Morphosyntactic Echoes: Language Contact Phenomena in Navajo Child English

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
In Linguistics

By

Corinne L. H. Hutchinson, M.S.

Washington, DC
August 16, 2013
MORPHOSYNACTIC ECHOES: LANGUAGE CONTACT PHENOMENA IN NAVAJO CHILD ENGLISH

Corinne L. H. Hutchinson, M.S.

Thesis Advisor: Donna Lardiere, Ph.D.

ABSTRACT

This dissertation explores the morphosyntactic properties of the language produced by a group of Navajo students participating in a Navajo language and culture program. The dissertation finds that the students employ a non-standard English with morphosyntactic parallels to other non-standard ethnically-linked English vernaculars such as African American Vernacular English (AAVE). The dissertation argues that the vernacular speech documented in this project shows influences from Navajo and should be classified as a contact language.

The study participants attend an off-reservation New Mexico public elementary school where they receive 30 minutes daily of instruction in Navajo language and culture. The participants in the project all identified as heritage learners of Navajo, and were all strongly English-dominant. Analyzing two months of first- and second-grade students’ speech recorded during Navajo language and culture class, the dissertation addresses the following three research questions:

• RQ1: What are the morphosyntactic properties of children’s speech in this Navajo immersion classroom?

• RQ2: To what extent do the levels of morphosyntactic regularity support classification of the children’s dialect of English as a contact language?

• RQ3: What are the implications of the observed language patterns for bilingual families and educators?

The data set collected for this dissertation shows occasional use of Navajo, primarily in the form of simple insertional code-mixing, with the great majority of the children’s utterances (over 90%) being entirely in English. The dissertation focused primarily on several areas (do-support, use of the 3rd-person verbal agreement morpheme –s, use of the morpheme –ed on past tense verbs, and nominal number marking) where patterns of non-standard constructions were observed. Arguing that the rates of non-standard usage exceed those that would be expected from age-matched learners of standard English, the dissertation claims that the data supports a contact language analysis for the speech used by the participants. The dissertation concludes with a discussion of the social implications for Navajo community members choosing to use a non-standard contact language in educational and community contexts.
The research and writing of this thesis is dedicated to the many people who have helped and supported this work. As is probably the case with any major creative project, the process of completing this dissertation has been a truly collaborative effort which could not have been accomplished without a strong network of family, friends, mentors, and co-enthusiasts for the research goals.

Special thanks go to:

My adviser,
Donna Lardiere,
whose kindness, patience, careful advice, and persistent encouragement guided me through every stage of the research and writing process. This dissertation owes its very existence to her.

My husband,
Paul Obrecht,
whose support and faith made this project possible.

My committee member,
Paul Platero,
for sharing his wealth of research experience and linguistic knowledge with me.

My committee member,
Ruth Kramer,
for her insightful comments and advice on multiple drafts of the dissertation.

My parents,
Meg Hachmann and Bill Hutchinson,
for their belief and support.

The many school administrators, teachers, families, students, and community members who shared their time and knowledge with me and helped this project take place: Beverly Becenti-Pigman, Keri James, Antonio Martinez, Marjorie Mitchell, Tamarah Pfeiffer, Geraldine Piasso, Ferman Silago, Amanda Mather, and many more people who cannot be listed here.

The National Academy of Education and the American Philosophical Society for providing greatly appreciated funding to support this work.

Many thanks to all,

**CORINNE L. H. HUTCHINSON**
# Table of Contents

Introduction ........................................................................................................................................... 1

Chapter 1: Languages in Contact ......................................................................................................... 6

1.1. From Bilingualism to Blended Languages: Code-mixing and Mixed Codes .............................. 7

1.1.1. Modeling Code-mixing ........................................................................................................... 11

1.1.2. Mixed Languages .................................................................................................................. 29

1.1.3. Bilingualism: Conclusions .................................................................................................... 36

1.2. Restructured Languages ............................................................................................................. 37

1.2.1. Pidgins .................................................................................................................................... 38

1.2.2. Creoles ................................................................................................................................... 39

1.2.3. Semi-creoles .......................................................................................................................... 44

1.3. Languages in Contact: Conclusion ............................................................................................ 50

Chapter 2: Navajo Language & Linguistics ......................................................................................... 52

2.1. Navajo Linguistics: Past and Present ......................................................................................... 52

2.1.1. History of Navajo Linguistics ............................................................................................... 53

2.1.2. Navajo Linguistics Today ..................................................................................................... 55

2.1.3. Navajo Language in Use: Education and Cultural Attitudes ............................................ 57

2.1.4. Efforts towards Navajo Language Maintenance and Revitalization ................................ 61

2.1.5. Concluding Remarks on the Field of Navajo Linguistics ................................................. 62
2.2. Navajo Morphosyntax .......................................................... 62

2.3. Navajo and English: Languages in Contact .............................. 76

   2.3.1. Code-mixing ............................................................... 76

   2.3.2. Non-Standard English in the Navajo Community ............... 83

2.4. Navajo Language & Linguistics: Conclusion ............................. 86

Chapter 3: Research Design & Methods ........................................ 87

   3.1. Research in Indigenous Communities:
       Community-Based Participatory Research ............................. 88

   3.2. Data Collection .................................................................. 94

   3.3. Data Analysis ..................................................................... 96

   3.4. Research Design & Methods: Conclusion .............................. 98

Chapter 4: Language Data ................................................................ 99

   4.1. Describing the Data ........................................................... 99

       4.1.1. The Verbal Complex .................................................. 100

       4.1.2. Nominal Number Marking ......................................... 107

       4.1.3. Negation ................................................................. 109

       4.1.4. Code-mixing ........................................................... 110

       4.1.5. Data: Conclusion ....................................................... 111

   4.2. Origins of Non-standard Features: Developmental Effects? ....... 112

   4.3. Language Data: Conclusion ................................................. 114
Chapter 5: Discussion ................................................................. 116

5.1. Theoretical Implications: Contact Language Phenomena .......... 116

  5.1.1. The Case for Navajo Origins............................................. 117

  5.1.2. The Case for Origins in Contact Language Universals......... 122

  5.1.3. Code-mixing.................................................................. 129

5.2. Practical Implications: Language in the Community................. 131

  5.2.1. Implications for Families and Communities ...................... 131

  5.2.2. Educational Implications............................................... 137

5.3. Conclusion ......................................................................... 139

Bibliography ............................................................................. 143
The goals of this dissertation are relatively straightforward: this project\(^1\) is intended to provide a targeted exploration of the non-standard English used by heritage Navajo speaking children in a Navajo immersion classroom. Even recently, researchers (e.g. Leap, 1982; Platero, 2001; Schaengold, 2004; Haskan, 2007) have reported that high numbers of Navajo children have limited proficiency in both standard English and in Navajo. However, it is not always clear how these findings should be interpreted: are children demonstrating limited English proficiency due to incomplete acquisition of standard English, or because their target language materially differs from standard English? This dissertation uses naturalistic data collected in first- and second-grade Navajo language and culture classrooms to construct a structural profile of the language being used by young Navajo children in one community. The dissertation goes on to argue that the vernacular speech used by the participants in this project shows systematic differences from standard American English, and these differences are characteristic of a category of contact languages sometimes known as “semi-creoles”. The dissertation is structured around the following three research questions:

- **RQ1**: What are the morphosyntactic properties of children’s speech in this Navajo immersion classroom?
- **RQ2**: To what extent do the levels of morphosyntactic regularity support classification of the children’s dialect of English as a contact language?
- **RQ3**: What are the implications of the observed language patterns for bilingual families and educators?

---

\(^1\) This work was supported by the greatly appreciated generosity of the American Philosophical Society—whose Phillips Fund Grant provided the equipment and travel expenses associated with data collection—and the National Academy of Education, whose Spencer Foundation Fellowship supported the writing of this dissertation.
The project aims to benefit both linguistic researchers and community members interested in promoting minority language education, and the research questions reflect this attempt to balance the interests of these two target audiences.

The theoretical goals of this project are simple; in arguing that the target language differs materially from standard American English, the project provides a profile of a closely-related grammatical system and documents a range of available morphosyntactic constructions which differ from their standard English counterparts. The discussion of these changes is relevant to researchers working on contact language phenomena, as well as researchers working on vernacular forms of standard languages such as American English.

This project also aims to benefit Navajo community members, as well as members of other minority language communities grappling with creating a balance between their heritage languages and a majority language such as English or Spanish. There is no single “Navajo speech community” (there are over a quarter of a million registered members of the Navajo tribe (U.S. Census, 2010), and the Navajo reservation—where many, but not all Navajos live—covers an area roughly the size of Austria), and the findings presented in this dissertation are taken from data collected in a single small community; it is important to note that the particular linguistic findings are specific to the single community which served as the data collection site. Rather, these specific linguistic findings are presented as a foundation for discussing the way communities use ethnically-linked non-standard vernacular speech; the dissertation argues that the children who participated in this project—all of whom were effectively monolingual English speakers—may use non-standard “Navajo English” in contexts where the use of standard Navajo is encouraged.
The paper is organized into five chapters. Chapters 1 and 2 provide background information: chapter 1 reviews the literature on contact language phenomena, and chapter 2 provides information about the Navajo language and the history of Navajo linguistics. The remaining chapters are focused on the work conducted for this dissertation: chapter 3 reviews the research design and methods, chapter 4 presents the data, and chapter 5 discusses the theoretical and practical implications of the collected data and reviews both the limitations of this current study and some directions for future research.

The first chapter of the dissertation, chapter 1, provides an overview of the literature on contact language phenomena including the practice of code-mixing and the formation and linguistic profiles of mixed languages, pidgins, creoles, and semi-creoles. Beginning with a review of pertinent terminology related to bilingualism and multilingualism, the chapter reviews theoretical approaches to modeling code-mixing. The next section of chapter 1 discusses mixed languages, a class of contact languages characterized by their apparently split genetic parentage, and argued by some researchers (e.g. Auer, 1999; Myers-Scotton, 2000; O’Shannessy, 2005) to arise out of wide-spread code-mixing. Languages falling into this category are argued to show the systemic inheritance of lexical and grammatical material from two different languages, such that the resulting language—which shows little or no structural simplification, such as that associated with pidgins and creoles—cannot be clearly assigned to any single genetic tree. The final portion of chapter 1 reviews the relatively well-known set of contact language phenomena known as pidgins, creoles, and semi-creoles. Unlike mixed languages, all three of these phenomena are characterized by systemic structural simplification, including simplified systems of verbal and nominal inflection. A number of researchers (e.g. Penfield, 1977; Bartelt, 1982;
Schaengold, 2004) suggest that the English spoken in Navajo communities shows signs of being a creole or semi-creole language; the final portion of chapter 2 reviews these arguments and discusses their relevance to the current project.

Chapter 2 turns to a discussion of the Navajo language and the history of Navajo linguistics. The first section of this chapter reviews the history of research into the Navajo language, and modern-day work on Navajo linguistics. The chapter then provides some discussion on the modern sociolinguistic profile of the Navajo language, and ends by discussing approaches to modeling Navajo morphosyntax.

Chapter 3 reviews the research design for this project, and discusses some of the considerations which are relevant to research relating to indigenous communities. The first portion of this chapter discusses a proposed research project which was cancelled during the participant recruitment phase; the particular reasons for cancelling this earlier project are pertinent for any researchers hoping to conduct projects in connection with members of indigenous communities. The second portion of the chapter reviews the research plan for the current project and discusses the means of data collection and data analysis.

Chapter 4 presents these data, and discusses the possible role of developmental factors influencing the use of non-standard morphosyntactic constructions. The chapter concludes by arguing that the rates of non-standard constructions exceed those found in the speech of age-matched children who are unambiguously acquiring standard English, suggesting that the target language of the participants in this study is something other than standard English.

The final chapter, chapter 5, is divided into two main parts. The first portion discusses the potential linguistic influences shaping the observed non-standard vernacular, and argues that the
documented speech shows influences from Navajo, but also shows similarities to other proposed semi-creole languages. The second portion of chapter 5 discusses the social and educational implications of the claim that students in Navajo language and culture classes may be using a regular, systematic, and ethnically-linked English vernacular. The chapter (and the dissertation) concludes by identifying the limitations of the current study and outlining several proposed directions for future research.
Chapter 1: Languages in Contact

Language contact—whether through trade, migration, colonialism, tourism, or warfare—has undoubtedly occurred as long as human languages have been spoken, and every human language almost certainly displays the evidence of language contact phenomena knit somewhere into its history. Over the past century and a half, Navajo speaking communities have had substantial contact with the English language, and most Navajo speakers today are also speakers of English (U.S. Census, 2000). The effects of language contact can be short-lived—despite centuries of close geographic proximity and presumably occasional instances of language contact effects, many researchers (e.g. Fennell, 2001) argue that effectively no Celtic influence is seen in modern English—or they can fundamentally change the structure and form of the language spoken in a community, as attested by the host of common Old Norse-derived modern English words such as until, they, give, or leg. Many researchers (e.g. Faarlund, 1990; Thomason, 2000, 2001; Chamoreau & Leglise, 2012) have suggested that the social forces guiding the adoption or rejection of given lexical and structural changes are so varied and complex that models should be expected only to predict what types of contact phenomena can occur, but not what will occur, and Thomason (2000: 173) explicitly claims that “although it is often possible to state necessary conditions for change, it is never possible to state sufficient conditions for change”.

Research into language contact effects is notoriously messy, and there is often a lack of field-wide agreement on the appropriate categorization of given phenomena. The list of language contact phenomena ranges from basic lexical borrowings to endemic code-mixing to the birth of fully-fledged contact languages such as mixed languages, pidgins, creoles, and semi-creoles;
researchers such as Leap (1975, 1982) and Holm (2004) have suggested that some modern-day Native American speakers may speak varieties of non-standard English which could be classified as semi-creoles or partially restructured vernaculars. While the phenomena listed above are certainly not all present in the Navajo-speaking community, it seems helpful to provide a basic overview of language contact phenomena.

This current chapter provides a brief discussion of these main types of language contact phenomena, roughly grouping these phenomena according to the environments in which they develop: the first group being those that arise out of (widespread) multilingualism and the second group being those that arise out of a need for communication in communities (initially) having no single shared common language. This chapter is organized according to this rough sociohistorical framework; section 1.1. discusses the language contact phenomena predominantly associated with widespread individual-level bilingualism, and section 1.2. discusses the language contact phenomena predominantly arising in multilingual speech communities where most individuals have command of only a single functional linguistic system.

1.1. From Bilingualism to Blended Languages: Code-mixing and Mixed Codes

As indicated above, section 1.1. reviews language contact phenomena that arise in communities with high rates of individual-level bilingualism. Often associated with a history of widespread migration or geopolitical instability, predominantly bilingual communities foster a range of systematic language contact phenomena, the most clearly defined categories being code-mixing and the genesis of mixed languages. An in-depth discussion of the social histories
of bilingual communities is not provided in this dissertation\(^2\), but the following few paragraphs do briefly discuss what is usually understood by the term “bilingual”.

Unsurprisingly, the “bilingual” label is often used to describe a wide range of linguistic competencies. Researchers vary in their definitions/descriptions of “types” of bilingualism, and as with many areas of linguistic terminology, a soup of often controversial labels and sub-categories exists. Generally speaking, the so-called *balanced bilingual* is the (rare) speaker who is fully and equally fluent in each of his/her two languages; however, numerous researchers (e.g. Fishman, 1971; Grosjean, 1989; Romaine, 1999) have pointed out that effectively no bilingual speaker uses both of his/her languages in exactly the same circumstances, and that there are almost always subjects and communicative settings in which one language is preferred. As Romaine (1999: 62-63) writes, “Any society which produced functionally balanced bilinguals who used both languages equally well in all contexts would soon cease to be bilingual because no society needs two languages for the same set of functions.” The great majority of bilingual speakers are what is often referred to as *unbalanced bilinguals*, having greater fluency in one language, or greater fluency in a particular language domain. At the opposite end of the spectrum from balanced bilingualism lies so-called *semilingualism*, a controversial condition in which a speaker supposedly fails to gain fluency in any of the languages to which s/he is exposed. This term’s historically frequent application to members of Native American speech communities motivates the following short digression.

Although the earliest explicit formal definition of such a condition is generally attributed to the Swedish researcher Hansegård (1962)—who first employed the term *halvspråchigkeit*—both popular and academic work (e.g. Bloomfield, 1927, Dorothy Salisbury Davis’s 1970 *Where the Dark Streets Go*) illustrates the entrenched perception that multilingual individuals frequently suffer from impaired linguistic abilities. A chorus of researchers (e.g. Cummins, 1979b; Cummins and Swain, 1983; Romaine, 1999) have pointed out that the label *semilingual* is fundamentally pejorative and politically charged, and the very theory of semilingualism has been reprovingly denounced as “a half-baked theory of communicative competence” (Martin-Jones & Romaine, 1986: 36) and as being “essentially indistinguishable from classical prescriptivism” (MacSwan, 2000:3). Under these views, ‘semilingualism’ is a prescriptive failure to acknowledge a speaker’s fluency in one or more non-standard codes.

Unsurprisingly, the epithet of semilingualism has been effectively exclusively levied against bilingual speakers of minority languages. Fishman (1977) drew attention to the dichotomy between the often stigmatized “folk bilinguals” (native speakers of a minority language often with low status outside the minority language community, usually driven to multilingualism by economic demands) and “elite bilinguals” (speakers of the majority language who have chosen to learn and/or use a second language). While the terms themselves are not without controversy, there is general acknowledgement of a difference between societal perceptions of the multilingual status of highly educated speakers—typically biliterate users of standard language varieties—and the multilingual status of speakers from minority and non-standard language backgrounds. This discrepancy between the perceived status of elite bilingualism and folk bilingualism—which closely mirrors that of the socioeconomic status
typically held by members of each of these groups—appears to be rooted in the stigma often associated with minority languages and non-standard varieties of majority languages.

Mainstream linguistic research has shifted towards general promotion of the benefits (or at least the non-detrimental nature) of bilingualism, but even this acceptance encodes a persistent bias towards the bilingualism of elite speakers. Romaine (1999) points out that the vast majority of modern linguistic research on the development of childhood bilingualism has worked with elite bilingual families who are typically biliterate and have primary access to standard spoken and written registers of each of their languages. Folk bilinguals, however, often have primary access to non-standard varieties of one or both of their languages and are often not literate in at least one of their languages. In contrast to the perceived multilingual legitimacy of elite bilinguals—supported by literary and academic norms—the semilingualism label reframes the multilingualism of linguistic minorities as a fundamental linguistic deficit. Under this view, then, it is perhaps not too radical to suggest that the condition of semilingualism might simply be a disparaging term for a multilingual linguistic minority.²

Returning to the discussion of bilingualism, this subsection introduction has provided some preliminary remarks on bilingualism and language practices associated with bilingual

² While the condition of semilingualism is certainly suspicious from a purely linguistic standpoint, it has been suggested that the question itself is perhaps misframed. Cummins (2000:99) argues that “The vast majority of those who have argued that ‘semilingualism does not exist’ have failed to realize that theoretical constructs are not characterized by existence or non-existence but by characteristics such as validity and usefulness, or their opposites.” Studies (e.g. Verhallen & Schoonen, 1993) have found differences (not to be confused with deficits) in the performance of bilingual and monolingual speakers on monolingual tests of formal language ability. Cummins suggests that “many . . . bilingual students tend to gain less access to literate/academic registers in both L1 and L2. The real issue is how do we challenge the coercive social and educational structure that gives rise to this pattern”. (Cummins, 2000:105) The condition of semilingualism may be a misframed fiction, but it is important not to let that fiction obscure the fact that many bilingual individuals—including otherwise marginalized linguistic minorities—do experience limited access to standard language forms.
speech communities. The distinction drawn out between elite and folk bilinguals should not distract from the fact that there are practices—such as code-mixing—which are engaged in at least occasionally by nearly every bilingual speaker. Other phenomena—such as the use of mixed languages—are almost exclusively associated with folk bilinguals, and the specter of semilingualism is particularly present for speakers of mixed languages. Subsection 1.1.1. reviews the history of code-mixing research and morphosyntactic modeling approaches, while subsection 1.1.2 reviews the phenomenon of mixed languages and attempts at modeling this small class of languages.

