence*, edited by Nancy C. Dorian, 181–196. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620997.016>.

- Dauenhauer, Nora Marks, and Richard Dauenhauer. 1998. "Technical, Emotional, and Ideological Issues in Reversing Language Shift: Examples from Southeast Alaska." In *Endangered Languages: Language Loss and Community Response*, edited by Lenore A. Grenoble and Lindsay J. Whaley, 57–98. Cambridge University Press.
- Dorian, Nancy C. 1973. "Grammatical Change in a Dying Dialect." *Language* 49 (2): 413–438. <https://doi.org/10.2307/412461>. JSTOR: 412461.
- Dorian, Nancy C. 1977. "A Hierarchy of Morphophonemic Decay in Scottish Gaelic Language Death: The Differential Failure of Lenition." *Word* 28 (1–2): 96–109. <https://doi.org/10.1080/00437956.1977.11435851>.
- Dorian, Nancy C. 1978. "The Fate of Morphological Complexity in Language Death: Evidence from East Sutherland Gaelic." *Language* 54 (3): 590–609. <https://doi.org/10.2307/412788>. JSTOR: 412788.
- Dorian, Nancy C. 1981. *Language Death: The Life Cycle of a Scottish Gaelic Dialect*. Philadelphia: University of Pennsylvania Press.
- Dorian, Nancy C. 1994. "Purism vs. Compromise in Language Revitalization and Language Revival." *Language in Society* 23 (4): 479–494. <https://doi.org/10.1017/S0047404500018169>. JSTOR: 4168553.
- Dorman, Caroline. 1931. *Tunica Stories in English*. MS 4940. American Philosophical Society. Philadelphia.
- Drechsel, Emanuel J. 1996. "An Integrated Vocabulary of Mobilian Jargon, a Native American Pidgin of the Mississippi Valley." *Anthropological Linguistics* 38 (2): 248–354. JSTOR: 30028932.
- French, B. F. 1846. *Historical Collections of Louisiana, Embracing Translations of Many Rare and Valuable Documents Relating to the Natural, Civil and Political History of That State*. Vol. 3. New York: Wiley and Putnam.
- Gatschet, Albert S. 1884. *A Migration Legend of the Creek Indians: With a Linguistic, Historic, and Ethnographic Introduction*. Kraus Reprint Company.
- Gatschet, Albert S. 1886a. Letter to the Director of the Bureau of Ethnology. Lecompte, Louisiana. Oct. 24, 1886. Manuscript 1347, National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Gatschet, Albert S. 1886b. *Tunica Vocabulary, Phrases and Sentences, and Texts*. October–November 1886. Manuscript 1301, National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Gatschet, Albert S. 1889. "Sex-Denoting Nouns in American Languages." *Transactions of the American Philological Association (1869-1896)* 20:159–171. <https://doi.org/10.2307/2935676>. JSTOR: 2935676.