1.1.1. Modeling Code-mixing

Code-mixing research has taken a wide range of approaches to understanding the phenomenon of code-mixing: work on the social and communicative factors influencing language mixing (e.g. Blom & Gumperz, 1972; Auer, 1984), work on the psycholinguistic properties of language mixing (e.g. Grosjean, 1995), work on the uses of language mixing in bilingual and language learning classrooms (e.g. Mati, 2004; Moodley, 2007; Probyn, 2009; Raschka et al., 2009; Tien, 2009), and finally work on the linguistic properties that constrain language mixing (e.g. Poplack, 1980; MacSwan, 1999; Muysken, 2000; González-Vilbazo & López, 2011).

Research on the grammatical properties of code-mixing traditionally targets cases of intrasentential code-mixing; intersentential code-mixing—where language alterations occur between sentences—has been generally been ignored by morphosyntactic research, perhaps due to the complexity of analyzing intersentential, discourse-dependent morphosyntactic
dependencies. Following the typology outlined in Muysken (2000:3), regardless of the language pairs involved, three basic types of intrasentential code-mixing may be distinguished: **insertion** of words or phrases from one language into a sentence from the other language, **alternation** between languages where the sentence switches from one language to another, and what Muysken terms **congruent lexicalization**, where words and phrases from two languages alternate according to grammatical rules that are the same in both languages. Examples of each of these types are shown in (1a-c).

(1)  

a. **insertion:**  

[TURKISH/NORWEGIAN]  

Biz klatrebutikk-ler-i-ne git-ti-k  

We climbing shop-pl-compound marker-DAT go-PAST-1pl  

‘We went to the climbing shops’

---

4 A troubling issue in code-mixing research is distinguishing instances of code-mixing from the less morphosyntactically interesting cases of lexical borrowing. While defining these two terms is relatively simple—code-mixing involves lexical items and morphosyntactic material from two or more grammars, while borrowing involves only a single grammar and lexicon (into which items may have been adopted)—empirical distinction between these two cases is far trickier. In some cases, borrowed status may be conclusively established based on monolingual speaker use. However, a substantial number of researchers (e.g. Poplack et al., 1988, Sankoff et al., 1990; MacSwan, 2005) have distinguished between *established loanwords*—characterized by their presence in monolingual speakers’ lexicons—and *nonce borrowings*, characterized by their non-established use, by their appearance only in the speech of bilingual speakers, and (often) by their morphological, syntactic, and/or phonological integration.

The distinction between (nonce) borrowing and code-mixing is particularly challenging in the case of single-word insertions, even more so in cases where such words exhibit none of the morphosyntactic and/or phonological features of their native language. In cases where lexical items show phonological assimilation, there is general consensus (e.g. Poplack, 1988, MacSwan, 2002; Jake et al., 2002) that these items at least temporarily form part of the recipient language’s lexicon. Cases of morphological integration (without phonological integration) are, however, more controversial. Some researchers (e.g. Poplack, 1980; MacSwan, 2005, 2010) assert that apparent single-word insertions showing morphological integration are in fact instances of nonce borrowing (and that word-internal code-mixing does not occur); others (e.g. Muysken, 2000; Jake et al., 2002, 2005; González-Vilbazo & López, 2011) suggest that morphological integration is not a valid test of adoption into the lexicon, and argue that code-mixing may in fact occur word-externally. Although it has been argued that “[t]he effort spent on distinguishing [code-mixing] and [nonce borrowing] . . . is not always proportional to the insights generated by such a distinction” (Boumans, 1998: 58), the distinction bears on the validity of popular theoretical code-mixing models and continues to generate substantial debate.

The current project follows Muysken (2000: 78), González-Vilbazo & López (2011: 840), and others in considering single occurrences of morphologically integrated lexical items to be potential instances of code-mixing.
b. **alternation:**

[**JAPANESE/ENGLISH**](Nishimura, 1985: 86)

*Nihon no kaisha ni hataraku to* they have bonuses at the end of the year  
Japan POSS company in work when they have bonuses at the end of the year  
‘When they work for a Japanese company, they have bonuses at the end of the year.’

c. **congruent lexicalization:**

[**SPANISH/ENGLISH**](Pfaff, 1976: 250)

*Bueno, in other words, el flight que sale de Chicago around three o’clock*  
good in other words the flight that leaves from Chicago around three o’clock  
‘Good, in other words, the flight that leaves from Chicago around three o’clock’

While describing these broad types of code-mixes is fairly straightforward, a host of conflicting theoretical models have been proposed to account for details such as exactly what types of words and phrases may be mixed between specific language pairs, or locations where a switch may take place.

Effectively all theoretical approaches to modeling code-mixing share a common goal of describing and predicting the possible types of code-mixing between any sets of languages, but many approaches do differ in their fundamental assumptions about the relationship of cross-linguistic speech to monolingual speech and the theoretical models best suited to describing and predicting naturalistic language. These morphosyntactic modeling approaches may be roughly divided into three broad categories: the early constraint-based systems, often working within generative frameworks (e.g. Poplack, 1980; Sankoff & Poplack, 1981; Woolford, 1983; DiSciullo et al., 1986; Belazi et al., 1994), the Matrix Language Frame (MLF)/4-M models (Myers-Scotton, 1993; Myers-Scotton & Jake, 1995, 2000; Jake et al., 2002, 2005), and the null-hypothesis approach (e.g. Mahootian, 1993; Mahootian & Santorini, 1996; MacSwan, 1999, 2000, 2005, 2009). Many of the proposals falling into the first group are based on largely outdated theoretical frameworks, such as the Aspects model (Chomsky 1965, 1972) or
Government and Binding Theory (Chomsky, 1981), and are not discussed in detail in this current work. The remaining two approaches—the MLF/4-M model and the null hypothesis approach—currently continue to enjoy substantial academic support. These last two approaches differ in their basic assumptions about the nature of cross-linguistic speech: the MLF/4-M model assumes that code-mixing is guided by specific rules pertaining to bilingual speech, while the null hypothesis model works from the basic assumption that cross-linguistic speech is guided by exactly the same principles as monolingual speech, and requires no additional constraints, rules, or stipulations.

**Matrix Language Frame Model**

Broadly speaking, the Matrix Language Frame (MLF) model incorporates intuitions from theoretical and sociolinguistics to create a single model targeting both competence- and production-associated code-mixing phenomena. As its base assumption the MLF model states that for any instance of code-mixing, a Matrix Language and an Embedded Language may be defined, where the matrix language “is the source of the abstract grammatical frame of the constituent” (Myers-Scotton, 2001: 24) and the embedded language “can only contribute limited material (largely only content morphemes and embedded language islands within the larger constituent” (Myers-Scotton, 2001: 24). The matrix/embedded language distinction is variably defined in different versions of the MLF; Myers-Scotton (1993) loosely defines the matrix language as “contributing relatively more morphemes ... at a discourse level” (Myers-Scotton, 1993: 20), while Myers-Scotton & Jake, 2000b suggest that the matrix language may instead be identified by its exclusive suppliance of a narrowly defined set of ‘system’ morphemes, discussed in greater detail below. Although earlier versions of the MLF appear to suggest that the
matrix language is defined for a speaker, Jake et al., 2002 declare that “[t]he ML may change within successive CPs, even within a multi-clausal sentence, but we stress that the ML does not change within a single bilingual CP” (Jake et al., 2002: 73). The current project will follow later versions of the MLF in identifying the ML by CP rather than by speaker.

The MLF approach considers a code-mixing discourse to contain three types of constituents: ML islands (effectively monolingual matrix language constituents), EL islands (effectively monolingual embedded language constituents), and ML+EL islands (constituents where the constituent cannot be assigned to one language or the other, such as in the mixed Spanish/English utterance “Aquí, right?”, where an equally strong case might be made for a primarily Spanish constituent as for a primarily English constituent).

Following a constraint-based approach to modeling code-mixing, the MLF proposes two fundamental principles controlling the morphosyntactic properties of mixed utterances:

(2) **Morpheme Order Principle**: Surface morpheme order will be that of the ML in ML + EL constituents. (Myers-Scotton, 1995: 239)

(3) **System Morpheme Principle**: All syntactically or externally relevant system morphemes come only from the ML in ML + EL constituents. (Myers-Scotton, 1995: 239)

While these ‘system’ morphemes—and their complements, the ‘content’ morphemes—are somewhat loosely defined in earlier versions of the MLF, Myers-Scotton & Jake (2000a, b, 2001) refine to the system/content morpheme distinction with a classification scheme termed the 4-M model. Presumably so called because it identifies four basic types of morphemes, the 4-M model uses data from code-mixing, aphasic speech, and second language learners to support a

---

5 The abbreviation CP refers to the “complementizer phrase”, a unit of analysis in generative syntax which is broadly equivalent to the clause.
four-way production-based division of morphemes into content morphemes, early system morphemes, bridge late system morphemes, and outsider late system morphemes.

(4) 4-M Model Morphemes

a. **Content morphemes** are defined as being involved in thematic role assignment: both theta role assigners and theta role receivers are content morphemes. Verbs, nouns, and adjectives are proposed to generally be content morphemes, although Myers-Scotton & Jake warn that the content morpheme label exhibits cross-linguistic variation.

b. **Early system morphemes** bear semantic and/or pragmatic features, but do not participate in theta role assignment. Examples of suggested early system morphemes are English the and up.

c. **Bridge late system morphemes** “connect content morphemes with each other without reference to the specific semantic/pragmatic properties of a content head” (Myers-Scotton & Jake, 2000b: 1064). English of, possessive ‘s, and French de are all proposed to be instances of bridge late system morphemes.

d. **Outsider late system morphemes** are defined as being dependent on syntactic information outside of their immediate maximal projections. English 3rd person singular –s is proposed to be an outsider late system morpheme.

The 4-M model is used to explain distinctions within the MLF; for example, the Matrix Language is defined as the only language able to provide outsider late system morphemes in a given utterance.

The MLF model has enjoyed a great deal of popularity since its inception, and—supported by researchers including Clyne (2003), Deuchar (2006), etc—is generally considered to be the current mainstream approach to modeling code-mixing. Critics do exist, however; other researchers have argued that the MLF makes inaccurate predictions for given data sets. For example, Campos (2003: 138) argues of the MLF that “the embedded island trigger hypothesis prediction of obligatory islands is too strong, and that the functional head constraint prediction of
no switches between a functional head and its complement is not reflected in performance.” Zabrodskaya (2009) argues that the MLF model’s exclusion of phonological criteria for morpheme classification results in failed matrix language identifications for Russian/Estonian codemixing data, Mwamba (2011) claims that the MLF fails to account for the appearance of Lingala verbal inflection in apparently ML French constructions, MacSwan (2005) argues that the MLF makes erroneously liberal predications concerning negation in Nahuatl/Spanish code-mixing, and González-Vilbazo & López (2011a,b) suggest that the MLF is unable to account for word order in Spanish/German light verb constructions. Finally, the theoretical assumptions of the MLF have been vigorously critiqued in a series of papers by MacSwan (e.g. 2002, 2005, 2009, 2010); these criticisms include the claim that Jake, Myers-Scotton, & Gross (2002)—whose paper argues that the MLF is compatible with the theoretical framework of the Minimalist Program (Chomsky, 1995)—demonstrate a fundamental misunderstanding of the Minimalist process of feature checking (discussed further in section 1.2.1.).

Exploring the Null Hypothesis: Constraint-Free Code-mixing

The models discussed up to now—the constraint-based approaches and the MLF/4-M model—outline tangible and explicit systems for capturing code-mixing data. In contrast, the constraint-free approaches explore the null hypothesis, namely that no special morphosyntactic model is needed to capture code-mixing data, and that such data may simply be treated with the same morphosyntactic modeling tools used for monolingual data.

One of the earliest researchers to articulate this stance was Mahootian, whose 1993 dissertation A Null Theory of Codeswitching argues that “no special constraints or mechanisms are required to account for codeswitching” (Mahootian, 1993: 138). Working with a tree-
adjoining grammar (TAG) system, Mahootian uses data from Farsi/English codemixing to support the argument that any instance of code-mixing can be predicted from exactly those principles required to predict monolingual speech.

Drawing on the same fundamental intuitions, MacSwan (1999) outlines a null-hypothesis approach to modeling code-mixing using the theoretical framework of the Minimalist Program (Chomsky, 1995a). Subsequent work applying the Minimalist Program to code-mixing data includes further developments by MacSwan (e.g. 2002, 2005, 2009, 2010), González-Vilbazo (2005), and González-Vilbazo & López (2011a,b). After a short orientation to the Minimalist Program, the remainder of this section discusses specific applications of the Minimalist framework to code-mixing data.

The Minimalist Program

Briefly, the Minimalist Program is a generative approach to syntax which aims to eliminate elaborate and unnecessary syntactic machinery. Like many earlier generative approaches, the Minimalist Program subscribes to the principle of Universal Grammar (UG)—generally attributed to Chomsky

which asserts that all human languages are assembled from the same set of universal linguistic elements which are biologically ‘hardwired’ into the human brain.

The basic model of the Minimalist Program can be represented with a simple schema as shown in (5).

---

While the modern concept of Universal Grammar is generally attributed to Chomsky (1965, 1966), Pesetsky (2000: 2) points out that both the term and the concept itself are effectively continuations of “an earlier grammatical tradition that explicitly sought universal semantic roots of syntax (for example, the 1660 Port-Royal Grammaire générale et raisonnée).”
Within this framework, the prime mover is the lexical item, and the derivation is fueled by these lexical items’ accompanying features (syntactic, i.e. \( \varphi \)-features (person/gender/number), case, or tense; or semantic, i.e. \([-\text{animate}]\)). Cross-linguistic differences are attributed solely to the different features associated with lexical items rather than to any language-specific principles of morphology or syntax\(^7\). Not only do lexical features vary in type, but they vary in whether they are ‘interpretable’ or ‘uninterpretable’. Interpretable features are effectively inherent semantic properties of the lexical item, so a noun might bear the interpretable features \([+\text{animate}]\) and \([+\text{plural}]\), or a \(T^0\) might bear the interpretable feature \([+\text{PAST}]\). Uninterpretable features, on the other hand, are not inherent semantic properties of the lexical item, and serve only a syntactic purpose, so a pronoun might simply bear an unvalued uninterpretable case feature (since case serves only to establish a syntactic relationship between the verb and a nominal). Following Pesetsky & Torrego (2001), uninterpretable case features are identified by a lowercase \(u\), so for our preceding examples the feature on the pronoun would be represented as\([u\text{CASE:} ]\).

The features of lexical items are extremely important, since they determine the course of the derivation. In the Minimalist framework a syntactic structure is built from the bottom up; to

---

\(^7\) Baker (2008) labels this assumption—that all parametric variation stems from variation in lexical features—the Borer-Chomsky Conjecture, expressing it in the following terms: “All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon.” (Baker, 2008: 253). Baker attributes this hypothesis to the work of first Borer (1984) and later Chomsky (1995a).
begin a derivation an array termed the ‘numeration’ is populated with feature-bearing lexical items. (These ‘lexical’ items may here include functional items such as $T^0$, $D^0$, and so on.) The features of these items trigger the operation of the functions Merge (which merges lexical items or pieces of syntactic structure into binary sets) and Move (which moves lexical items or already formed pieces of syntactic structure).

Syntactic dependencies (such as subject-verb agreement) are established via probe-goal relationships between lexical items with uninterpretable features and the lexical items bearing corresponding (valued) features; for a syntactic derivation to succeed all lexical items must have their features valued and checked, so the operation Agree values and/or checks the features of the probe against those of its goal. Unvalued features may be valued by either interpretable or uninterpretable features, so lexical items bearing either interpretable or uninterpretable features may serve as goals. The checking of the uninterpretable features allows these features to be cancelled, which prevents the derivation from ‘crashing’ (the term applied to the processing of an attempted ungrammatical derivation, such as *We reads it); because only interpretable features can be processed by the conceptual-intentional interface, any unchecked (and therefore uncancelled) uninterpretable features will cause the derivation to crash at the semantic interface.

The end product of a syntactic derivation is Spell Out, where relevant components of the derivation are processed by two systems of interpretation: LF (logical form), which processes the semantic features of the items in the derivation, allowing the speaker to understand the semantic value of the derivation and PF (phonetic form), which processes the phonological features, allowing the speaker to pronounce the derivation. Following work by Chomsky (2001), the derivation is thought to be sent to Spell Out in discrete chunks, or phases. Only certain syntactic
constituents (such as CP and transitive vP) function as phases; the phase is argued to be the unit of analysis for the interfaces LF and PF, such that the phase head complement, or *spell-out domain*, of each phase—once complete—is sent off to Spell Out for interpretation. This process of cyclic interpretation can be thought of as reducing the strain on a language user’s computational system and is often used as a tool for explaining the unavailability of certain morphosyntactic operations which would involve material on both sides of a phase boundary. Evidence for the identification of many putative phases remains a subject of academic discussion.

Although a number of details have admittedly been glossed over in this overview, this brief orientation to the Minimalist Program should provide sufficient background to undertake a review of the Minimalist approaches to modeling code-mixing.

**Minimalist Approaches to Code-mixing**

Working within the Minimalist framework, the basic intuition guiding MacSwan’s approach is the notion that code-mixing is derived from exactly the same principles as monolingual speech: no more, and no less. Specifically, MacSwan (1999:146) asserts that:

\[(6) \quad \text{Nothing constrains code switching apart from the requirements of the mixed grammars.}\]

While earlier generative researchers (e.g. Poplack 1980, 1981; Belazi et al. 1994) had proposed specific constraints and principles to predict and describe licit code-mixing phenomena, MacSwan works from the assumption that morphosyntactic patterns of a code-mixed utterance are determined solely by the properties of the original languages which contribute to that
utterance, or in other words, the grammar of the mixed code is determined solely by the grammars of the monolingual codes.

The proposal builds on the basic Minimalist intuition that apparent cross-linguistic syntactic variation is nothing more than the natural result of different features associated with the items in a language’s lexicon; the computational or syntactic system is taken to be cross-linguistically invariant, and all observed morphosyntactic variation is attributed to the idiosyncratic features of lexical items. Applying this principle to cross-linguistic constructions, MacSwan suggests that the morphosyntactic properties of any such instances of code-mixing are predicted solely by the features of the individual lexical items assembled from any participating languages (requiring no special code-mixing stipulations or accommodations). By way of illustration, MacSwan presents code-mixing data involving an SVO and an SOV language to suggest that the word order in the mixed utterance simply falls out of the featural requirements of the verb.

---

8 MacSwan (1999) works with the Minimalist assumptions outlined in Chomsky (1995a). As in this version of the Minimalist Program all lexical verbs are generally assumed to move either overtly or covertly from V to T, and in some cases to C in the process of derivation, and as code-mixing is argued to be impossible within X₀ compounds or under X₀—MacSwan (1999: 224) argues that “code switches below X₀ are ill-formed”, and that “apparent morphological compounds may not be true X₀’s”—MacSwan suggests that V, T, and for relevant languages C, must be from the same language. These claims have been challenged by work on code-mixing in language pairs including French/Wolof (e.g. Legendre & Schindler, 2010), French/Lingala (Mwamba, 2011), Spanish/German (Cantone, 2005), and a collection of language pairs including German/English, Croatian/English, and Turkish/Norwegian (e.g. Jake et al., 2005). In subsequent work MacSwan (e.g. 2000, 2005, 2010) proposes a refinement to the system which he terms the PF Disjunction Theorem:

(i) The PF component consists of rules/ constraints which must be (partially) ordered/ ranked with respect to each other, and these orders/rankings vary cross-linguistically.

(ii) Codeswitching entails the union of at least two (lexically encoded) grammars.

(iii) Ordering relations are not preserved under union.

(iv) Therefore, codeswitching within a PF component is not possible.

(MacSwan, 2000: 45)
In both (7a) and (7b), the observed word order is argued to be the natural result of features associated with the specific V, T, and C chosen from the two lexicons available to the bilingual speaker. In (7a), the SVO word order is predicted to be the natural result of the speaker having selected a verb from English, a language where V is taken to be a weak case assigner (necessitating post-verbal objects), and T and C are taken to have weak v-features (thus allowing only covert verb movement, visible at LF but not PF). Similarly, in (2b), the SOV word order is predicted to be the natural result of the speaker having selected a verb from Korean, a language where V is instead taken to be a strong case assigner (thus permitting pre-verbal objects), but T and C are still taken to have weak v-features.

Subsequent researchers working with Minimalist approaches to code-mixing have included Chan (2003, 2008), who, working primarily with English/Cantonese code-mixing data builds on the analysis by suggesting that lexical and functional items behave in fundamentally different ways when they are code-mixed; functional items—e.g. C⁰, T⁰, v⁰—maintain the structural head positions required by their native features, while lexical items—e.g. N, V—appear in the positions dictated by the surrounding functional items. Other researchers include

MacSwan takes pains to assert that this theorem should not be thought of as a constraint or other code-mixing specific requirement, but instead as an extension of the principle of Full Interpretation outlined in Chomsky (1995a), which requires that all elements have interpretable phonetic features in order to be processed by PF. Since a mixed phonological system cannot be constructed, code-mixes under X⁰ are argued to produce unpronounceable elements which crash at PF. MacSwan defends the continued constraint-free integrity of his proposal, saying that the PF Disjunction Theorem “is a theory about the relationship between the phonological components of a bilingual’s linguistic system” (2000: 46) and not a syntactic constraint on code-mixing.
González-Vilbazo, whose 2005 dissertation reports that a Minimalist approach provides accurate predictions for German/Spanish code-mixing. More recently, González-Vilbazo & López (2011a) have suggested a refinement to the Minimalist approach to code-mixing, using data from German/Spanish light verb constructions to suggest that code-mixing may draw on grammatical options made available by UG, even when such constructions (such as light verbs) are not available in the monolingual grammars of either of the participating languages. Expanding on this argument, González-Vilbazo & López (2011b) suggest that German/Spanish light verbs provide evidence that \( v^0 \) is in fact the locus of grammatical features for its complement VP, and that word order in the VP is dictated by the \( v^0 \), not the VP itself.

The Minimalist model (and its underlying constraint-free linguistic model) is not without its own critics; for example Jake, Myers-Scotton, & Gross (2002, 2005) point out that there is no obvious reason that a bi- or multi-lingual language system should be subject to the exact same principles as a monolingual language system, and the authors raise concerns about the wisdom of forcing bilingual data into a single-language linguistic model.

**Child Code-mixing: Modifying the Models?**

The two models discussed above—the MLF and the Minimalist model—are the two approaches to modeling adult code-mixing which continue to receive mainstream academic support. Because this current project works with data collected from children, the following paragraphs briefly review the literature on modeling children’s code-mixing.

When code-mixing data is collected from children, there is disagreement about whether developmental factors need to be taken into consideration, and whether the same models may be used for adult and child data. This brief subsection divides work on morphosyntactic modeling of
child code-mixing into three groups: those working with essentially unmodified versions of the MLF/4-M models (e.g. Paradis et al., 2000; Namba, 2008), those working with unmodified versions of the null hypothesis/Minimalist model (e.g. Cantone, 2007; Cantone & Müller, 2008), and those working with the Bilingual Bootstrapping Hypothesis (Gawlitzek-Maiwald & Tracy, 1996; Gawlitzek-Maiwald, 2000, 2003), which argues that child code-mixing is primarily motivated by incomplete acquisition of vocabulary and morphosyntax in one or both of the child’s languages.

**MLF Approaches to Child Code-mixing**

Studies applying the MLF to child data include Paradis et al. (2000), Takagi (2000), Pert & Letts (2006), and Namba (2008). Working with French-English codemixing data collected from bilingual preschool-aged children, Paradis et al. (2000) applies the data to the MLF, concluding that “the children obeyed all the constraints set out in the Matrix Language Frame model the majority of the time” (Paradis et al., 2000: 245). The study concluded that there was no need to modify the MLF to account for the child codemixing data, attributing any differences in accuracy between the child data and adult data to simple incomplete acquisition of the morphosyntactic systems of the participating language. For example, Paradis et al. (2000) found that French-English bilingual children were likely to violate the MLF’s System Morpheme Principle (SMP) outlined in (3) and reprinted here as (8):

(8) System Morpheme Principle (SMP): All syntactically or externally relevant system morphemes come only from the [Matrix Language] ML in ML + [Embedded Language] EL constituents” (Myers-Scotton, 1995: 239)

The SMP predicts that elements such as quantifiers, tense and agreement morphology, auxiliary verbs, copulas, do-support, negative operators, and pronominal clitics should all be provided by
the matrix language. However, Paradis et al. found a number of instances where French-English bilingual children produced data such as that shown in (9).

(9) a. *Il a* put this.  
    he has put this  
    ‘He put this”

b. I like *pas* strawberries  
    ‘I don’t like strawberries’

Example (9a) shows a child producing a French pronominal clitic and auxiliary verb with an English verb and demonstrative pronoun, while (9b) shows a child producing an all-English sentence with French negation (which notably occurs post-verbally, as appropriate to French but not English negation). The SMP predicts that neither utterance should occur, since the auxiliary verb in (9a) and the negation in (9b) should be provided by the ML.\(^9\)

However, Paradis et al. (2000) suggest that the apparent violations can be explained by the uneven rates at which INFL-related items are acquired in French and English. Noting that the great majority of SMP violations involved so-called system morphemes from French inexplicably appearing in ML English sentences, Paradis et al. point out that INFL-related items are acquired sooner in French than in English, and suggest that the bilingual children were “employing a gap-filling strategy . . . to increase the communicative competences of [their] English sentences” (Paradis et al., 2000: 254)\(^{10}\). By drawing on the preceding explanation, the

\(^9\) An important and tricky issue for the MLF is determining the matrix language for a given utterance. Paradis et al. (2000) adopt the ML definition of Myers-Scotton (1993), which argues that “the determination of the ML should be based on both sociolinguistic and psycholinguistic factors for a stretch of discourse (not for an individual utterance)”. A ML was identified for each participant by means of individual assessment. Both utterances in (4) were produced by children whose ML was assessed to be English.

\(^{10}\) This explanation is similar to the approach suggested in the Bilingual Bootstrapping Hypothesis described in section 1.3.4., and raises the question of whether Paradis et al. (2000) implicitly suggest that the MLF applies to child data only insofar as the child speakers are adult-like in their use of code-mixing.
apparent violations of an MLF principle are attributed to independent developmental principles, and Paradis et al. (2000) conclude that no child-specific modifications to the MLF are warranted.

**Unmodified Minimalist Models in Child Code-mixing**

A small number of studies have assessed unmodified versions of the Minimalist approaches with child code-mixing data. Cantone (2005) works with code-mixing data from Italian/German bilingual children and reports that the child data supports the notion that “two separate lexicons in the bilinguals make use of one invariant Computational System” (Cantone, 2005: 493). Cantone (2007) expands on this earlier work, presenting three main claims: (a) that child grammars are organized identically to adult grammars, and that the morphosyntactic properties of child code-mixing are identical in kind (if not in frequency) to those of adult code-mixing, (b) that code-mixing does not involve any linguistic constraints other than those of the monolingual grammars, and (c) that functional categories (e.g. $C^0$, $T^0$) determine the structural properties of their complements.

**Child Code-mixing: The Bilingual Bootstrapping Hypothesis**

Grounded primarily in language acquisition research rather than in morphosyntactic theory, the Bilingual Bootstrapping Hypothesis (Gawlitzek-Maiwald & Tracy, 1996; Gawlitzek-Maiwald, 2000, 2003) suggests that code-mixing in young bilingual children is motivated by incomplete acquisition of syntactic structures in one of their languages, leading the child to press into service comparable constructions from the other language. Examples such as (10) are presented to support the argument that young children’s code-mixing can be attributed to incomplete acquisition of their monolingual grammars.
Example (10), produced by a bilingual German/English speaking child, shows English-like word order despite the use of all German lexical items, and Gawlitzek-Maiwald & Tracy argue that this is an instance of English interference. More canonical examples of code-mixing such as that shown in (11) are argued to be instances of lexical gaps in the target language, which in the case of example (11) is presumably German.

Gawlitzek-Maiwald & Tracy argue that children’s code-mixing is fundamentally a developmental phenomenon, and they report that the rate of code-mixing in their data drops precipitately as the child acquired specific monolingual-like structures in one or the other of the two languages, such as the German CP. Working with the fundamental belief that a bilingual child does have two separate grammatical systems, the authors argue that before the child has become a fluent user of his/her languages, “something that has been acquired in language A fulfills a booster function for language B” (Gawlitzek-Maiwald & Tracy 1996:903).

The Bilingual Bootstrapping Hypothesis is specifically a theory concerning language mixing phenomena in developing speakers, and does not claim to address code-mixing data from fluent bilingual speakers. As noted in a footnote to section 1.3.1., Paradis et al. (2000) appear to...
suggest that while some child data conforms to the predictions of the MLF model, other child data would be better described by a model such as the Bilingual Bootstrapping Hypothesis.

**Code-mixing: Conclusion**

This subsection reviews current approaches to modeling the morphosyntax of code-mixing in both adult and child speech. Both the MLF model and the null hypothesis/Minimalist model have been argued for as appropriate tools for modeling the code-mixing data from adults and fluent children, while the Bilingual Bootstrapping Model has been proposed as a tool for assessing the code-mixing of children with developing language skills. Chapter 5 revisits this conversation, briefly discussing the varying suitability of these code-mixing models in capturing the Navajo-English data produced by the child participants in this study. The next subsection turns to the phenomenon of mixed languages.

**1.1.2. Mixed Languages**

Unlike the spontaneous and often idiosyncratic morphosyntax of code-mixing, the somewhat controversial category of mixed or split languages is characterized by conventionalized morphosyntactic patterns involving lexical and grammatical material from multiple source languages. Mixed languages are generally described as having a so-called “split ancestry” (Matras & Bakker, 2003; Thomason, 1995, 1997; Bakker & Mous, 1994), where no primary genetic classification can be assigned. Mixed languages do not show the structural simplification that is often popularly associated with creoles or other contact languages; data from argued mixed languages show rich systems of inflectional and derivational morphology rooted in the morphosyntactic systems of one or both parent languages. While there is no current indication that any Navajo-speaking community has adopted a mixed language, the practice of
code-mixing—a prevalent practice in many Navajo communities—has been argued by some researchers to be diachronically linked to mixed languages; the end of this section discusses the bearing this connection might have on the Navajo speech community.

Classification systems vary, but even the most conservative researchers describe roughly a half dozen to a dozen mixed languages in use today; the most widely accepted examples of modern mixed languages include Media Lengua, spoken in Ecuador and having Spanish and Quechua roots (e.g. Muysken, 1994, 1997, 2012), Michif, spoken in scattered Canadian communities and having French and Cree roots (e.g. Bakker, 1997; Matras & Bakker, 2003), Ma’a (a dialect of Mbangu), spoken in northern Tanzania and having Bantu and Cushitic roots (e.g. Mous, 1994, 2003; Bakker & Muysken, 1994; Croft, 2003), Mednyj Aleut, spoken on Copper Island and having Russian and Aleut roots (e.g. Golovko, 1994; Matras & Bakker, 2003; Croft, 2003), and a family of underdocumented mixed Romani languages, spoken in Roma communities across Europe and having Romani roots mixed with various European languages (e.g. Bakker & Muysken, 1994). Other languages argued to be mixed languages include Paraguayan Jopará, a largely unanalyzed language employing fundamentally Guarani morphosyntax with a high proportion of Spanish lexical items (e.g. Lustig, 1996), Light Warlpiri, spoken in Australia and having Kriol and Warlpiri roots (e.g. O’Shannessy, 2005, 2008, 2009), Gurundji Kriol, spoken in Australia and having Kriol and Gurundji roots (e.g. McConnell & Meakins, 2005; Meakins & O’Shannessy, 2010), and the now-extinct language Petjo, spoken in Indonesia and having Dutch and Malay roots (e.g. Croft, 2003; van Rheeden, 1994).

Isolated examples of mixed-language data often strongly resemble the code-mixing examples discussed in the previous section; only in the context of a larger corpus of data does the
distinction between the two types of phenomena become clear. Example (12) illustrates this similarity with an utterance in Media Lengua (perhaps the best documented mixed language today), along with the monolingual Spanish and Quechua equivalents.

(12)  

    “I come to ask a favor.”


c. Sp:  Vengo para pedir un favor.  

_comparison of (12a) to the monolingual Quechua and Spanish equivalents shows that the Media Lengua verbal and nominal inflection is effectively identical—with some largely phonological differences—to that of the comparable Quechua utterance, while the lexical items are all of Spanish origin.

Taken in isolation, example (12a)—where Quechua-origin material is indicated in italics—is quite reminiscent of the interwoven style of code-mixing. However, it is important to note that mixed languages are not simply instances of widespread code-mixing; a mixed language is a single language system, and speakers of a mixed language do not choose which phrases to draw from one source language or the other, instead having access to exactly one system of verbal inflection, one system of nominal inflection, one lexicon within that linguistic system.

Although nearly all documented instances of mixed languages involve speech communities where at least one of the parent languages is still widely used—for example most speakers of Ecuadorian Media Lengua are also speakers of both Spanish and Quechua (e.g. Muysken, 1981; Dikker, 2008, Shappeck, 2011)—counterexamples exist. Michif, a mixed
language with French and Plains Cree roots, has been argued by Matras & Bakker (2003) to be the single example of a mixed language whose speakers no longer have meaningful access to either of the parent languages; Bakker (1997:6) points out that “. . . virtually no Michif speakers are able to speak or understand Cree, and only a minority speak and understand French,” while essentially all Michif speakers are fluent in English.

**Modeling Mixed Languages**

Identifying structural universals in mixed languages is challenging, with one of the most obvious limitations being the extremely small sample size. Only a handful of mixed languages have been described, and even fewer have been documented with any sort of rigor. Furthermore, even among the handful of specialists working with mixed languages, there is a lack of field-wide consensus on what constitutes a “true” mixed language. The narrowest definition of a mixed language, supported by some researchers (e.g. Myers-Scotton, 2003), is one in which the lexical material stems entirely (or almost entirely) from one language and the grammatical system from another. However, the structural profiles of many mixed languages challenge this clean-cut distinction.

Mixed languages do appear to inherit their lexical and grammatical material on a largely class-by-class basis. As suggested in example (12) above, Media Lengua draws nearly all of its lexical material from Spanish while still employing predominantly Quechua morphosyntax. In fact, Muysken (1997) writes of Media Lengua “. . . it is essentially Quechua with the vast majority of its stems replaced by Spanish forms.” (Muysken, 1997: 365). The table below shows a rough breakdown of the distribution of Quechua- and Spanish-origin material in Media Lengua.
(13) **Structural Profile of Media Lengua** (reprinted from Matras & Bakker, 2003: 5)

<table>
<thead>
<tr>
<th>Source Language</th>
<th>Quechua</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexicon</td>
<td>85% of lexicon (Muysken, 1997 gives 90+%)</td>
<td></td>
</tr>
<tr>
<td>Inflection</td>
<td>Verb &amp; noun inflection</td>
<td></td>
</tr>
<tr>
<td>Function words</td>
<td>(Plural pronouns)</td>
<td>Pronouns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Demonstratives</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indefinite pronouns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Negators</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Adverbs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conjunctions</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Numerals</td>
</tr>
<tr>
<td>Syntax</td>
<td>Word order</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Subordinations</td>
<td></td>
</tr>
</tbody>
</table>

As example 13 shows, Quechua has provided most of Media Lengua’s functional material and Spanish most of the lexical; although there are categories—such as the conjunctions—with members adopted from both parent languages, the overarching pattern appears to support the generalization that a mixed language shows a functional/lexical split in its inheritance.

However, other mixed languages show different patterns of inheritance which are more challenging to fit into the narrowly defined paradigm; example (14) below shows an utterance from the French/Cree-based language Michif.

(14) ki-nipi-ya-wa son frère aspin kâ-la-petite-fille-iwi-t
     ST-die-OB-3sg his/her brother since COMP-the-little-girl-be-3sg
     “Her brother died when she was a young girl.”

(Bakker & Muysken, 1994: 45)
In the Michif utterance reprinted in example (14), nominal material is derived from French, but all verbal material—lexical verbs and verbal inflection—is derived from Cree. The table below provides an outline of the etymological roots of the Michif language.

(15) **Structural Profile of Michif** (reprinted from Matras & Bakker, 2003: 2)

<table>
<thead>
<tr>
<th>Source Language</th>
<th>Cree</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexicon</td>
<td>Verbs, Few nouns</td>
<td>Nouns, Adjectives</td>
</tr>
<tr>
<td>Inflection</td>
<td>Verb inflection</td>
<td>Noun inflection</td>
</tr>
<tr>
<td>Function words</td>
<td>Personal pronouns, Possessives with Cree nouns, Demonstratives</td>
<td>Possessives, Definite articles, Indefinite articles</td>
</tr>
<tr>
<td></td>
<td>Indefinite pronouns, Some negators, Most adverbs</td>
<td>Most negators, Few adverbs, Numerals</td>
</tr>
<tr>
<td>Syntax</td>
<td>VP word order</td>
<td>NP word order</td>
</tr>
<tr>
<td></td>
<td>Some subordinations, Conjunctions</td>
<td>Some subordinations, Conjunctions</td>
</tr>
</tbody>
</table>

As the preceding two tables suggest, there is considerable variability between mixed languages in the inheritance patterns of different lexical and functional categories; looking just at these two examples, it can be seen that Media Lengua draws effectively all its lexical material from Spanish and effectively all its inflectional material from Quechua, while Michif draws effectively all its verbal material (lexical and inflectional) from Cree and all its nominal material.
(lexical and inflectional) from French. Other argued mixed languages show still other patterns; Matras & Bakker (2003) report that Mednyj Aleut, for example, takes the majority of its lexicon and all its nominal inflection from Aleut, but all its verbal inflection from Russian.

**Code-mixing and Mixed Languages**

The superficial similarities between a code-mixed utterance and an utterance from a mixed language can be so striking, it might seem elementary to conclude that one is derived from the other. However, the diachronic relationship between code-mixing and mixed languages is rather murky; and there is considerable disagreement as to whether wide-spread code-mixing represents an early stage in the genesis of a mixed language. While a number of researchers (e.g. Myers-Scotton, 1998, 2000; Auer, 1999, 2000, Gardner-Chloros, 2000) have suggested that mixed languages clearly must arise out of code-mixing, Backus (2003:40) points out that mixed languages should not be thought of as a simple fossilization of code-mixing, and explicitly writes that “the compartmentalization that would turn a mixed lect\textsuperscript{11} containing mainly alternational [code-switching] into an unambiguous Mixed Language seems to be a linguistic impossibility.” Bakker (2003:129) goes further, stating that he is “convinced it is not the case [that mixed languages derive from code-mixing]”, and argues that, among other concerns, there is “no documentation of a transitory phase between the supposed CS behavior preceding the genesis of the mixed language” (Bakker, 2003: 129).

\textsuperscript{11} Backus (2003) makes a distinction between a mixed lect (“any kind of bilingual speech that is the unmarked way of speaking in the community in question” (Backus, 2003:2)) and a mixed language (characterized by being a single language system). A salient distinction between these two types of language systems is that a mixed lect is produced by bilingual speakers, whereas use of a mixed language does not imply bilingual competence.
Narratives concerning the genesis of mixed languages are somewhat limited by a lack of historical data; for most of the mixed languages spoken today, there is quite limited evidence of the linguistic behavior in the speech communities immediately prior to the development of those mixed languages. However, in the past decade several researchers have published work arguing that at least two Australian speech communities provide modern examples of the transition from a so-called mixed lect to a mixed language. In a 2005 paper, O’Shannessy presents data suggesting that a Warlpiri community in Northern Australia has transitioned over a generation or so from a speech community that “typically code-mixed between Warlpiri and Kriol or English” to one speaking “a new language, which is now learned by children as one of their two first languages.” (O’Shannessy, 2005: 32). Termed Light Warlpiri, this argued language combines lexical and morphosyntactic material from English, Kriol (an English-based creole), and Warlpiri (an indigenous Australian language in the Pama-Nyungan family), and demonstrates a unique auxiliary system not immediately attributable to its parent languages. A similar argument is made in McConvell & Meakins (2005) for the language Gurundji Kriol; their paper discusses the differences over a generation between the prolific Gurundji/Kriol code-mixing practiced by older speakers in the community and what they argue is a newly-minted mixed language spoken by a younger generation.