- Golla, Victor, James A. Matisoff, and Pamela Munro. 1997. "Mary R. Haas." *Language* 73 (4): 826–837. <https://doi.org/10.1353/lan.1997.0056>.
- Haas, Mary R. 1933. Letter to Edward Sapir, August 29, 1933. Mss.Ms.Coll.94, Mary Haas Papers. American Philosophical Society. Philadelphia.
- Haas, Mary R. 1940. *Tunica*. Vol. IV. Handbook of American Indian Languages. J.J. Augustin.
- Haas, Mary R. 1946. "A Grammatical Sketch of Tunica." In *Linguistic Structures of Native America*, edited by Cornelius Osgood, 337–366. Viking Fund Publications in Anthropology 6. Viking Fund.
- Haas, Mary R. 1950. *Tunica Texts*. University of California Press.
- Haas, Mary R. 1951. "The Proto-Gulf Word for 'Water' (With Notes on Siouan-Yuchi)." *International Journal of American Linguistics* 17 (2): 71–79. <https://doi.org/10.1086/464109>. JSTOR: 1263263.
- Haas, Mary R. 1953. *Tunica Dictionary*. University of California Press.
- Haas, Mary R. n.d. "Gatschet-Swanton Tunica Texts." Mss.Ms.Coll.94. American Philosophical Society. Philadelphia.
- Harvey, Meg. 2022. "Centering Revitalization in Remote Documentation." *Coyote Papers* 24.
- Harvey, Meg. 2023. "Documenting Reawakening Languages: A Case Study of Tunica." PhD diss., University of Arizona.
- Heaton, Raina. 2013. "The Tunica Language Revitalization Project: Methods, Challenges, and Data Conflict in Language Recreation."
- Heaton, Raina, and Patricia Anderson. 2017. "When Animals Become Humans: Grammatical Gender in Tunica." *International Journal of American Linguistics* 83 (2): 341–363. <https://doi.org/10.1086/689832>.
- Hernández de Biedma, Luis. 1850. "A Translation of a Recently Discovered Manuscript Journal of the Expedition of Hernando de Soto into Florida." In *Historical Collections of Louisiana*, edited by Benjamin F. French, vol. 2. New York.
- Jespersen, Otto. 1922. *Language: Its Nature, Development and Origin*. New York: Henry Holt & Company.
- Karlgren, Jussi, and Douglass Cutting. 1994. "Recognizing Text Genres With Simple Metrics Using Discriminant Analysis." In *COLING 1994 Volume 2: The 15th International Conference on Computational Linguistics*. <https://doi.org/10.3115/991250.991324>.

- Kessler, Brett, Geoffrey Nunberg, and Hinrich Schutze. 1997. "Automatic Detection of Text Genre." In *35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*, 32–38. Madrid, Spain: Association for Computational Linguistics. <https://doi.org/10.3115/976909.979622>.
- Klopotek, Brian. 2011. *Recognition Odysseys: Indigeneity, Race, and Federal Tribal Recognition Policy in Three Louisiana Indian Communities*. Duke University Press. <https://doi.org/10.1515/9780822394082>.
- Kroeber, A. L. 1940. "The Work of John R. Swanton." In *Essays in Historical Anthropology of North America*, 1–9. Smithsonian Miscellaneous Collections 100. Washington: Smithsonian Institution.
- Kroskrity, Paul V. 2004. "Language Ideologies." Chap. 22 in *A Companion to Linguistic Anthropology*, edited by Alessandro Duranti, 496–517. Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470996522.ch22>.
- Kroskrity, Paul V. 2009. "Language Renewal as Sites of Language Ideological Struggle: The Need for "Ideological Clarification"." In *Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned*, edited by Jon Allan Reyhner and Louise Lockard, 71–83. Northern Arizona University Press.
- Kuhpani Yoyani Luhchi Yoroni. 2023. *Rowinataworu Luhchi Yoroni / Tunica Language Textbook*. Bloomington, Indiana: Indiana University Press.
- Le Page du Pratz, Antoine-Simon. 1758. *Histoire de La Louisiane*. Vol. 2. Paris: De Bure, l'aîné.
- Loiseau, Sylvain. 2018. *interlineaR: Importing Interlinearized Corpora and Dictionaries as Produced by Descriptive Linguistics Software*. R package version 1.0.
- Malrieu, Denise, and François Rastier. 2000. "Genres et variations morphosyntaxiques." *Revue française de linguistique appliquée* V (2): 101. <https://doi.org/10.3917/rfla.052.0101>.
- Martin, Jack B. 2004. "Languages." In *Handbook of North American Indians, Vol. 14: Southeast*, edited by William C. Sturtevant and Raymond D. Fogelson. Washington: Smithsonian Institution.
- McLendon, Sally. 1997. "Mary R. Haas: A Life in Linguistics." *Anthropological Linguistics* 39 (4): 522–543. JSTOR: 30028484.
- Mikesell, Lisa, Elizabeth Bromley, and Dmitry Khodyakov. 2013. "Ethical Community-Engaged Research: A Literature Review." *American journal of public health (1971)* (United States) 103 (12): e7–e14. <https://doi.org/10.2105/AJPH.2013.301605>.