1.1.3. Bilingualism: Conclusions

This subsection has provided a brief overview of two types of language contact phenomena which arise in communities of predominantly bilingual individuals (even if such bilingualism dissipates over time). Code-mixing is common among Navajo-speaking communities, and—as discussed in chapter 2—some research (e.g. Schaengold, 2004; Haskan, 2007) has suggested that
certain subgroups within the Navajo community may have access only to a mixed Navajo/English code (“Bilingual Navajo”) which appears to meet the criteria outlined by Backus (2003) for a “mixed lect”. While not a focus of this dissertation, future careful documentation of this Bilingual Navajo could provide evidence as to whether the Bilingual Navajo speech community might be on the brink of grammaticalizing a new mixed language. In light of the recent research on Light Warlpiri and Gurundji Kriol, the prevalence of Navajo/English codemixing (and the declining numbers of Navajo community members who report fluency in standard Navajo) suggest a hospitable linguistic environment for an embryonic mixed language. Although I am aware of no current research or data documenting a (developing) mixed language in the Navajo speech community, research on bilingual Navajo speech currently remains quite limited and the absence of such findings may not be particularly informative.

1.2. Restructured Languages

This next section turns to language contact phenomena which arise in newly-formed communities where functional individual-level bilingualism is initially rare. In such linguistically fragmented speech communities—where the majority of community members do not share a common language—the need for communication can serve as a catalyst for the creation of innovative linguistic systems known as pidgins or creoles; the resulting languages are widely believed to display a certain degree of morphosyntactic simplification when compared to their genetic forebears. This section reviews the phenomena of pidgins, creoles, and finally the less clearly defined group of what have variably been called “semi-creoles” or “partially restructured languages”. Because the divisions between these three types of contact languages are not necessarily clearly defined (and because Navajo English has been suggested to fall under the
typological umbrella of creole languages), this section provides a brief treatment of each of the three typological classes.

The notion that the English spoken by various Native American groups may show characteristics of creolization has been proposed by a number of authors (e.g. Leechman & Hall, 1955; Dillard, 1972, 1975; Bartelt, 1982), and more recently Holm (2004) has suggested various vernacular forms of English spoken by Native American communities across North America might fall into the specific category of semi-creoles. This section is divided into three subsections; subsection 1.2.1. discusses pidgins, subsection 1.2.2. reviews the research on creoles, and subsection 1.2.3. discusses semi-creoles and the research connecting Navajo English to this class of contact languages.

1.2.1 Pidgins

Under the strictest definition of the term “pidgin”, supported by many researchers such as Bickerton (1985), Thomason & Kaufman (1988), and Siegel (2008), there are no enduring pidgin languages; the term is used to refer to the spontaneous and largely unstructured communication systems which arise in a community with no access to a common language. Such communication systems are considered unstable, and have no monolingual speakers, with all speakers being second-language users of what is an effectively transactional code. Although most researchers agree that such pidgins are usually short-lived, Hall (e.g. 1962, 1966) points out that pidgin languages may exist as long as the engendering sociolinguistic conditions persist, resulting in potentially long-lived (if dynamic) pidgins.

Pidgins are typically associated with transactional communities such as those associated with the Atlantic slave trade, those associated with Indian and Pacific plantation economies, and
those that arise along trade routes. Examples of pidgins which persisted long enough to be named include Chinook Jargon (an Amerindian, English, and French-based pidgin used by traders during the mid-1800s; e.g. Thomason, 1983), Basque-Icelandic Pidgin (a Basque, Spanish, English, German, French, and Dutch-influenced pidgin used by sailors and traders during the 1600s; e.g. Bakker et al., 1991), and West African Pidgin English (a pidgin continuum based on English and a number of mostly Niger-Congo languages (such as Kwa, Akan, Ewe, and Ga-Dangme), spoken widely along the West African coast from the mid-1600s until being supplanted by more structured creoles such as Ghanaian Pidgin English; e.g. Huber 1999). Many contact languages are popularly termed “pidgins”, such as Ghanaian Pidgin English or Hawaiian Pidgin English, but are considered by most mainstream linguists to be examples of what are more properly termed creoles.

1.2.2. Creoles

Unlike pidgins, creoles are considered by most mainstream linguists to be fully-fledged languages with regular morphosyntactic systems and intergenerational transmission. There are dozens of creole languages in use today, with several well-known examples including Haitian Creole (e.g. Lefebvre, 1985), Louisiana Creole (e.g. Marshall, 1997; Klingler, 2003), Gullah (e.g. Cunningham, 1970), Tok Pisin (e.g. Murphy, 1985, Smith, 2002), Papiamento (e.g. Jacobs, 2012), and Jamaican Patois (e.g. Patrick, 1999). There is no universally accepted definition of what constitutes a creole language: Singler (2006: 159) argues that “No linguistic litmus test of creole status exists—or could exist”, while McWhorter (2005, 2011) suggests that creole languages are simply exactly the set of languages whose grammatical system shows the following specific characteristics:
- **Phonological**: Little to no use of tone to distinguish monosyllables or grammatical categories.

- **Morphosyntactic**: Little or no inflectional morphology.

- **Semantic**: Little or no noncompositional\(^{12}\) combinations of derivational markers and roots.

  McWhorter (2011: 6)

Constructing a widely accepted linguistic definition for a creole language is made challenging by the fact that many researchers support a more holistic definition of the category, taking into consideration not only the linguistic characteristics but also the sociolinguistic history. This latter view is straightforwardly summarized in Mufwene (2002: 11440), who states that “there are no features that are peculiar to p[idgins] and c[reoles]”, and argues that the linguistic class is fundamentally defined by its social history.

Generally believed to arise out of pidgins, creoles are typically described as having a primary genetic parent language, known as the *lexifier* or *superstrate*, which provides the bulk of the lexicon for the newly minted creole language. Additional genetic influences are contributed by the *substrate* language(s), which typically provide the phonological system of the new creole language along with some contributions to the lexical and morphosyntactic systems. Creoles are argued to arise in particular sociolinguistic environments, wherein a language of prestige—which serves as the superstrate—imposes a certain degree of linguistic hegemony over the indigenous languages, which serve as the substrate. Presumably due to the modern global history of European colonialism, the lexifier for majority of known creoles today is one of several

---

\(^{12}\) Examples of non-compositional (semantically opaque) derivational morphology in English are words such as *understand* (which does not literally mean “to stand under something”) or *remember* (which does not mean “to ‘member’ again”).
European languages—primarily English, French, or Spanish—although creoles based on other languages such as Arabic or Mandarin do exist.

There are a number of theories of creole genesis, ranging from the theory that creoles arise out of superstrate native speakers’ attempts to simplify their language for foreigners (e.g. Arends et al., 1995) to the theory that creoles arise out through the relexification of the substrate language with material from the superstrate (e.g. Singler, 1996; Parkvall, 2000). In general, most theories of creole genesis acknowledge the influence of three main factors—the superstrate language, the substrate language(s), and linguistic universals—while disagreeing on the importance of each factor in shaping the resulting creole. However, the subject of creole genesis continues to be highly debated, and the short discussion presented in this dissertation does not attempt to provide a thorough treatment of the various proposals. This subsection reviews one of the more influential theories of creole genesis: the Language Bioprogram Hypothesis, which gives primacy to the importance of linguistic universals, and continues to receive wide support (particularly among researchers working outside the field of creole studies). The latter portion of this subsection discusses criticisms of the Language Bioprogram Hypothesis and briefly reviews some of the more prominent alternatives supported by creolists today. However, as theories of creole genesis are largely tangential to the goals of this dissertation—namely, documenting and describing the vernacular speech of a community of Navajo children—an in-depth review of different theories of creole genesis is not provided here.

The Language Bioprogram Hypothesis (LBH), first outlined in Bickerton (1981, 1988) has been for many years a leading theory of pidgin and creole genesis, and many researchers (e.g. Jackendoff, 2002; Aronoff et al., 2005; Hudson Kam & Newport, 2005; McWhorter, 2011)
continue to discuss the LBH as a current model of creole genesis. In brief, the LBH—a theory rooted in generative linguistics—proposes that creole languages develop according to biologically determined linguistic universals. Working primarily with data from Hawaiian Creole English, Bickerton (1981, 1988, 1999) argues that creoles are created abruptly by the first generation of children raised in a pidgin-speaking community, who, receiving no linguistic input from a cohesive grammatical system, fall back on their innate linguistic knowledge and produce a fully-developed linguistic system to serve as their native language. Under this view, which draws on the Lexical Learning Hypothesis outlined in Borer (1984), the grammatical system of a creole language is derived primarily from the adoption of unmarked or default feature specifications for various functional and lexical categories and only minimally from genetically acquired influences. A corollary to the LBH is that creole languages form a typological class, sharing a set of synchronically identifiable characteristics.

Proponents of the LBH (e.g. Pinker, 1994; Senghas, 1995) often point to the example of Nicaraguan Sign Language (e.g. Kegl & Iwata, 1989), a signed language which arose in the 1980s in a Nicaraguan vocational school for deaf children. The children in this school—which exposed the deaf students to oral and crude manually coded Spanish—had no meaningful exposure to existing sign languages, and had been largely reliant on rudimentary homegrown gestural systems. In less than a decade after the founding of the school, students who had entered as young children were found to be producing grammatical constructions such as regularized verbal agreement and tense marking, and researchers such as Kegl & Iwata (1989) reported that the children were using a fully developed sign language. Characterized by an effectively agrammatical mixture of gestures and other signs, the conditions in which NSL arose have been
argued by some (e.g. Kegl & Iwata, 1989; Senghas, 1994) to be comparable to the conditions encountered by children being raised in a pidgin-speaking community. The emergence of a linguistic system containing grammatical elements not found in any of the linguistic input has been suggested to be strong evidence for the validity of the Language Bioprogram Hypothesis.

However, support for the LBH has fallen away among some researchers within the field of creole studies, and a number of critics (e.g. Roberts, 2000; Ansaldo & Mathews, 2007; Siegel, 2007, 2008; DeGraff, 2003, 2005; Mufwene, 2009) have dismissed the theory as failing to accurately describe particular historical instances of creole genesis. Some of the main complaints levied against the LBH are that it promotes the idea that a creole grammar emerges fully formed (rather than gradually) and that it promotes the idea of what is known as “creole exceptionalism”, namely that there is a set of features unique to creoles, such that the class of creoles forms a synchronically identifiable typological class. Critics of this last notion (e.g. Mufwene, 2007, 2008; DeGraff, 2005) argue that the class of creole languages is necessarily defined by a particular diachronic profile, and that attempts to define the class independently from the sociolinguistic history are unproductive.\(^\text{13}\)

Alternatives to the LBH include the account promoted by researchers such as Mufwene (2008), DeGraff (2003) which argues that creole languages have arisen through “a gradual development from the colonial koiné ancestors spoken as vernaculars by the Creole populations of both European and non-European descent in the homestead communities that preceded the plantation communities” (Mufwene, 2007b: 67) This view—which gives primacy to the role of

---

\(^{13}\) Supporters of the notion that creoles form a legitimate synchronic class (e.g. McWhorter, 2005; Krämer, 2013) point out that defining creoles as a purely sociohistorical class would suggest that creole studies should be the purview of anthropologists and historians rather than linguists, as there would be no meaningful linguistic reason to differentiate between the study of Jamaican Patois and the study of Hungarian.
the superstrate language in shaping the creole—suggests that creole languages do not differ categorically from other non-standard dialects, and that many of the so-called “creole” features are simply non-standard features inherited from vernacular forms of the so-called superstrate; this view further supposes that there is no true linguistic distinction between the putative creole and any other synchronically variable descendants of a language. Other alternatives include the relexification hypothesis (e.g. Lefebvre, 1998) which proposes that creole languages have adopted the lexicon of the superstrate language but the morphosyntactic structure of the substrate(s); this proposal has been criticized for a failure to account for variability in the morphosyntactic structures of the substrate languages.

1.2.3. Semi-Creoles

While there is wide support for the belief that pidgins and creoles are distinct classes of linguistic phenomena, there is less consensus about the argued category of semi-creoles, or partially restructured languages. For terminological consistency in this following discussion, the term ‘semi-creole’ is arbitrarily chosen, although this choice is not intended to reflect any theoretical distinction. Supporters of this view (e.g. Holm, 2001, 2004; Winford, 1997; Lipski, 2006; Sutcliffe, 2007) suggest that certain languages lack the necessary sociolinguistic and linguistic characteristics which distinguish a true creole, but nevertheless show significant contact-induced structural changes\(^{14}\). Languages argued by Holm (2004) to fall into this category include African American English, Afrikaans, Brazilian Vernacular Portuguese, Non-standard

\(^{14}\) Mixed languages and semi-creoles are similar phenomena, but it is important to note that these are typically considered to be two separate and distinct classes of contact languages. Mixed languages are argued to show a regular pattern of genetic inheritance from their two parent languages, with entire categories of lexical items being inherited without material morphosyntactic simplification. Semi-creoles, on the other hand, are argued to show the patterns of phonological and morphosyntactic simplification attributed to creolized languages. The distinction between semi-creoles and creoles is further discussed throughout this subsection.
Caribbean Spanish, and various vernacular lects of Réunion French; Holm (2004:4) further notes that the “nonstandard English of American Indians” seems “likely to have undergone a similar process”—namely that of partial restructuring—while Lipski argues for the inclusion of Afro-Bolivian Spanish in this class.

The proposal for a specific category of semi-creole languages is somewhat controversial, with many of the arguments in its favor focusing primarily on the sociolinguistic history rather than on the absence or presence of specific linguistic properties. Taking, for example, the widely debated research on the diachronic origins of African American English (also known as African American Vernacular English, and variably abbreviated AAE or AAVE)\(^\text{15}\); many researchers have argued that the modern vernacular is not a semi-creole, but instead is the result of a process of decreolization (e.g. Dillard, 1972, 1992; Labov & Harris, 1986; Rickford, 1998), wherein prolonged contact with a creole’s lexifier results in gradual loss of distinguishing creole features and ultimate assimilation into the superstrate language. Arguing against this idea, at least with regard to AAE, Holm (2004:2) claims that “no one could find reliable historical evidence of the widespread, stable creole from which AAE has supposedly decreolized”, and instead proposes that languages such as AAVE must instead be considered African-influenced dialects of English which have undergone the process of ‘partial restructuring’ rather than complete creolization\(^\text{16}\).

---

\(^{15}\) While Navajo English and AAE are not generally believed to share common linguistic roots—other than a shared superstrate language—this discussion is relevant to the current project in that Navajo English has been suggested to fall into this putative category of semi-creoles. As the linguistic research on AAE dwarfs that on Navajo English, the following paragraphs explore the more thoroughly researched case of AAE as an illustrative case of the active and ongoing debate surrounding both the synchronic and diachronic validity of the semi-creole class.

\(^{16}\) An alternative argument for the genesis of AAVE, more widely supported as an alternative to the post-creole conjecture (e.g. Montgomery, 1991; Tagliamonte and Poplack, 1988, 1993; Poplack & Tagliamonte, 2001; McWhorter, 2005), holds that AAVE is the descendent of English dialects spoken by both white and black Americans in the 1800s, that the African roots of the modern-day vernacular are not significant, and that the current
The debate concerning the origins of AAVE highlights the strong sociolinguistic components to the proposed semi-creole profile; supporters of the distinction suggest that a creole arises in situations where native speakers of the superstrate language are overwhelmingly outnumbered by native speakers of the substrate language(s), while a semi-creole arises in situations where “more than a small minority (20-25 percent) of a colony’s founding population are native speakers of the colonists’s language” (Holm, 2010: 258). However, this focus on sociolinguistic factors should not obscure the fact that semi-creoles are also argued to represent a typologically distinct class from full creoles. The contrast between AAVE and the fairly uncontroversially categorized creole Gullah (e.g. Turner, 1949; Weldon, 1998, 2003, 2007; Frank, 2007; Hackert & Holm, 2009; Moody, 2011) serves as an illustrative example of the purely linguistic factors suggested to distinguish the two types of contact languages; argued areas of difference include the prevalence of null copula use, the characteristics of the system for marking tense, mood, and aspect in the verbal complex, and the prominence of West African etymological roots in the lexicon.

As this discussion on AAVE has suggested, an objection to the proposal for a discrete category of semi-creoles is the difficulty of identifying a clear linguistic delineation between semi-creoles and ‘true’ creoles; a set of functional linguistic diagnostics would strengthen the argument for semi-creoles being structurally distinct rather than simply arising in particular sociolinguistic circumstances. As evidence of the linguistic (rather than purely social or historical) validity of the semi-creole category, Holm (2004) presents a set of 18 features which

ethnic associations are attributable to ethnically-linked processes of linguistic assimilation and dissimilation. This latter view suggests that the supposed African influences—such as serial verb constructions—in modern AAVE are largely artifacts of language-internal change with only coincidental parallels to West African morphosyntax.
he argues are commonly found in partially restructured languages. The table below shows these features, along with a tabulation of their presence or absence in each of Holm’s argued set of partially restructured languages, as well as their superstrate and presumed substrate languages.

16. Key Morphosyntactic features in partial restructuring
Reprinted from Holm (2004: 138)

<table>
<thead>
<tr>
<th>Morphosyntactic Feature</th>
<th>AAE Sb</th>
<th>AFR Sb</th>
<th>BVP Sb</th>
<th>NSCS Sb</th>
<th>VLRF Sb</th>
<th>E</th>
<th>D</th>
<th>P</th>
<th>S</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb Phrase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal Morphology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero 3s PRES infl.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zero 1p PRES infl.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zero PAST infl.</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aux./preverbal marker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic influence</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Negation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative concord</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Discontinuous double</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Non-verbal predicates</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Noun Phrase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero plural infl.</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unbound pluralizer</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Associative plural</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No agreement in NP</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0 +</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Possession</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[possessor 0 possessed]</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0 +</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>[possessed 0 possessor]</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>0 +</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pronouns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reduced case marking</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zero reflexive pronoun</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Order</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QW S-V/Aux (direct)</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dependent clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero subject REL</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Zero subordinator “that”</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total number of +’s</td>
<td>15</td>
<td>9</td>
<td>13</td>
<td>9</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sb = relevant substrate language(s)  
iinf. = inflection  
AAE = African American English  
AFR = Afrikaans  
BVP = Brazilian Vernacular Portuguese  
NSCS = Non-Standard Caribbean Spanish  
VLRF = Vernacular Lects of Reunionnais French

E = English  
D = Dutch  
P = Portuguese  
S = Spanish  
F = French
As table (16) illustrates, a disappointing observation about these proposed semi-creole features is that many of them are in fact found in greater numbers in the presumed substrate languages than the argued semi-creoles themselves; between 14 and 16 ‘semi-creole features’ are found in each of the presumed substrate languages, while each of the proposed semi-creoles is argued to have between 9 and 15. Given this apparent lack of semi-creole uniqueness, it is not clear that the features presented by Holm are sufficient as a linguistic diagnostic to identify members of this presumed class. Writing of the debate over accepting a semi-creole origin or a post-creole origin for languages such as AAVE, Mufwene (2007: 73) writes that “The greatest problem with the semi-creolization hypothesis as the reversal of the decreolization hypothesis lies in difficulties in determining what “creole features” really are or what a “prototypical creole” is supposed to be”, and it is indeed not clear that the features identified by Holm are any more characteristic of languages in the proposed class of semi-creoles than they are of so-called full creoles or of any randomly sampled languages.

Native American English: Evidence of Creolization or Semi-creolization?

The relevance of this current discussion, of course, is the potential applicability of the semi-creole typological class to Navajo English. As mentioned in the introduction to this section, the vernacular English spoken by some Native Americans has also been suggested by Holm (2004) to fall into the category of semi-creoles. Explicit discussion of this claim, however, is limited; Holm does not elaborate further, and it is not clear whether he is referring to any specific speech communities.

Researchers such as Penfield (1977), Bartelt (1981, 1982a,b) and Schaengold (2004) have reported a robust non-standard vernacular form of English used by members of Navajo
communities. Schaengold (2004:31) claims that “the kinds of differences found between [Navajo English] and Standard American English can be attributed mainly to Navajo-English contact . . . the particularly identifiable characteristics can be explained best as substrate interference from Navajo.” Perhaps most obviously characterized by prominent phonological effects, this so-called Navajo English is also argued to show influences from Navajo throughout its morphological, syntactic, semantic, and pragmatic systems. While many of these features might be easily explainable as incomplete acquisition of English as a second language, Schaengold points out that many modern Navajos are effectively monolingual English speakers, and have little or no functional command of the Navajo language. Describing exactly this phenomenon, Leap (1975:1) writes that “. . . I view Indian English as a continuation of the speakers’ native language tradition; in situations where the Indian language itself is no longer a part of the community’s verbal repertoire, it may be the only continuation of the tradition which now remains.”

The suggestion that Navajo English in particular might show some characteristics of a creole dates back to at least the 1980s, when Bartelt (1982: 71) argued that “for some at least, [Navajo English] has become a creole of a sort.” Discussing a range of tense, mood, and aspect (TMA) markers found in Navajo English examples taken from letters and other written communication, Bartelt notes that the Navajo English TMA system shows many features associated with prototypical creoles or post-creoles, generalizing from this that “creolization as well as decreolization may be processes relevant to American Indian English” (Bartelt, 1982:79). As evidence for this suggestion, Bartelt refers to the presence of non-standard features which he attributes to Navajo influence, but clarifies that there are far more commonalities with standard
English than in acknowledged creoles such as Hawaiian English. However, as has been pointed out for AAVE, there appears to be no meaningful evidence of any earlier, more “fully” creolized form of Navajo English, making a decreolization analysis more challenging to accept.

In concluding this subsection, the linguistic characteristics of the proposed category of semi-creoles seem to be of limited utility in identifying proposed semi-creoles; a holistic assessment incorporating both sociolinguistic and purely linguistic factors appears necessary for distinguishing the category of semi-creoles from the category of ‘true’ creoles. However, the lack of obviously discrete categories is not seen by all linguists as a death knell for the typological division, and McWhorter (1998: 811) offers the following commentary in support of the synchronic validity of the semi-creole subclass: “There are surely no distinct lines between the classes creole, semi-creole, and regular language. The inherently gradient nature of language restructuring is such that it would be quite futile—and ultimately of unclear utility—to propose any metric of creoleness. This, however, no more invalidates the terms of their usefulness than the non discrete nature of growth invalidates the terms PUPPY and DOG.”

1.3. Languages in Contact: Conclusion

This chapter has provided an overview of language contact phenomena. Working first through contact language phenomena in communities with high levels of individual-level bilingualism in the parent/source languages, the first section of this chapter is dedicated to a discussion of code-mixing and the arguably derived phenomenon of mixed languages. The second section of this chapter treats contact language phenomena that occur in communities where the parent languages are historically spoken by mutually exclusive groups of people; this second section outlines the phenomena of pidgin languages and creoles, and finally the case of semi-creoles or
partially restructured languages. It is important to note that each of these language contact phenomena is a field unto itself, and is it impracticable to provide a thorough treatment in the few pages that have been assigned here. Rather than serving as an exhaustive treatment of research on contact languages, the information covered in this chapter has the modest goal of providing a basic framework for assessing the merits of the data-driven discussions in chapters 4 and 5.
Chapter 2: Navajo Language & Linguistics

This chapter provides an overview of the research on Navajo linguistics, the sociolinguistic context in which Navajo is spoken, the morphosyntax of the Navajo language, and the literature on Navajo-English contact phenomena. Section 2.1 addresses the history of Navajo linguistics and language use, Section 2.2 provides an overview of Navajo morphosyntax, Section 2.3 reviews the literature on Navajo-English code-mixing and other language contact phenomena, and Section 2.4 reviews and concludes the chapter.

2.1. Navajo Linguistics: Past and Present

One of the greatest boons to students of the Navajo language—the language has been more thoroughly studied and documented than any other indigenous language—has also been a rather mixed blessing to scholars and often research-weary Navajo communities. Willie (1991:11) observes somewhat pointedly that “[t]he Navajo, as a people, and their language, have been subjected to much study; in fact, there is more published research on this Native American tribe than any other.” This is not to suggest that the subject of Navajo linguistics has been exhausted, of course; the amount of research on Navajo—or in fact the combined body of work on Navajo and all the other polysynthetic languages in the world—is dwarfed by the massive amount of research into languages such as English. However, years of research, often only minimally informed by the interests and concerns of the tribal members themselves, have fostered deep-rooted community suspicions concerning non-Navajo researchers. Greater tribal involvement in the research process, as is discussed briefly in this chapter and in more detail in chapter 3, is helping to develop more community-focused research approaches in the field of Navajo linguistics. This subsection provides a short orientation to the field of Navajo linguistic research;
Section 2.1.1 reviews the history of Navajo linguistics, Section 2.1.2 discusses current trends in the field, Section 2.1.3 reviews the educational and sociolinguistic context of Navajo use, and Section 2.1.4. discusses current efforts towards language revitalization and language maintenance. Section 2.1.5. provides a brief summary of the subsection.

2.1.1. History of Navajo Linguistics

The earliest known documentation of Navajo\textsuperscript{17} comes from a brief 39-word list appended to a military journal (Simpson 1849), followed by a series of word lists compiled by anthropologists or members of the military, including a detailed 15-page word list prepared in 1852 by Captain J. H. Eaton\textsuperscript{18}. While a number of other word lists were compiled over the latter half of the nineteenth century (Pilling’s 1892 survey of Native American language research identifies 24 Navajo word lists in existence, varying in breadth and quality), more sustained documentation was not carried out until the early part of the twentieth century. Beginning shortly after the arrival of the first Franciscan missionaries into Navajo territories in 1898 (Bahr, 2004), Franciscan friars began translating religious texts into Navajo and producing the first Navajo

\textsuperscript{17} An appendix to Barreiro (1832) presents a list of 10 words ostensibly in Navajo with Spanish equivalents, noting that “[l]a lengua de estos gentiles es muy facil de pronunciar” (Barreiro 1832:10). This list is perplexing for its lack of any perceptible correspondence to modern-day Navajo, with such words listed as tinde for “man” (the modern Navajo word is hastiin), chihuata for “woman” (the modern Navajo word is asdzani), nortin for “father” (the modern Navajo is azhé’ê), and thastia for “mother” (the modern Navajo being amá). Diachronic variation seems unlikely to have produced such drastic changes in the Navajo language, and it is probable that this list represents some other unidentified geographically proximate language. Given these inconsistencies, the Barreiro word list may be reasonably omitted from the record of written Navajo.

\textsuperscript{18} Eaton’s word list contains a number of basic errors, such as listing the Navajo word hahs thin (which in fact means “man”) for the English “old”, or the Navajo to she kái da (which means “(s)he’s not my kin”) for the English “enemy”. However, it is fair to note that Captain Eaton made no claims of being a trained linguist, and appears to have prepared the word list solely in response to a questionnaire distributed by ethnographer Henry Schoolcraft. As pointed out by Young & Morgan, the word list was almost certainly prepared through a series of Navajo-Spanish and Spanish-English interpreters, as “[f]or obvious reasons, there was no Navajo in 1852 who could speak or understand the English language, and there was no speaker of English who knew Navajo” (Young & Morgan 1987:iv). Given the handicaps with which Eaton was working, the errors are quite understandable.
reference materials. In 1910 a group of Franciscan friars led by Father Berard Haile published the *Ethnologic Dictionary of the Navajo Language*, followed in 1912 by the *Vocabulary of the Navajo Language*, in 1926 by the *Manual of Navajo Grammar*, and in the 1940s by the *Learning Navajo* series. In 1910 a Protestant missionary by the name of Fred Mitchell published *Diné Bizâd: a Handbook for Beginners in Navajo*. Around this time a group of Wycliffe translators began work on a Navajo translation of the Bible, although it was not until 1956 that a complete version of the New Testament was finished.

These early texts were soon followed by a flurry of academic interest in the language, and a number of articles were published, including Goddard (1911), Hoijer (e.g. 1943, 1945a,b, 1946a), and Sapir (1936, 1949, 1967). However, the most important development in Navajo linguistics was the 1943 publication of the first edition of *The Navajo Language*. Compiled by Navajo linguist William Morgan and the non-Navajo linguist Robert Young, the massive single-volume text reached far further than any of the earlier reference materials and provided a richly detailed grammar addressing the complex verbal inflection system, as well as a cross-referenced colloquial dictionary including invaluable examples of lexical items collected from naturalistic native speaker speech. This text was the first in a series of reference materials produced by the fruitful Young & Morgan partnership, with a number of updated volumes appearing over the next half century (Young & Morgan, 1980, 1987, 1992, Young 2000). The breadth and depth of the information provided in Young and Morgan’s work is unparalleled for any other indigenous language, and has fostered much of the subsequent research into Navajo linguistics and the Navajo language. In fact, Holm (1996) points out that beyond having provided the tools that have greatly facilitated the study of Navajo linguistics, Young and Morgan’s grammars and
dictionaries may be the primary reason that any native Navajo speakers today are able to read and write in Navajo. The contribution of these reference materials to the body of work on Navajo—and to the Navajo language community—cannot be overemphasized.

2.1.2. Navajo Linguistics Today

As discussed above, considerable academic effort has been dedicated to documenting and analyzing the Navajo language. Today, hundreds of papers have been published on the Navajo language, with topics ranging from the ritual language in traditional stories and songs, to the efficacy of immersion programs and other second-language learner programs, to theoretical topics such as the morphosyntax, semantics, pragmatics, phonetics, and phonology of the language. Thanks to the popularity of the Navajo language as a research subject, the bibliography of Navajo texts has ballooned to proportions no longer feasibly itemized in a few paragraphs. The substantial research already conducted has made linguistic research on Navajo easier than on many other indigenous languages that do not have a comparable groundwork of basic documentation and analysis.

The Navajo Language Academy (NLA) is currently perhaps the most important promoter and facilitator of Navajo linguistic research. A small group of Navajo linguists and advocates, the NLA describe themselves as a “non-profit educational organization devoted to the scientific study and promotion of the Navajo language,” and to this end they organize annual workshops on topics in Navajo linguistics, provide support and training for Navajo language researchers and teachers, promote Navajo language research, and maintain reference and research materials related to Navajo linguistics and language research.
Research on theoretical aspects of the Navajo language such as morphosyntax continues—more recent research in this area includes Aissen (2000), Fernald & Perkins (2006), Goertz et al. (2006), Hale (2000)—but the greatest bulk of modern Navajo research deals with issues related to language identity and language revitalization and maintenance. Researchers working on language acquisition include Saville-Troike (1980, 1996), Courtney & Saville-Troike (2002), Gentner & Boroditsky (2009), and Shepard (2012), and those working on bilingualism and language identity include Lee (e.g. 2007, 2009), McCarty, Romero-Little, & Zepada (2006), Schaengold (2004), Spolsky (2002), Webster (e.g. 2008, 2010a, 2010b), and Jacobsen (2012). In general, the field of Navajo linguistics seems to be trending towards promoting language revitalization and bilingualism, and most modern linguistic work related to Navajo is rooted in practical questions concerning language identity and education. This trend is not terribly surprising; as members of an endangered language community, many Navajo tribal members—who are increasingly responsible for directing linguistic research into their heritage language—are extremely interested in issues related to the role of the Navajo language in the modern world, and are at most passingly interested in questions pertaining to theoretical linguistics.

Unlike in earlier eras, where indigenous language research was primarily conducted by non-indigenous researchers and often had little connection with the concerns and preferences of the speakers themselves, modern Navajo linguistics is strongly influenced by the interests of the Navajo people. As discussed in greater detail in chapter 3, the Navajo tribe increasingly maintains the right to regulate all research related to the Navajo cultural heritage (including the Navajo language), and operates under a general policy of approving only research that is deemed pertinent to the well-being of the Navajo people. The trend towards practical issues of language
education and language revitalization is a reflection of the tribe’s asserted right to control research into their language; esoteric linguistic questions (such as those arising out of pure syntax or semantics) are often not seen as having an obvious connection to the concerns and interests of the Navajo people, and are not as likely to receive research approval\textsuperscript{19}.

\textbf{2.1.3. Navajo Language in Use: Education and Cultural Attitudes}

With roughly 175,000 speakers (2000 U.S. Census), Navajo is currently the most widely-spoken indigenous language in North America, yet educators, researchers, and community members have unanimously observed (e.g. McCarty & Dick, 1996; Platero, 2001; Schaengold 2004; Benally & Viri, 2005; Haskan, 2007; Jackson, 2008, Shepard, 2012) that the Navajo-speaking community is rapidly undergoing a process of \textit{language shift}; the process through which a community moves from speaking one language to another. The roughly 25\% of Navajo speakers who reported themselves to be effectively monolingual on the 2000 Census are almost all over the age of 65; tribal members under the age of 18 are most likely to be monolingual speakers of English or English-dominant with a basic knowledge of Navajo, and only a handful of younger Navajo tribal members are even relatively balanced bilingual speakers of Navajo and English. Surveys and anecdotal accounts indicate that the rate of change is astonishingly rapid; Lee (2009) reports that in one rural community roughly 90\% of kindergarteners at the local community school were Navajo speakers in the 1980s, while roughly 10 to 15 years later a former school director reported that 90\% of the kindergarteners were identified as primarily

\textsuperscript{19} As discussed in greater detail in chapter 3, the Navajo tribe currently claims the right to regulate any research conducted within the bounds of the Navajo Nation, whether linguistic or otherwise. It is common practice to consider already-published research and other linguistic material as public domain, and therefore out of the jurisdiction of any individual tribe or other group. It is worth noting that there is some murkiness surrounding the goal of regulating access to linguistic knowledge, where in some cases the researchers may be native speakers of the language in question and are themselves arguably rightfully empowered to control access to their language.
English speakers. While there is certainly geographic variation—children from smaller, more isolated communities are widely believed to have somewhat higher rates of Navajo proficiency—there is little disagreement that nearly all Navajo children today are far more fluent in English than in their heritage language.

The current language landscape on the Navajo reservation is the arguably natural result of a century and a half of linguistic hegemony. From 1882 to the latter half of the twentieth century, schools operated on the Navajo reservation by the Bureau of Indian Affairs enforced strict English-only policies for their students and “[s]chool was . . . the principal weapon of linguistic colonialism” (Spolsky, 2002: 143). Attendance at such schools was enforced by government officials, who would collect children—often against their families’ wishes—and forcibly enroll them in the residential schools (e.g. McCarty, 2002; Shepard, 2012), where they were punished for speaking any language other than English.

By the early 1900s a distinctive “boarding school Navajo” characterized by English lexical and grammatical influences was being described (Harvey, 1974; Schaengold, 2004). Unfortunately, the lack of documentation means that almost no examples of this early English-influenced Navajo are available for reference, and the only recorded evidence of this mixed code comes from passing mentions in other texts. Today, some older speakers can still speak boarding school Navajo, but there is limited documentation of this code. Working with these speakers—an opportunity which may not exist for much longer—would provide valuable insights into the early history of Navajo/English bilingualism.20

---

20 An in-depth investigation of boarding school Navajo is outside the scope of this current research project. Personal observation and informal communication suggests that even modern standard Navajo may in fact show material influences from English, many of which may be traceable to the early boarding school code, but this has not
While the economic incentives associated with English cannot be overlooked, anti-
Navajo attitudes in the schools have certainly contributed to the accelerating language shift
(Henry, 2003; Schaengold, 2004). Today, English-Navajo code-mixing is robustly attested,
particularly in informal settings (Schaengold, 2004). Cross-linguistic utterances are frequently
viewed negatively by more traditional Navajo speakers; Young & Morgan (1987:7) report that
“[t]he Tribal Council has generally insisted upon linguistic purity, sometimes stopping speakers
in the middle of their discourse to insist that they speak only one language at a time.” However,
there is evidence (e.g. Haskan, 2007) that many of the youngest speakers of Navajo have access
only to this mixed code, and Parsons-Yazzie (1995) reports that one of the most common reasons
cited by parents for lack of Navajo use in the home is children’s embarrassment over imperfect
Navajo. 21

Some researchers speculate that code-mixing could accelerate language shift (e.g. Minett
& Wang, 2008), and it is true that linguistic environments which foster high levels of code-
mixing can be unstable over time, shifting away from bilingualism and towards monolingualism.
However, speech communities such as the Old Order Amish in Pennsylvania—where

currently been explored in any sustained manner. Certainly, social practices have had a noticeable linguistic impact; for example, the word yiits’iqz meaning “he kisses her”—was traditionally reserved for non-romantic kissing (such as that between a mother and child), which was the only type of kissing typically practiced in traditional Navajo communities, but the term has undergone semantic extension and is now used by monolingual Navajo speakers to refer to the now-common practice of romantic kissing. Sustained research into the original boarding school Navajo and the early stages of Navajo/English bilingualism would be a valuable direction for future research into Navajo language contact effects.

21 A further complication limiting effective inter-generational Navajo communication is the diachronic change that occurs in any speech community. Overlaid across the changes that occur within any monolingual speech community are those changes attributable to English influences on the Navajo used by younger speakers. Lee (2009) points out that Navajo community members of all ages have a strong tendency to perceive the Navajo spoken by older, often monolingual speakers to be purer and more correct than that spoken by younger, usually bilingual community members. Aside from a few studies on code-mixing, almost no research has looked at the influences English has had on the modern Navajo language. The converse, Navajo influences on the English used by tribal members, is discussed in section 2.3.2 of this chapter.
community members have maintained relative stability as bilingual speakers of speakers of English and Pennsylvania German for over 200 years—stand as clear reminders that bilingualism can be a long-term linguistic situation for a community, and that the mere presence of code-mixing should not be considered a death knell for one language or the other. In fact, there is reason to suspect that code-mixing—perhaps in specific contexts—may be a valuable means of promoting heritage language use. A recent and rapidly expanding body of research into the use of code-mixing in the language classroom (e.g. Bloom, 2008; Boztepe, 2009; Lin, 2008; Macaro, 2001; Mati, 2004; Moodley, 2007; Palmer, 2009; Probyn, 2009; Raschka et al., 2009; Reyes, 2001; Tien, 2009; Wei & Martin, 2009) suggests that code-mixing can be a useful pedagogical tool in the foreign language classroom and in bilingual immersion settings.

Schaengold (2004) reports that some Navajo language teachers, apparently following the intuitions explored in the research cited above, encourage the use of code-mixing as a useful path to Navajo fluency. Personal observation suggests that this in fact common practice in Navajo immersion classrooms, and students are encouraged to use Navajo in whatever capacity possible. This pedagogical approach notwithstanding, code-mixing is a stigmatized practice in the Navajo-speaking community, and has historically been labeled a sign of ‘semilingualism’, or inability to speak either language well. Indeed, the question of ‘semilingualism’ is related to the question of whether young speakers do, in fact, have access to two separate codes—juxtaposed by means of code-mixing—or whether these speakers are lapsing into a (semi-)regularized ‘mixed’ language similar to Paraguayan Jopará or Gurundji Kriol. While the term ‘semilingual’ is both incendiary and misleading, the question of whether young Navajo/English speakers may be truly considered bilingual is quite pertinent. As discussed above in section 1.1.3, it is not unknown for historically
diglossic communities to undergo a shift towards monolingual use of a grammaticalized mixed code—such as the pattern argued to be observable in modern-day Light Warlpiri communities (e.g. O’Shanessy, 2005, 2008, 2009)—and some community members have expressed concerns that such a trend might be present in the Navajo community. While there are—as discussed further in this chapter—many examples of clear code-mixing that do not appear to represent a regularized mixed code, it is quite plausible that at least some Navajo speech communities may be in the process of developing a mixed code. This question is revisited in the discussion section of this dissertation.

A further hindrance to the Navajo language revitalization effort is the widely acknowledged fact that many younger Navajos associate the Navajo language with lower socioeconomic status, and even some who are functional Navajo speakers may prefer to speak in English (Lee, 2007). As in many minority communities whose traditional language has been stigmatized and repressed, community members themselves may be cautious about using a language with strong historical associations of poverty, limited access to education, and reduced employment opportunities.

2.1.4. Efforts towards Navajo Language Maintenance and Revitalization

Supporting language revitalization and maintenance is a topic of great interest to many Navajo tribal members. While it is true that many community members have internalized the stigmatization of the Navajo language, it is also true that Navajo fluency is considered to be an extremely important mark of cultural authenticity (e.g. Aronilth, 1994; Henry, 2003) for many tribal members. Many positions within the tribe—including most elected positions within the tribal government, eligibility for the Miss Navajo Nation beauty pageant, and service as a
traditional religious leader—are only open to those who demonstrate some level of Navajo fluency. A number of tribally-administrated scholarships and forms of educational support—such as the Chief Manuelito Scholarship—require knowledge of the Navajo language. As the rapid acceleration of language shift becomes more apparent, tribal members have been facing concerns about the potential loss of their heritage language and have been working to encourage and promote the use of the Navajo language in tribal communities. A generation ago it seemed inconceivable that the Navajo language could become endangered; today it is rare to find a child who is considered a truly fluent speaker of Navajo.