- Mougeon, Raymond, and Edouard Beniak. 1989. "Language Contraction and Linguistic Change: The Case of Welland French." In *Investigating Obsolescence: Studies in Language Contraction and Death*, edited by Nancy C. Dorian, 287–312. Studies in the Social and Cultural Foundations of Language. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620997.023>.
- Nettle, Daniel. 1999. "Is the Rate of Linguistic Change Constant?" *Lingua* 108 (2): 119–136. [https://doi.org/10.1016/S0024-3841\(98\)00047-3](https://doi.org/10.1016/S0024-3841(98)00047-3).
- Nettle, Daniel. 2012. "Social Scale and Structural Complexity in Human Languages." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1597): 1829–1836. <https://doi.org/10.1098/rstb.2011.0216>.
- Pierite, Chief Joe. 1964. *Letter from Chief Joe Pierite to Mary Haas*.
- Pride, Kalinda, Nicholas Tomlin, and Scott AnderBois. 2020. "LingView: A Web Interface for Viewing FLEx and ELAN Files." *Language Documentation & Conservation* 14:87–107.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rankin, Robert L. 1990. "De Soto's Capaha: Siouan or Tunican?" In *Papers of the Mid-America Linguistics Conference*, 564–572. University of Kansas.
- Sapir, Edward. 1925. "Sound Patterns in Language." *Language* 1 (2): 37–51. <https://doi.org/10.2307/409004>. JSTOR: 409004.
- Schmidt, Annette. 1985a. "The Fate of Ergativity in Dying Dyirbal." *Language* 61 (2): 378–396. <https://doi.org/10.2307/414150>. JSTOR: 414150.
- Schmidt, Annette. 1985b. *Young People's Dyirbal: An Example of Language Death from Australia*. Cambridge Studies in Linguistics. Cambridge, UK: Cambridge University Press.
- Shea, John Gilmary. 1861. *Early Voyages up and down the Mississippi, by Cavalier, St. Cosme, Le Sueur, Gravier, and Guignas*. Albany: J. Munsell.
- SIL. 2020. *Fieldworks Language Explorer*. <https://software.sil.org/fieldworks>.
- SIL. 2023. *Webonary*. <https://www.webonary.org/>.
- Stanton, Timothy K. 2008. "New Times Demand New Scholarship: Opportunities and Challenges for Civic Engagement at Research Universities." *Education, Citizenship and Social Justice* 3 (1): 19–42. <https://doi.org/10.1177/1746197907086716>.
- Supahan, Terry, and Sarah E. Supahan. 2001. "Teaching Well, Learning Quickly: Communication-Based Language Instruction." In *The Green Book of Language Revitalization in Practice*, edited by Leanne Hinton and Kenneth Hale, 195–197. Academic Press.