In response to the advancing language shift, Navajo-immersion and Navajo language programs have been springing up across the reservation and nearby communities. Early Navajo immersion programs followed a dual-language immersion model, where the incoming class of kindergarteners contained monolingual speakers of both Navajo and English, allowing a peer-to-peer learning model where both English- and Navajo-speaking students learned their second language from native speaking peers. Although such programs were viable as recently as 20 years ago—in 1992, Tuba City Public Schools in Arizona began a two-way immersion program for Navajo students—today there are too few monolingual Navajo-speaking kindergarteners to support dual-immersion education (Wiles, 2006; Shepard, 2012). Current Navajo immersion programs follow typical single-language immersion models, generally providing instruction entirely in Navajo for first-year students (usually the incoming kindergarten class), then adding in English in the higher grades until students receive half or more of their instruction in English and the remainder in Navajo. The handful of true Navajo immersion programs include the tribally-run Rough Rock Community School in Rough Rock, AZ (serving students from
kindergarten through 12th grade), the Tséhootsooi Diné Bi’ólta’ (Window Rock Navajo School) in Fort Defiance, AZ22 (serving students from kindergarten through 8th grade), Rock Point Community School in Rock Point, AZ (also serving students from kindergarten through 12th grade), a recently opened immersion program at Eva B. Stokely Elementary in Shiprock, NM (currently offering only lower level immersion classes, but planning to offer immersion classes up to 5th grade), and the off-reservation Puente de Hozho Language Academy, operated by the Flagstaff Unified School District (offering Navajo immersion classes from kindergarten through 8th grade).

While only a handful of full Navajo immersion programs are in operation, in response to a 1986 Tribal Council ruling that Navajo must be taught at every school in the Navajo Nation (Reyhner, 1994), some form of Navajo language classes are offered at every school on the reservation. These non-immersion programs vary greatly—some offer as little as 30 minutes of Navajo instruction each week while others provide an hour or more daily—making the transmission of significant language skills challenging in some cases. However, the institutional support for the Navajo language is a powerful indication to students that their language and culture are welcomed and encouraged in the schools; a stark contrast to the reception Navajo children received only a few decades ago.

Navajo language classes are also offered at the college level in a number of communities. The tribally-run Diné College offers Navajo classes at six of its eight campuses scattered across the main Navajo reservation. Navajo language and culture classes are also offered by several non-tribal universities near the reservation, including the University of New Mexico (UNM),

22 Tséhootsooi Diné Bi’ólta’ is run on the site of The Navajo Times—a tribally-run daily newspaper—reported in February of 2013 that financial pressures threaten the future of the immersion program at Tséhootsooi Diné Bi’ólta’.
Arizona State University (ASU), Northern Arizona University (NAU), the University of Utah, and (intermittently) Brigham Young University. The range of courses offered varies; Diné College offers courses from basic introductory classes to literacy for native speakers, UNM offers classes for all levels of fluency, as well as literature classes and classes on Navajo linguistics, while the courses at the University of Utah are generally limited to beginning and intermediate classes for non-native speakers.

Support for the Navajo language is not limited to the classroom. In 2010, the commercial language-learning software giant Rosetta Stone released Navajo language learning CDs for beginning and lower-intermediate students. The materials were produced in partnership with the non-profit group Navajo Language Renaissance as part of the Rosetta Stone company’s Endangered Language Projects program, and currently consist of software CDs marketed as levels 1 and 2 (for comparison, Rosetta Stone produces levels 1-5 for global languages such as French or Spanish). The materials have received mixed reception in the Navajo community; while many Navajos have expressed excitement about the opportunity for children and younger tribal members to study their heritage language, the Navajo Times reported in late 2010 (Zah, September 11, 2010) that some tribal members had expressed concerns about non-Navajos having access to the Navajo language. Other concerns include the fear that, despite strong Navajo involvement in the development and production of the materials, the language CDs fail to take traditional Navajo cultural concerns into consideration in lessons such as teaching the conjugation of the verb meaning “to drink” first in its intransitive form (usually interpreted as meaning “to drink alcohol”) rather than its transitive form (where the verb has a more neutral interpretation). Still another concern is that the software prioritizes a single dialect of Navajo;
modern Navajo in fact exhibits a good deal of linguistic heterogeneity\textsuperscript{23}, and by selecting the generally accepted standard dialect for the Rosetta Stone software, the program—which rejects alternative or regional pronunciations of words—deemphasizes all those not represented. While innumerable Navajo reference materials have perpetuated exactly such preferential status for the de facto standard dialect (and avoiding such a bias is arguably both impracticable and even undesirable), the largely oral methodology of the Rosetta Stone program may make this bias more apparent to non-specialists. Overall, though, the community reception to the Rosetta Stone materials has been extremely positive. The long-term impact and efficacy of the materials as a language revitalization tool remains to be seen.

\textbf{2.1.5. Concluding Remarks on the Field of Navajo Linguistics}

Since the early years of linguistic research into the Navajo language, the field has shifted from focusing on theoretical aspects of the language to addressing community concerns about teaching and supporting Navajo as a living community language. While research into theoretical aspects of Navajo linguistics does continue, the greatest community support falls behind research concerning the maintenance and revitalization of the language.

\textbf{2.2. Navajo Morphosyntax}

To ensure that the discussion of Navajo/English language contact phenomena are as meaningful as possible, a few pages are first dedicated to outlining some basic facts about Navajo morphosyntax. As anyone reading this work presumably has a functional understanding

\textsuperscript{23} Although Navajo has a relatively long written tradition and a substantial body of language support materials such as dictionaries and even published children’s books, most fluent speakers are not literate in Navajo. Unsurprisingly for a primarily oral language whose speech community is spread across more than 27,000 square miles, there is a good deal of regional variation, and speakers from geographically distant communities often have trouble understanding one another.
of English morphosyntax—and is therefore not in need of a general introduction to the subject—a comparable overview of English is omitted, but Navajo grammar is generally unfamiliar to non-specialists and a brief overview seems appropriate.

Navajo is a morphologically complex language with a rich system of verbal and nominal affixes. The status of overt nominals in Navajo is somewhat controversial. Navajo has often been argued (e.g. Willie, 1991; Hale, 2003) to be a pronominal argument language as described by Jelinek (1984), where verbal arguments are expressed as pronominal affixes and lexical NPs or DPs optionally appear as adjuncts. Other researchers (e.g. Platero, 1982; Speas, 1990; Rice, 2001; Rice & Saxon, 1991; Saxon & Rice, 2005) argue that Navajo nominals in fact do serve as verbal arguments, and that the optionality of nominals is comparable to the pro-drop patterns seen in languages such as Spanish or Japanese; under this view the verbal morphemes are taken to be agreement markers. Examples (17,18) illustrate the optionality of lexical nominals serving both as subjects and objects.

(17) a. Neinišché.  
    3.3.chase  
    ‘s/he/it is chasing him/her/it’

    b. Nāshdói bižh neinišché.  
    wildcat deer 3.3.chase  
    ‘The wildcat is chasing the deer.’

(18) a. Bits’áníjaa’.  
    3.3.1sg.take-away-from:PERF  
    ‘I took him/her/it away from him/her/it.’

When overt nominals do appear, the most common word order in Navajo is SOV, as shown in (19a). However, Navajo word order is somewhat flexible, and as shown in (19b) speakers may rearrange the order of lexical nouns.

(19) a. Ashkii at’ééd yiztsós
   boy girl 3.3.kiss
   ‘The boy kissed the girl.’

b. Ashkii at’ééd biztsós
   boy girl 3.3.kiss
   ‘That boy, the girl kissed him.’

Navajo exhibits several other properties pointed out by Dryer (e.g. 2007) to be characteristic of head-final languages, such as the appearance of manner adverbs in pre-verbal position (20), the appearance of genitive noun phrases before the nominals being modified (21), and the use of postpositions.

(20) díí hastiin ‘ayóó ‘ahááh yáltí’
    this man very fast 3.talk:IMPF
    ‘This man talks very fast’

(21) ashkii bi-zhé’é
    boy 3-father
    ‘boy’s father’

(22) tsin bi-yaa
    tree 3-under
    ‘under the/a tree’

When both the subject and object are third person, deviation from the unmarked SOV word order triggers the so-called yi-bi alternation, in which the verbal prefix yi- is replaced with the prefix bi-. Athabaskan literature generally follows Klaiman (1992) in terming the yi-bi alternation the direct/inverse voice alternation (see Willie, 2000b or Aissen, 2000, for more in-depth discussions of the Navajo yi-bi alternation) or the proximate/obviative alternation. As pointed out by a number of authors (e.g. Speas, 1990; Jelinek, 1997), this so-called voice alternation does not correspond to a passive interpretation and has no effect on verbal valence, and is generally employed for either pragmatic reasons such as topic prominence (as shown in (17b)), or out of compliance with the Navajo animacy hierarchy which requires lexical nouns to appear in the order humans > human infants/larger animals/animals with ‘superior intelligence’ > medium-sized animals > smaller animals > insects > natural forces > inanimate objects > abstractions. (Creamer, 1974; Young & Morgan, 1987:65). Navajo does have a true grammatical passive, as in ch’íshi’deedzí ‘I was dragged out’ (Young, 2000:64), which exhibits a canonical passive interpretation.
A rich and extensive system of enclitics are used to express (among other things) directionality, location, temporality, nominalization, and clause type, as shown in (23a–g). Enclitics may be appended to nominals (23a–c) as well as verbs (as in (23d,e)) and postpositions (as in (23f,g)).

(23) a. hooghan-di
    house-at
    'at home/at the house'

b. hooghan-góó
    house-towards
    'towards home/the house'

c. Chiyáán-ísh íínílaa?
    Food-INT 3.2sg.make:PERF
    'Did you make the food?'

d. yidláa-go
    3.3.drink-while
    'while he’s drinking it'

e. ashkii alhosh-ígií
    boy 3.sleep-REL
    'the boy who is sleeping'

f. tsé bi-yaa-déé’
    rock 3.under-from
    'from under a/the rock'

g. Tsédáa-jí’ doodago dził bi-gháá’-jí’-ísh nííí’na’? (Fernald & Perkins, 2006:31)
    cliff-to or mountain 3- on- to-INT 1sg-crawl:PERF
    ‘Did you crawl to the edge of the cliff or to the top of the mountain?’

As shown in (24), English-like adjectives do not exist in Navajo, and stative verbs are instead used for comparable nominal modification.

(24) ŋééchąą́’ lítso’-ígií tó yidlá
    dog 3.yellow-REL water 3.3.drink
    ‘a/the yellow dog is drinking water’
Navajo has no grammatical gender, and the great majority of nouns do not show number marking. Navajo verbs undergo far richer inflectional processes than do nouns, and serve a far wider range of semantic functions. While most verbs indicate the number of the subject through inflectional morphology, a subset of verbs (primarily of motion/location) indicate the number of the subject through stem changes, as shown in (25). As shown in (26), verbs mark both singular, dual, and plural number for both the subject and the object.

(25)  
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nádááh</td>
<td>b. nát’ásh</td>
<td>c. nákah</td>
</tr>
<tr>
<td>‘he/she/it arrives often’</td>
<td>‘the two of them arrive often’</td>
<td>‘they (3 or more) arrive often’</td>
</tr>
</tbody>
</table>

(26)  
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. nish-ch’id</td>
<td>e. shoh-ch’id</td>
<td></td>
</tr>
<tr>
<td>2sg.1sg.scratch:IMPF</td>
<td>1sg.2dpl.scratch:IMPF</td>
<td></td>
</tr>
<tr>
<td>‘I’m scratching you’</td>
<td>‘You two are scratching me’</td>
<td></td>
</tr>
<tr>
<td>b. yish-ch’id</td>
<td>f. woh-ch’id</td>
<td></td>
</tr>
<tr>
<td>3.1sg.scratch:IMPF</td>
<td>3.2dpl.scratch:IMPF</td>
<td></td>
</tr>
<tr>
<td>‘I’m scratching him/her/it’</td>
<td>‘You two are scratching him/her/it’</td>
<td></td>
</tr>
</tbody>
</table>

---

Nominal number marking is not optional in Navajo; rather it is obligatory for the small subset of nominals—terms for kinship and human age/gender groups—which have overt number-marked forms, and is unavailable for all other nominals. The vast majority of Navajo nominals do not permit overt dual/plural marking, and for these nominals number is expressed solely through verbal morphology. This pattern is illustrated in example (i) with a Navajo noun that has both a singular and a plural form, and in (ii) with a Navajo noun that does not have a number-marked form.

i.    
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. hastiín gohwééh yidlá</td>
<td>ii. a. bá’óltá’í gohwééh yidlá</td>
<td></td>
</tr>
<tr>
<td>man</td>
<td>coffee</td>
<td>3.3.drink</td>
</tr>
<tr>
<td>“the/a man is drinking coffee”</td>
<td>“the/a teacher is drinking coffee”</td>
<td></td>
</tr>
</tbody>
</table>

b. hastóí/*hastiín gohwééh deidlá | b. bá’óltá’í gohwééh deidlá |
| men/man | coffee | PL.3.3.drink |
| “(the) men are drinking coffee” | “(the) teachers are drinking coffee” |

There is no “regular” plural nominal marker which corresponds semantically to the English nominal plural, nor is there a default way to pluralize a nominal that does not already have an established plural-marked form. It is worth noting that Navajo does have a distributive plural morpheme which may be appended to nominals, but the meaning differs from the regular English nominal plural. As discussed in Young & Morgan (1987: 7-8), the prefix da- may be appended to certain nominals such as k’os ‘cloud’, producing forms such as daak’os meaning ‘many individual clouds’, but not to others such as bee’atsidi ‘hammer’ (*beeda’atsidi ‘hammers’).

For most verbs, the subject markings for third person singular and dual are homophous.
Navajo verbs perform many of the communicative functions played by nouns or adjectives in languages such as English. As shown in (27), while a small subset of Navajo nominals—mostly words indicating kinship terms, body parts, wild animals, and what Young and Morgan (1987:357) term “universals” (words such as tó ‘water’, sq ‘star’, k’os ‘cloud’)—are monomorphemic canonical nouns, the great majority of Navajo nominals are derived from either nominal stems compounded with additional nominal stems, postpositions, enclitics, or verbs (as shown in 28, 29), or are deverbal nouns, as shown in (30).

(27)  
a. tl’oh chin  
‘grass, hay’ ‘smell, have an odor’  
=>  
 tl’ohchin  
‘onion’

b. jeeh lid  
‘pitch, resin’ ‘smoke’  
=>  
 jeelid  
‘soot’

(28)  
łéé chā’ í  
pet/domesticated animal defecate NOM  
=>  
łééchā’í  
‘dog’
(literally, ‘excrement pet’, presumably from the canine habit of eating excrement)

(29)  
bił níjoobalí  
3sg.with 3sg.4sg.spin-NOM  
‘cell phone’ (literally, ‘one spins around with it’)

(30)  
adinídíín  
‘light’ (N) (literally, ‘there is light’)

The core of the Navajo language is unarguably the verb, which has been the focus of perhaps the greatest amount of Navajo linguistic research. The internal structure of the Navajo
verb is generally described using a template with around ten positions (the exact number varies according to the analysis) for verbal prefixes. Example (31), modified from Young & Morgan (1987: 37-38), shows a basic ten-position representation where a morphophonological distinction is made between the so-called disjunct and conjunct morphemes.  

(31) Navajo Verbal Template

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>a</td>
<td>b</td>
<td>c</td>
<td>d</td>
<td>e</td>
<td>a</td>
<td>b</td>
<td>c</td>
<td></td>
</tr>
</tbody>
</table>

**DISJUNCT PREFIXES**

<table>
<thead>
<tr>
<th></th>
<th>Object of postposition</th>
<th>Adverbial-Thematic</th>
<th>Reflexive</th>
<th>Reversivory</th>
<th>Semiliterative</th>
<th>Iterative</th>
<th>Distributive-Plural</th>
<th>Direct Object Pronouns</th>
<th>Subject pronouns</th>
<th>Iterative</th>
<th>Translational/Semelfactive Aspect Markers</th>
<th>Modal-Aspectival Conjugation markers</th>
<th>Subject Pronouns</th>
<th>Subject Pronouns</th>
<th>Classifier (transitivity, voice)</th>
<th>Enclitics (negative, relativizer, usually, etc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Null postposition</td>
<td>Adverbial-Thematic</td>
<td>Reflexive</td>
<td>Reversivory</td>
<td>Semiliterative</td>
<td>Iterative</td>
<td>Distributive-Plural</td>
<td>Direct Object Pronouns</td>
<td>Subject pronouns</td>
<td>Iterative</td>
<td>Translational/Semelfactive Aspect Markers</td>
<td>Modal-Aspectival Conjugation markers</td>
<td>Subject Pronouns</td>
<td>Subject Pronouns</td>
<td>Classifier (transitivity, voice)</td>
<td>Enclitics (negative, relativizer, usually, etc)</td>
</tr>
</tbody>
</table>

The template provides for the formation of exceptionally complex verbs, as shown in example (32), although generally the slots are not all filled at the same time, as shown in example (33a). Slots 1 and 6—which are filled by adverbial-type morphemes—may be multiply filled, as shown in (33b).

---

28 Disjunct morphemes are considered to be “less tightly bound as components of the verb prefix complex than the [c]onjunct elements” (Young & Morgan, 1987: 39). They are clitic-like in their morphosyntactic behavior and generally speaking are more resistant to morphophonological modification during the process of incorporation. While conjunct morphemes introduce verbal arguments, modify verb valency, or provide information about mood, disjunct morphemes add adverbial information, introduce oblique arguments to the verbal structure, or provide information about iterativity or distributive properties of the verb.

29 The term *fourth person* is used in the Athabaskan literature to refer to a set of subject and object morphemes which, among other uses, are used to express obviative/antiobviative reference, disjoint reference, and certain types of deixis. See Willie (1991) or Aissen (2000) for a more detailed survey of the discourse functions and morphosyntactic properties of the Navajo fourth person.

30 Athabaskan literature traditionally uses the term “classifiers”; these morphemes are effectively transitivizers, and have no relationship to the morphemes used to indicate conceptual classification of nouns in many East Asian languages. Despite the potential for confusion with the more widely-known use of the term “classifier”, this current project retains the traditional Athabaskan term in an effort to maintain continuity with the Athabaskan literature.
(32)

{
0  1  2  3  4  5  6  7 / 8  9  10  Enc.
}

yi    sh    d    na’
PERF  1.sg  CL    Crawl:PERF

to-safety  IT  PL  1.sg  4.sg  INCHO  FUT / 3.sg  CL  move:FUT

“They (indefinite) will repeatedly get me back to safety.”
(modified from Hale, 2003:14)

(33) a.  0  1  2  3  4  5  6  7  8  9  10  Enc.

yi    sh    d    na’
PERF  1.sg  CL    Crawl:PERF

“1 crawled to a point”

b.  0  1  2  3  4  5  6  7  8  9  10  Enc.

shí  áldo’  da  ‘í  Ø  l  tsóód  go
shá  1.sg  also  PL  3.sg  3.pl  CL  feed:IMPF  while

“While they were feeding it to me too”
(modified from Courtney & Saville-Troike, 2002:624)

It is important to clarify that this template is intended to represent descriptive
generalizations about the typical linear order of morphemes, and does not make theoretical
claims about the status of any particular Navajo morphemes or their relationships to one another.

McDonough (2000b) outlines several reasons to avoid treating the template as a model of the
Navajo verbal system, the first of which being that the template was designed by Li (1946) as a
tool for diachronic and synchronic comparison between Athabaskan languages, and was never
intended to provide a model of speaker competence. Even discounting the origins of the
template, McDonough points out that in general, position class morphologies—morphological
systems based on templates—exist only “because of ordering considerations among morphemes:
they arise crucially when concatenation cannot be accounted for either by prosodic morphology
or affixation to stem without the prosthesis of the template” (2000b: 140). In the case of Navajo,
“the template requires that elaborate theoretical and procedural adjustments be invented to
handle it, mainly because the linear order of positions cannot be used as a concatenation order and because the concatenation of template morphemes often requires extensive rewrite rules” McDonough (2003:18). Furthermore, certain positions (such as 1 or 6) are filled by a collection of morphemes having no real semantic or functional properties in common and being assigned to the same position only by virtue of habitually appearing between other sets of more clearly defined morphemes. Additional concerns voiced by McDonough include the fact that earlier researchers (e.g. Morice and Sapir) also believed that a simpler structure was underlying, and, perhaps most importantly, the fact that native Navajo speakers “find the template counterintuitive and opaque” (McDonough, 2000:141).

An alternative system proposed to align more closely with the underlying organizational system of native speakers’ grammars is outlined in a series of papers by McDonough (1990, 1996, 2000a, 2000b, 2003). Termed the ‘bipartite’ or the ‘bipartite constituent’ model, McDonough’s model is based on the non-controversial morphophonemic divisions of the Navajo verb into a verb stem, a conjunct domain, and a disjunct domain. As shown in (34), the bipartite model relabels these domains, recasting the verb stem as the canonical verb, the conjunct domain as a phonologically synthesized auxiliary verb, and the disjunct domain as a morphosyntactically optional proclitic complex.

(34)  Disjunct   #   [ Conjunct ]   [ Verb Stem ]   (McDonough 2003:23)
      D       #   [ Aux ]   [ Verb ]

Beyond simply relabeling the domains, the bipartite model uses phonological evidence to

---

31 In fact McDonough (e.g. 2003) proposes that the bipartite model is applicable for all Athabaskan verbs, not just those in Navajo. Many features of verbal morphology are shared by languages across the Athabaskan language family, and cross-family proposals for verbal templates and other verbal morphological models are common (e.g. Kari, 1989, 1992; Rice, 2000).
support a reorganization of morphemes proposed to underlie the Navajo verbal complex. The model proposes that, rather than being composed of a string of concatenated morphemes, a Navajo verb is composed of a verb and a bound auxiliary. This verb is comprised of a verbal root and, where applicable, a ‘classifier’ (voice/valence marker); the auxiliary is comprised of an inflectional stem (a synthetic bundle of tense, mode, and subject agreement), one or more ‘qualifiers’ (aspectual or adverbial-like morphemes) when relevant, and any agreement markers for direct objects or deictic subjects. Example (35), modified from (McDonough, 2003: 19), shows this proposed reanalysis of the Navajo verbal complex.

(35) proclitics # [(AGR) [(qual) Aux]_{AUX} [( CL) Verb]_{VERB}]_{WORD}

Example (36), modified from McDonough (2003: 38), shows the correspondence of the proposed bipartite model to the traditional template positions, where the underlined positions correspond to the proposed auxiliary and verb stems.

(36) 0-III # [ IV/V-VI(abc) – VII/VIII ] [ IX Verb stem ] <= template positions
D   # [ Aux ] [ Verb ] <= bipartite model

Example (37) shows how this model maps the verb *yisdánidashizhdohtéél*, which was shown in the template model in example (32) above.

(37) yisdá-ní-da # [[[ shi-zh-d-o0 ]_{AUX} [ l-téél }_{VERB}]_{WORD}
to safety-repeatedly-PL # [1sgOBJ.3sgSUB.INCHO.FUT] [move-it:FUT]
‘They/someone will repeatedly return me back to safety’

As shown in (37), under the bipartite model the verb is treated as being composed of a verb (-htéél), an inflected auxiliary (shizhdo-), and a string of proclitics (yisdá-ní-da-).

This bipartite model is in fact not wildly radical, McDonough (2003) points out, and is in many ways a reframing and revision of the so-called Base Paradigms identified by Young and
Morgan (1980, 1987, etc). Young and Morgan (1987:200) write that “[e]ssentially, the [Navajo] verb is conjugated in 16 basic paradigms . . .” and continue by identifying 16 ‘base’ and ‘extended base’ paradigms, each comprised of the morphemes corresponding to either the template positions 7–9 (the ‘base’ paradigms) or 3–9 (the ‘extended base’ paradigms). The divisions within the verbal complex proposed for these base paradigms differ from those identified by the bipartite model, but both approaches operate from the underlying intuition that Navajo is an inflectional language.

Adopting the bipartite model over the traditional position class model allows the Navajo verb to be analyzed as participating in inflectional alternations, rather than simply presenting itself with a string of morphophonologically opaque prefixes. Not only does the bipartite model arguably offer greater theoretical elegance, its proponents argue that it has greater correspondence to a native speaker’s understanding of the grammar. While no rigorous data is available to support the bipartite model’s claims on this latter point, informal observation suggests that some native speakers do find the bipartite model to be more intuitive than the position class template, and some preliminary research (McDonough & Sussman, 2006) has found adult native speakers to access unanalyzed verbal units corresponding to the proposed bipartite auxiliary\(^{32}\).

\(^{32}\)Limited first language acquisition data (Saville-Troike, 1996; Courtney & Saville-Troike, 2002) do not appear to provide clear support either for or against the bipartite model. Although there is evidence that child speakers do produce verbs that consist of only a verb stem and a subject marker—which, under the bipartite model, should not occur, as the subject marker is argued to be bundled with tense/mode to form the auxiliary stem—two facts (that the subject markers are homophonous with the corresponding unbound emphatic pronouns, and that the youngest child speakers do sometimes produce bare verb stems), render the theoretical impact of these data somewhat uncertain. Other child verb forms which are glossed as containing bare tense/mode markers without the accompanying subject marker are also difficult to interpret, as many of these forms are homophonous with the adult-like forms of non-context-appropriate or non-stem-appropriate person/tense/mode bundles, raising the question of whether the child’s error involves having chosen a semantically/pragmatically incorrect verb form, or having selected the incorrect
2.3. Navajo & English: Languages in Contact

After the brief orientation to Navajo morphosyntax provided in section 2.2., this current section now turns to a review of Navajo/English language contact phenomena. Navajo speakers have been interacting with the English language for several generations, and a reasonable body of literature exists on observed contact effects. This section first reviews the literature on Navajo-English code-mixing and then turns to the non-standard English used by many Navajo tribal members. Section 2.3.1 reviews the literature on code-mixing, while section 2.3.2 discusses the language contact effects often called “Navajo English” or “Navglish”.

2.3.1. Code-mixing

As discussed in chapter 1, code-mixing occurs in speech communities with high levels of functional bilingualism. Researchers working in such communities on and near the Navajo reservation have reported phenomena falling into several of the categories outlined in Muysken (2000), such as the alternations shown below. In example (38), a Navajo clause is followed by an English clause:

auxiliary stem (rather than having produced an orphaned tense/mode marker.) This last case is illustrated in example (i), which shows a verb produced by a child aged 2 years, 10 months. (The bolded text corresponds to the bipartite model’s proposed auxiliary stem.)

(i)  
i- ni- l- gesh (child form)  
k’i- ni- sh- gizh (adult-like form)
off PERF 1sg trans. cut it

‘I cut it off’

Ignoring the adverbial k’i (or i- in the child form), in example (i) the appearance of the perfective marker ni without the subject marking sh might be taken as evidence that instead of targeting the bipartite model’s proposed first-person singular perfective auxiliary stem ni\(\text{sh-}\), this child’s verb formation process involves selecting individual morphemes from the template, and that the child simply made several selection errors. However, Navajo has eight possible first-person singular perfective stems, one of which is ni\(\text{l-}\), which corresponds to the form ni\(\text{l-gesh}\) produced by the child. While the stem ni\(\text{l-}\) does not co-occur with the verb stem –gizh, an argument could be made that in this case the child might have simply selected the wrong auxiliary stem.
Alternations also occur in less canonical places, as shown in the following slogan used during a 2007 public health campaign on the Navajo Nation. While the sentence in example (39) begins with the Navajo exclamation and possessive, the lexical noun is produced in English, and the sentence continues in English.

(39) Shóó shí heart, buckle up! (Webster, 2008: 532)

hey 1sg.POSS heart, buckle up
‘Hey, my heart [my love], buckle up!’

In monolingual Navajo, possessive $shí$- would be a proclitic (as opposed to, say, the emphatic pronoun $shí$, which is a free morpheme) and would be attached to the lexical noun. The authors’ orthographic choice to represent it as an unbound morpheme is suggestive of an underlying intuition ruling out word-internal code-mixing, although no definitive research exists on this point.

Data falling into Muysken’s category of insertions represent a more varied group of code-mixing phenomena. Example (40) shows the insertion of a Navajo adpositional phrase into an English sentence.

(40) The priest wore a breastplate, bikáagi éiyá precious stones
The priest wore a breastplate, 3.on-it-at dem precious stones
‘The priest wore a breastplate with precious stones on it . . .’ (Schaengold, 2004:71)

Single-word insertions are also extremely common, as shown in (41) below.

(41) a. ei moon yikáá’ naakai
    dem moon 3.on it 3pl.walk:IMPF
    ‘They’re walking on the moon’ (Holm et al., 1971:81)
b. Town-góó déyá
town-towards 1sg.go:IMPF
‘I am going to town’ (Schaengold, 2004: 47)

c. ni-paper
2.POSS paper
‘your paper’ (Saville-Troike, 1996:161)

While little phonological information is provided for most of these items, and so nonce borrowing cannot be definitively ruled out, it is worth noting that such single-word insertions occur sometimes in the same conversation as their Navajo counterparts, as shown below in (42a,b), where the four-year old speaker uses both the English word *horse* and its Navajo counterpart *líí’* during the same recorded session. The proximity of the corresponding lexical items suggests that, at a minimum, the speaker’s use of *hersi* should not be attributed to a gap in the Navajo lexicon.

(42) a. Hersi #ka’ sida (Saville-Troike 1996: 158)
  horsey on 3.sit:IMPF
  “It is sitting on the horse”

b. líí’ nanazi (Saville-Troike 1996: 161)
  horse ?.sit:IMPF
  ‘It is standing by the horse’

In other cases, the corresponding English and Navajo lexical items are used in the same sentence, as in (43), making a nonce borrowing interpretation more challenging.

(43) daadiílkaal door-ísh close ada’ílaah (Schaengold, 2004: 61)
  door door-INT close 3rd.pl.make:PERF
  ‘The door, did they close the door?’

Without phonological evidence it is difficult to reach a conclusive verdict on whether single-word insertions such as those above represent borrowing or code-mixing (and it is probable that
ultimately each instance should be treated individually), but it seems unlikely that all cases of single-word insertions should be attributed to nonce borrowing.

Finally, perhaps the most intriguing set of data comes from the instances of verbal code-mixing. The only productive strategy for verbal code-mixing is to use forms of the Navajo verbs áshlééh ‘make/do’ and ásht’į‘act/do’ with English lexical verbs to form light verb constructions, as shown above in (43) and below in (44, 45).

(44) Bookshelf la’ shá save ání-lééh. (Schaengold, 2004: 53)
    bookshelf one 1sg.BEN save 2sg.make:IMPF
    ‘You’re saving me (one) bookshelf’

(45) Swimming ásht’į (Canfield, 1980: 219)
    swimming 1sg.act:IMPF
    ‘I’m swimming’

Interestingly, light verb constructions such as those shown in (43) and (44) exhibit what has been termed feature spreading, where “the [English] VP selected by the [Navajo] light verb has features typical of [Navajo] with regards to . . . word order, prosody and information structure. (González-Vilbazo & López, 2011: 833). In other words, the observed word order within the English VP is not the expected monolingual English VO order, but the monolingual Navajo OV order. As noted above in section 2.2.2, González-Vilbazo & López (2012) claim that such word order differences suggest that the grammatical features of \( v^0 \)—rather than any features associated with the VP itself—determine word order in a complement VP.

---

33 Example (45) is not a clear example of a light verb construction, as it is not clear whether swimming should be treated as a nominalization or as a verbal element. If swimming is considered to be a nominalization—as would be common in a light verb construction—the utterance follows the same OV word order as seen in (43 and 44). If swimming is considered to be verbal, though, the utterance in example (45) is clearly not illustrative of verb-object order, but is still indicative of Navajo-like word order, as a comparable English construction would employ an auxiliary-lexical verb word order.
Canfield (1980) observes that forms of áshlééh are reserved for code-mixing involving transitive verbs, and forms of ásht'į for intransitive verbs, although data such as (46)—which by most interpretations would be intransitive, yet uses the third-person singular form of áshlééh (not ásht'į)—suggest that this generalization may require some refinement.

(46) jump iyiliaa
    jump 3sg.make:PERF
  ‘he/she/it jumped’

Holm et al. (1971) report that, in monolingual Navajo, light verb constructions such as those above are not common. In their words, “[t]he more usual way of saying this in Navajo would be dah neeshjiid, ‘he jumped’ . . . there is an awkward but acceptable circumlocution: dah ná’nljiįį iyiliiaa . . . where the content verb is nominalized.” (Holm et al., 1971:7) The cross-linguistic adaptation of this ‘awkward’ and rare monolingual Navajo circumlocution, however, is extremely common in mixed Navajo-English code, and its high frequency make an ‘awkward’ attribution unlikely.

Cross-linguistically, there is an extremely strong precedent for verbs that are borrowed or code-mixed to trigger the formation of a light verb construction. González-Vilbazo & López (2011, 2012) give evidence from Spanish/German code-mixing that such light verb constructions may be a part of the mixed code even when light verb constructions do not exist in either of the participating monolingual grammars.

Furthermore, mirroring observations for numerous other language pairs (Spanish/German: González-Vilbazo & López, 2011, 2012; Telegu/English: Den Dikken & Rao, 2003; Hindi/English: Ritchie & Bhatia, 1996; Marathi/English: Joshi, 1985; Turkish/Dutch: Boeschoten & Verhoeven, 1985), this pattern appears to be asymmetrical in that parallel
constructions—employing an English light verb and a Navajo lexical verb or VP—are unattested in the already published data on Navajo/English code-mixing\textsuperscript{34}. More detailed code-mixing data would be needed to determine whether such a gap is systematic or accidental; if systematic, the Navajo/English data would support the models outlined in González-Vilbazo & López (2011, 2012).

While such cross-linguistic light verb constructions are an extremely common tool used by bilingual speakers to code-mix within the verbal compound, the use of verbal morphology from one language with a verbal stem from the other is rarer\textsuperscript{35}. Interestingly, Canfield (1980) presents several examples of English lexical verb stems bearing Navajo inflectional morphology.

\begin{enumerate}
\item [47] a. na’ish-\textit{crash} lá
   \begin{tabular}{l}
   1.FUT:\textit{crash} EMPH  \\
   “I’m about to pass out”
\end{tabular}
\item b. shił naweasy
   \begin{tabular}{l}
   1.with 3:\textit{queasy}  \\
   “I feel sick” (literally, ‘it is queasy with me’)
\end{tabular}
\end{enumerate}

\textsuperscript{34} González-Vilbazo & López (2011a) suggest that these asymmetries may be traced to different featural specifications of verbal roots and $v^0$s in each of the participating languages; if a given feature (such as conjugation class) is specified for verbal roots and $v^0$s in one language but not the other, the language with more highly specified verbal roots and $v^0$s is predicted to contribute light verbs but never lexical verbs, while the less specified verbal roots are predicted to function as lexical verbs in light verb constructions. Working from the idea that a $v^0$ specified for features unmatched by those of a verbal root would result in feature mismatch, the authors propose that light verb constructions may only be formed from a lexical verb which has the same features or fewer features than the $v^0$. When light verb constructions are formed through code-mixing, the prediction is that a language asymmetry should generally be observed unless there is alignment between the two languages’ verbal feature specifications.

\textsuperscript{35} This is certainly not to suggest that cross-linguistic verbal constructions are non-existent; for example, in a language attrition case study following an L1 Hebrew-speaking child acquiring English, Kaufman & Aronoff (1991) report multiple instances of Hebrew verb stems being used with English morphology, such as in their example “I’m \textit{me-nagev-ing} myself” (Kaufman & Aronoff, 1991: 182), meaning “I’m drying myself”, where the Hebrew present tense marker \textit{me} and the Hebrew verb stem \textit{nagev} are used in conjunction with the English -\textit{ing} morpheme. However, such single-word cross-linguistic verbal examples are widely believed (e.g. González-Vilbazo & López (2011) to be rarer than the relatively common cross-linguistic light verb construction.
c. shilééchaa’í anáyiítum
   1.POSS.dog  3.turn
   “my dog died” (a play on the Navajo idiom anáyiildéél, literally ‘threw/tossed away’)

d. Háágóó ŋsini-go
   where-to  2.PERF:go
   “Where did you go?”

(48) “H’áát’íship baa  nan-doing?
    What  3.on  IT-doing
   “What are you doing?” (monolingual Navajo = “H’áát’íship baa naniná)

However, Canfield takes pains to clarify that constructions of this nature are in fact rather
marked, and that the juxtaposition of an English verb stem with Navajo inflectional morphology
“tend[s] to be more idiomatic and playful, often getting a laugh” (Canfield 1980: 220). The
success of such humor is arguably thanks to the use of a morphosyntactically marked
construction, similar to the infamous idiosyncratically strong verbs of early 1900’s baseball
announcer Dizzy Dean; many linguists (e.g. Pinker, 1994; Lederer, 1996) have discussed how
the consciously non-standard verbal forms in utterances such as “the pitcher flang the ball” or
“he slud into third base” act as a playful mark of unpretentiousness.

    The playfulness of these examples of English verb stems with Navajo inflection is
bolstered by a comparison with the findings of Saville-Troike (1996). Working over a two year
period with five children ranging between 13 and 55 months of age, Saville-Troike found that
“although there was considerable borrowing from English . . . in no case was an element of
English incorporated in a Navajo inflected verb.” (Saville-Troike 1996: 161) While it is certainly
possible that the discrepancy between the findings of Canfield and those of Saville-Troike is
coincidental, it seems far more likely that differences between the speaker populations
participating in these studies plays a role.
Canfield’s observation—that insertion of English verb stems into the inflectional complex is a playful act, apparently undertaken “just for the fun of it” (Canfield, 1980: 220)—points towards such constructions being pragmatically regulated. Pragmatic constraints are notoriously late to be acquired by young speakers, and there is some evidence that children may not acquire an adult-like understanding of pragmatic rules until their early teenage years: one study working with speakers of Standard American English found that “. . . children do not tend to enjoy teasing or use teasing in a playful way until 11 to 12 years of age.” (Keltner et al., 2001) Although researchers such as Heath (1983) report teasing-like behavior at much younger ages, teasing language may be produced and interpreted differently by child and adult speakers. If the use of Navajo-inflected English verb stems reflects a specific playful use of language, it is quite plausible that immature pragmatic fluency is responsible for the failure of Saville-Troike’s young bilingual speakers to produce any such constructions.

2.3.2. Non-standard English in the Navajo Community

As discussed in chapter 1, extended contact between two languages can produce effects perhaps less dramatic than code-mixing—which is starkly apparent to monolingual listeners—which are often dismissed as being simply improper or imperfect forms of one language. Contact effects range from phonological changes, lexical borrowing, and in some cases, morphosyntactic changes based on structural analogy to a construction in the source language

---

36 An example of this type of structural analogy might be found in possessive constructions in Modern Hebrew. Zuckermann (2009) argues that possessive constructions in Modern Hebrew show the influence of language contact with Yiddish and other Indo-European languages. In classical Hebrew, possession is indicated by means of an existential clause, as shown in example (i), where the possessed object bears no case marking and serves as the grammatical subject. Modern Hebrew, however, employs the construction shown in (ii), where the possessed object is marked by accusative case.
Perhaps the most comprehensive treatment of language contact effects in the Navajo community is currently provided in Charlotte Schaengold’s 2004 dissertation. The dissertation makes a distinction between the so-called “Bilingual Navajo”—described as being a “non-standard variety of Navajo”—and the non-standard English used by Navajo tribal members, clarifying that “bilingual Navajo is . . . the result of code switching phenomena among fluent bilinguals” (Schaengold 2004:4), while the non-standard English is described as having “some non-standard features . . . shared with other non-standard American English dialects, [but] particularly identifiable characteristics [which] can be explained best as substrate interference from Navajo.” (Schaengold, 2004:31).

This subsection focuses on the non-standard variety of English spoken on the Navajo reservation. Spoken by both Navajo speakers and non-Navajo speakers, the dialect is characterized by its almost exclusive use of English lexical items. A small number of common Navajo-origin words are in use, such as chidi (car), chizhi (dry, ashy), or nalis (paternal grandparents), but the bulk of the lexical items are English in origin.

Leaving aside the phonological differences between the Navajo English (NE) and Standard American English (SAE), many differences between the two dialects are arguably attributable to the influence of Navajo morphosyntax. Examples include these in (49), where the

(i) yésh l-i ha-séfer ha-zè  
EXIS DAT-1sg DEF-book DEF-msg PROX
“I have this book”

(ii) yésh l-i et ha-séfer ha-zè  
EXIS DAT-1sg ACC DEF-book DEF-msg PROX
“I have this book”

(Zuckermann, 2009:51)

Zuckermann argues that the use of the accusative case marker on the possessed object stems from analogy to Yiddish, where the possessive construction uses a form of the verb hobn, meaning “to have” and the possessed object is marked with accusative case.
redundant plural marking is suggested by Schaengold to be an instance of hypercorrection (as Navajo has no productive nominal number marking).