- Swadesh, Morris. 1948. "Sociologic Notes on Obsolescent Languages." *International Journal of American Linguistics* 14 (4): 226–235. <https://doi.org/10.1086/464009>. JSTOR: 1262876.
- Swanton, John R. 1900. "Morphology of the Chinook Verb." *American Anthropologist* 2 (2): 199–237. <https://doi.org/10.1525/aa.1900.2.2.02a00010>.
- Swanton, John R. 1908. Letter to Frederick Webb Hodge, November 29, 1908. Mss.Ms.Coll.94, American Philosophical Society. Philadelphia.
- Swanton, John R. 1910. "Manuscript 4941." National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Swanton, John R. 1911. *Indian Tribes of the Lower Mississippi Valley and Adjacent Coast of the Gulf of Mexico*. Bulletin / Bureau of American Ethnography 43. Washington, DC: Government Printing Office.
- Swanton, John R. 1919. *A Structural and Lexical Comparison of the Tunica, Chitimacha, and Atakapa Languages*. Bulletin / Bureau of American Ethnography 68. Washington, D.C: Government Printing Office.
- Swanton, John R. 1921. "The Tunica Language." *International Journal of American Linguistics* 2 (1/2): 1–39. <https://doi.org/10.1086/463732>. JSTOR: 1263179.
- Swanton, John R. 1930a. "Copy of A. S. Gatschet's Tunica Texts and Swanton's Tunica Texts." MS 4123. National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Swanton, John R. 1930b. John R. Swanton to Franz Boas, October 17, 1930. Letter. Mss.B.B61. American Philosophical Society. Philadelphia.
- Swanton, John R. 1930c. "Tunica Texts of 'Thunder' and 'Orphan'." MS 4939. National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Swanton, John R. 1931. "Tunica Texts with Interlinear English Translation and Vocabularies." MS 4937. National Anthropological Archives, Smithsonian Institution, Suitland, MD.
- Swanton, John R. 1946. *Indians of the Southeastern United States*. Washington: U.S. Government Printing Office.
- Tunica-Biloxi Indian Tribe. 2017. *The Tunica-Biloxi Tribe: Its Culture and People*. Tunica-Biloxi Tribe of Louisiana.
- Tunica-Biloxi Tribe of Louisiana. 2016. *FAQs*. <https://www.tunicabiloxi.org/faq/>.
- Tunica-Biloxi Tribe of Louisiana. 2018. *Tunica-Biloxi Tribe of Louisiana Receives \$748,200 Grant to Assist with Revitalization of Tunica Language*. Press Release. <https://www.tunicabiloxi.org/tunica-biloxi-tribe-of-louisiana-receives-748200-grant-to-assist-with-revitalization-of-tunica-language/>.

- United States De Soto Expedition Commission. 1939. *Final Report of the United States De Soto Expedition Commission. Letter from the Chairman, United States De Soto Expedition Commission, Transmitting the Final Report of the United States De Soto Expedition Commission*. Washington: U. S. Govt. Print. Off.
- United States Department of the Interior. 1980. *Recommendation and Summary of Evidence for Proposed Finding for Federal Acknowledgment of the Tunica-Biloxi Indian Tribe of Louisiana Pursuant to 25 CFR 54*. Report. Washington, D.C.: Office of Federal Acknowledgment.
- Vega, Garcilaso de la. (1605) 1723. *La Florida del Inca*. Madrid: Nicolas Rodriguez Franco.
- Voegelin, C. F., and F. M. Voegelin. 1977. "Is Tübatulabal De-Acquisition Relevant to Theories of Language Acquisition?" *International Journal of American Linguistics* 43 (4): 333–338. <https://doi.org/10.1086/465504>.
- Whitaker, Tyler A. 2017. "‘We Are All Learners’ Discourses of Ownership and Strategies of Reinforcement in the Tunica Language Reawakening." Master's thesis, Southern Illinois University at Carbondale.
- Wikitungues. 2018. *WIKITONGUES: Donna and Elisabeth Speaking Tunica and English*. <https://www.youtube.com/watch?v=DHH91F52IYc>.
- Willett, Thomas. 1988. "A Cross-Linguistic Survey of the Grammaticization of Evidentiality." *Studies in Language* 12 (1): 51–97. <https://doi.org/10.1075/sl.12.1.04wil>.
- Wolfram, Walt. 2008. "Language Death and Dying." In *The Handbook of Language Variation and Change*, edited by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 764–787. Oxford, UK: Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470756591.ch29>.
- Youchigant, Sesostrie, and Mary Haas. 2011. *Hichut'una awachihk'unanahch*. Harahan, LA: Tunica-Biloxi Tribe of Louisiana.

BIOGRAPHY

Andrew Abdalian was raised in New Orleans. Before the siren call of graduate study beckoned him back, he earned his B.A. in linguistics from Swarthmore College and worked in Northern Virginia. His main research interests are language endangerment and revitalization, language documentation, and the creation and use of digital language learning and analysis tools. His scholarly interests also include language and the law, English double modal constructions, and computational linguistics.