(49) The womens are over on that side. (Schaengold, 2004: 34)

In example (50a), the possessive construction closely echoes the standard Navajo possessive printed in (50b).

(50) a. my aunt her house (Schaengold, 2004: 36)

b. shi-máázhí bi-ghan
   1.POSS-maternal aunt 3.POSS-house
   “my aunt’s house”

Finally, the examples in (51) suggest the influence of the Navajo aspectual system. As discussed above, Navajo does not mark verbal tenses, and instead employs temporal adverbs to indicate that an event took place in the past. Examples such as those printed in (51) are suggested by Schaengold to show analogous constructions, where the standard tense marking is absent from the verb\(^{37}\), and past time is instead indicated solely by the adverbials *yet* and *yesterday*.

(51) a. We haven’t cook it yet.

b. Yesterday we meet him in town. (Schaengold, 2004: 35)

Such examples suggest the lasting effect of Navajo grammar on the non-standard English used on the reservation. The non-standard dialect is used today by many speakers who have little or no command of Navajo; while many differences between Navajo English and standard English

\(^{37}\) In (51a), *cook* appears to be a past participle, and the (absent) –ed morpheme would presumably mark perfect aspect rather than past tense as suggested by Schaengold. It is not clear that the absence of perfect aspect marking should be attributed to Navajo influence, as perfect aspect is marked in Navajo. As discussed further in chapters 4 and 5, phonological effects may confound the question of overtly expressed –ed morphemes in Navajo English.
may be originally attributable to imperfect acquisition by second-language learners, many speakers today almost certainly are unaware of the origins of these constructions.

2.4. Navajo Language & Linguistics: Conclusion

Over the past century, Navajo linguistics has shifted from a field primarily focused on language documentation to a field strongly influenced by questions of language maintenance and revitalization. This chapter briefly reviews the history of Navajo linguistic research, provides an overview of the social and cultural context of Navajo/English bilingualism, and discusses some of the issues which are of interest to many members of the modern Navajo language community. The chapter concludes with an overview of Navajo morphosyntax and a discussion of the literature on Navajo-English code-mixing and language contact phenomena.
Chapter 3: Research Design & Methods

This chapter reviews the research plan, outlines the methods used for data collection and data analysis, and discusses some logistical and cultural considerations relevant to work involving Navajo communities and other indigenous communities.

As indicated in the introduction, this project was guided by the following research questions:

- **RQ1**: What are the morphosyntactic properties of children’s speech in this Navajo immersion classroom?
- **RQ2**: To what extent do the levels of morphosyntactic regularity support classification of the children’s dialect of English as a contact language?
- **RQ3**: What are the implications of the observed language patterns for bilingual families and educators?

This study used naturalistic data without any supplementary data from grammaticality judgment tests or other language intuition tasks. Naturalistic data is valuable in that it provides evidence of what can and does occur; however, sole reliance on naturalistic data in the case of contact language phenomena has been criticized by some (e.g. MacSwan, 1997) because it does not provide evidence of what cannot occur. Chomsky (e.g. 1965) has famously discussed the distinction between performance and competence; naturalistic data provides evidence of performance, but does not provide evidence of competence. However, research into potentially non-standard language varieties faces a challenge in the collection of grammaticality judgments; participants may respond according to the prescriptive norms of a standard language variety, and may not provide an accurate indication of the actual acceptability of a non-standard language.
feature in their speech community. As this study aimed to investigate the presence of non-standard English features and language contact effects, the research design prioritized the collection of naturalistic speech over grammaticality judgments of potentially questionable reliability.

However, before describing the research plan designed to address these questions, this chapter provides a brief discussion on designing and carrying out research in indigenous communities.

3.1. Research in Indigenous Communities: Community-Based Participatory Research

While the research carried out for this study was conducted in a community located off the Navajo reservation, many of these community members maintain cultural connections to the Navajo tribe, and some words on research relating to indigenous communities seem appropriate.

The research plan designed out for this dissertation was to some extent informed by an earlier research project which was cancelled during the participant recruitment phase. Targeting a rather different set of research questions (related to simultaneous Navajo/English bilingualism), the earlier protocol was found to be impracticable for logistical and cultural reasons, and was discarded prior to collecting data. Although not connected to the current study, some details of this earlier study’s failure may be instructive for researchers who hope to work with indigenous communities. The discussion in this section is based on research experience in the Navajo Nation, but much of it has broader applicability to research involving other indigenous communities.

In the past two decades more and more research carried out in indigenous communities has turned away from “pure” academic research protocols and instead turned towards practices
which engage community members as equal partners and collaborators in the research design and implementation. Such research protocols are often known as community-based participatory research plans (CBPRs), and are not only argued to be more effective and ethical than traditional academic research plans (e.g. Green et al., 2001; Minkler & Wallerstein, 2008; Harding et al., 2012), but are also increasingly mandated by tribal laws. Since the late 1960s Native American activists such as Vine Deloria have urged tribes to limit research activity conducted purely for the sake of research, and to instead focus on supporting only research activities that promote the interests of the tribe; Deloria’s 1969 book *Custer Died for Your Sins* was an early call for Native American tribes to assert control over research involving their cultural heritage and to involve tribal members in all stages of the research process. Starting in the late 1980s, the Navajo Nation has implemented a series of regulations designed to ensure the protection of tribal interests in research involving Navajo people, and now maintains careful control over all research activities on the reservation.

Beginning in 1996, the Navajo Nation has required that all research conducted on the reservation be conducted in accordance with the guidelines set out in the Navajo Nation Code 25:3251-3271 and be approved by the Navajo Nation Human Research Review Board (NNHRRB). The NNHRRB requires researchers to follow a community-oriented research plan, including raising community support for the research proposal, approaching local government officials to secure support, and (if the research relates to school-aged children, even if it does not take place in a school setting) securing support from the Navajo Nation Board of Education. Securing the approval of these governing bodies generally requires the delivery of presentations at scheduled public meetings and the preparation and dissemination of paperwork explaining the
research for the community. Furthermore, to ensure that research respects the cultural heritage of the Navajo people, researchers must apply for a permit from the Navajo Nation Historic Preservation Office. These steps, which can require a good deal of planning and may take months to accomplish, must all be completed before a research application can be presented to the NNHRRB. The board maintains an active involvement in research proceedings; once a protocol has been approved, the NNHRRB requires a research team to provide quarterly reports of any findings, with annual reports to be delivered in person. Upon completion of the project, all data and other materials collected (including consent forms) are to be handed over to the Navajo Nation.

The research protocol for the cancelled study followed the steps outlined in the paragraph above, and over the course of almost a year secured approval from two regional tribal councils, two different school boards, the Navajo Nation Board of Education, the Navajo Nation Historic Preservation Office, and finally the Navajo Nation Human Research Review Board. However, preliminary work under the protocol revealed that despite the careful compliance with the tribal requirements for community support, the plan still suffered from an inadequate consideration of sociocultural concerns.

The cancelled study had called for a research team to recruit parents living on the reservation to create recordings of their children’s speech in home-based, naturalistic settings; the goal of this plan was to reduce intrusion and interference from the research team and to protect family privacy. Under this plan, the research team would visit families periodically to collect recordings, but would not be physically present during recording sessions (allowing families to review recordings before sharing the data with the research team). However, the
research plan had failed to adequately consider the importance of individual connections in the community.

As described above, recent trends on the Navajo reservation have strongly supported research that is rooted in the community, rather than that carried out by non-community and particularly non-tribal members. As discussed in chapter 2, the past century has seen many indigenous culture and language projects—almost exclusively conducted by non-indigenous researchers—which promoted the interests of the research team over the interests of the community, and “reinforc[ed] longstanding deficit conceptions of the indigenous participants, their languages and cultures” (Hill & May, 2013: 47). Many Navajo tribal members are extremely interested in regaining control of research concerning their people; a widely circulated quote, popularly attributed to Beverly Becenti-Pigman (current chair of the Navajo Nation Human Research Review Board) and lauded by many indigenous activists (e.g. Walters et al., 2009) encourages tribes to “develop expert Indians rather than Indian experts”. During the process of securing research approval for this failed project, multiple tribal members approached me to discuss the importance of promoting Navajo involvement in the project; one educator admitted to being enthusiastic about the project until learning that I am not, in fact, a member of the Navajo tribe. Hill & May (2013) argue that non-indigenous researchers can be valuable participants in research conducted in indigenous communities, but only those researchers who demonstrate a long-standing commitment to the community over the course of years.

38 Being Navajo is not always correlated with specific physical characteristics or linguistic abilities; any individual having at least one Navajo grandparent is eligible to be enrolled as a member of the Navajo tribe (e.g. Spruhan, 2007-2008), and some members of the Navajo tribe do not have the physical appearance that is traditionally associated with being of Navajo ethnicity.
In preparing for this earlier project, well-established community membership emerged as an extremely important qualification for a researcher hoping to work within the Navajo reservation. The cancelled project had called for family recruitment and data collected to be carried out through regular visits from a research team consisting of me (a non-tribal member) and several native Navajo-speaking linguistics undergraduate students who were Navajo tribal members (but who were not members of the particular community where the research was to take place). However, it quickly became clear that this approach suffered from major flaws stemming from insufficient immersion in the specific community. Jacobsen (2012) notes that two full years of ethnographic immersion—including several months living with a Navajo family and a variety of work/volunteer engagements—were necessary to build adequate relationships with research participants on the Navajo reservation; in fact, a research plan calling for only nine to ten months of fieldwork was described by one of her participants as a “red flag” of a lack of commitment to the community (Jacobsen, 2012: 39). In contrast, the failed research project (prioritizing participant privacy), had selected a commuter approach to data collection, operating on the belief that regular visits to the community—and collaborative work with Navajo research team members—would provide an adequate connection between the research team and the participants. However, community members appeared cautious about the motives and intentions of the research team, and it quickly became clear that a successful version of that research project would require a much greater level of immersion into the community and active participation in the data collection process than the plan allowed. As it was not possible to adequately modify the research plan in a timely fashion, it was ultimately discarded due to
impracticability, and all copies of the consent forms and other intake materials were handed over to the Navajo tribe.

Returning to the question of tribal regulation of research, a further point of interest relates to the publication of research findings. The NNHRRB’s regulations stipulate that any findings from approved research—and therefore any findings from research on the reservation—be submitted for tribal approval before they are submitted for publication or presentation. As pointed out by Brugge & Missaghian (2006: 501), “there is a distinction between prohibiting research and halting or restricting release of findings of approved research”, and some concerns might be raised as to whether the requirements of the NNHRRB could result in restrictions of academic freedom and/or integrity. However, despite the appearance of vulnerability to such limitations, I am currently unaware of any attested instances of the NNHRRB suppressing research findings from approved research protocols, and Brugge & Missaghian report that the NNHRRB firmly denies an interest in such censorship.

This subsection has aimed to provide a brief overview of the process of conducting research on the Navajo Nation. Many other indigenous communities maintain the right to monitor and participate in research activities taking place within their borders, and researchers hoping to work in indigenous communities should be aware of the need to inform themselves of tribal regulations governing research, to collaborate closely with tribal communities throughout the research process, and to be careful of underestimating the importance of establishing long-term relationships in the community.
3.2. Data Collection

Data collection\(^{39}\) for this current project was carried out at an elementary school located in a town close to the Navajo reservation. While not carried out under the jurisdiction of the Navajo Nation, the project aimed to follow many of the same requirements outlined in the regulations of the NNHRRB, and worked with school administrators and the school Navajo language teacher to determine appropriate research practices, determine what services the research team might be able to offer to the community, and to share the information collected as a part of this project. Copies of this resulting dissertation are being made available to the school.

For reasons of privacy, the specific school and its location are not identified. Roughly 50% of the student body at this school self-identifies as Native American, and the remainder of the student body is roughly evenly split between those identifying as Hispanic and those identifying as non-Hispanic Anglo or some other ethnicity. Serving a geographically large community, the school has students who are bussed from 60 or more miles away, with many students commuting from the nearby Navajo reservation.

The school offers a Navajo language immersion program, in which students\(^{40}\) are offered a 30-minute daily lesson on Navajo language and culture. The immersion program—taught by a native Navajo-speaking instructor—follows a pull-out model, and each instructional period

\(^{39}\) This project received approval from the Georgetown University IRB as project 2011-484.

\(^{40}\) All students participating in the immersion program were described as being heritage speakers of Navajo. This term is potentially somewhat misleading; most linguists typically use the term *heritage speaker* in reference to individuals who have incompletely acquired a language, generally in a home-based context (such as the children of immigrants to countries that lack widespread support for the parents’ L1.) In the Navajo community, however, the term ‘heritage speaker’ is commonly used as a simple means of identifying individuals of Navajo ethnicity. Effectively anyone who is of Navajo heritage is considered a heritage speaker of Navajo in a language-learning context, with the term essentially being used to contrast with non-Navajo learners. In this context, the term ‘heritage speaker’ is a primarily cultural rather than linguistic descriptor, and does not provide meaningful information about a language learner’s linguistic knowledge.
serves students from a single grade level. The school has dedicated a classroom to the Navajo language and culture program, and the room features displays of Navajo artwork, Navajo language materials such as calendars and clocks, and news clippings and other printed materials relating to Navajo cultural heritage. The classroom is arranged with several stations around the room, including an area with a large low table (such that an entire class could sit around the table at one time) and a carpeted area where students could sit on the floor. All participants in this study were enrolled in the first or second grade at the time of data collection; the dominant language of all study participants was English.

Recordings were completed by means of Olympus 8100 digital recorders fitted with lavalier (clip-on) microphones. Participating students wore the microphones clipped to a collar or seam on their shirts, and were provided with assistance in setting up the equipment. Recorders were set either on the table, desk, or carpet, or in a pocket. Before each recording session, students were reminded that participation was voluntary, and that they did not have to use the recording equipment or participate if they did not wish to, and that they could turn off the recorders at any point if they wished. However, the recording equipment seemed popular, and students appeared to enjoy using the equipment. While the recorders initially presented a distraction from the lessons, most participants soon appeared to forget they were present. So as to minimize any perception of differential treatment, students who were not participating in the research project were given the option of using the recording equipment during the class period (with all such recordings being immediately destroyed in the presence of the class teacher) and all students, regardless of their participation, were allowed to select a small toy (such as an eraser or small character figurine) after the recording sessions. Twelve students ultimately participated.
over a period of two months, for a total of slightly over twenty hours of recordings. At the end of the data collection portion of the project, the recording equipment was donated to the school for use in the immersion program.

3.3. Data Analysis

Of the original 20 hours of recordings, just over five hours were set aside due to poor sound quality; the analysis is based on the remaining 14 hours and 53 minutes of recorded speech. The most common problem occurred when an external microphone became detached after a recorder had been placed in a participant’s pocket; the students were frequently physically active during the class and it was not possible to maintain uninterrupted visual monitoring of their equipment during the class. In these cases, the sound quality was critically diminished and the recordings were discarded. The other common problem occurred when settings on the recorder were changed after the recording session began; particularly during the earlier recording sessions, the children appeared interested in learning how the recording equipment worked, and some participants ended up changing settings on the recorders while exploring the capabilities of the recording equipment. A number of recordings had to be discarded due to the input volume having been lowered beyond intelligible levels.

Sound files were transcribed with the help of an assistant (an undergraduate student in anthropology), and each utterance was coded for the presence of Navajo lexical items and morphosyntactic material (i.e. all-Navajo utterances), mixed Navajo-English lexical items and morphosyntactic material (i.e. code-mixing), standard English lexical items and morphosyntactic material (i.e. all-English utterances), or non-standard English lexical items and morphosyntactic material (also all-English utterances, but with one or more non-standard features). English
utterances were further coded for a range of properties, including obligatory contexts for do-support, past tense, nominal number marking, auxiliary verb use, third person subjects, wh-word questions, or infinitival constructions. Four categories were chosen for the final analysis: do-support, nominal number marking, use of the past tense –ed morpheme, and use of the third person singular –s morpheme.

Utterances produced by the teacher or other children who were not participating in the project were not transcribed, while names or other potentially identifying information were removed from the transcription or replaced by an initial. Utterances that were exact repetitions of the teacher’s prompt, or utterances that represented the student reading from a text were not included in the analysis. The transcription conventions aimed to capture morphosyntactic features, but did not attempt to represent phonological features aside from those which might indicate an interaction with the morphosyntax of an utterance. Utterances that seemed ambiguous or unintelligible were checked by a second transcriber, and utterances that remained unclear—particularly those where phonetic details might indicate the presence or absence of certain morphosyntactic features—were checked using the acoustic analysis software Praat. Utterances that could not be reliably transcribed after these steps were left out of the analysis; the analysis presented in the remainder of this dissertation draws on a final data set containing 619 utterances.

Data were transcribed and coded in an Excel spreadsheet, allowing for sorting according to characteristics such as speaker or presence of non-standard English morphosyntax. Those utterances that were determined to be non-standard English were later recoded to distinguish between non-standard morphosyntactic material and/or non-standard lexical items in several
categories: nominal number marking, verbal inflection for subject agreement and/or tense marking, and *do*-support.

3.4. Research Design & Methods: Conclusion

This chapter has described the methods of data collection and analysis for this dissertation, and has briefly discussed research practices and sociocultural factors which should be considered when designing research projects taking place in indigenous communities. The next chapter presents the data collected for this dissertation.
Chapter 4: Language Data

As discussed in the preceding chapter, the data collected for this project came from approximately 15 hours of recorded naturalistic speech produced by first- and second-grade students participating in a Navajo language and culture program. The transcription documented only phonological features which appeared to interact with the morphosyntactic form of an utterance; this chapter (and this project in general) does not aim to address the so-called “Navajo accent” in any meaningful way. As this current study focuses only on morphosyntactic and lexical properties, only brief remarks are provided on phonological phenomena which could affect the realization of certain functional morphemes; more detailed work addressing the phonological and prosodic features of Navajo and Navajo English can be found in the work of other authors such as Leap (1982), Penfield (1977), McDonough (2000), or Schaengold (2004).

This chapter aims to address the first of the three research questions guiding this project:

(52) What are the morphosyntactic properties of children’s speech in this Navajo immersion classroom?

The data collected for this project are presented in section 5.1, which provides a profile of the observed morphosyntactic properties. Section 5.2 contains a discussion of developmental factors which might influence the language used by participants in this project, and section 5.3 concludes the chapter.

4.1. Describing the Data

As discussed in chapter 3, utterances were transcribed and coded for the presence of standard or non-standard English morphosyntax, or for the presence of overt Navajo lexical and morphosyntactic material. A fairly high level of intraspeaker morphosyntactic variability was
observed in the data set; while a majority (79%) of the 619 utterances produced by participants adhered to standard English morphosyntactic conventions, a significant number of utterances contained one or more morphosyntactic features not found in standard English. As standard American English may be safely assumed to be fairly well documented in the extensive literature on standard American English, this chapter (and the dissertation as a whole) focuses on those areas of the children’s speech which demonstrate departures from standard American English; the utterances in this data set which appear to correspond to standard American English morphosyntactic rules are discussed only insofar as they inform the discussion of the non-standard constructions. The section is organized as follows: subsection 4.1.1. discusses the verbal complex and the phenomena of do-support, subject-verb agreement, and tense/mood/aspect marking, subsection 4.1.2. discusses nominal inflection, subsection 4.1.3. discusses negation patterns, section 4.1.4. discusses examples of code-mixing, and section 4.1.5. provides a short conclusion and summarizes the findings in a table.

4.1.1. The Verbal Complex

The data contain a number of non-standard morphosyntactic constructions involving the verbal system. This subsection reviews the observed types of non-standard verbal constructions; while some preliminary commentary is provided, a more sustained discussion about the theoretical and practical implications of these patterns is reserved for chapter 6.

Verbal Inflection

As has been reported for a number of other non-standard English vernaculars (e.g. African American English (e.g. Labov, 1972), Appalachian English (e.g. Wolfram & Christian, 1976)), the data show a range of non-standard patterns of verbal inflection. A common pattern, occurring
in 21% of the utterances containing third-person singular subjects, is illustrated in the utterances shown in example (53), where the standard English third-person singular agreement morphology is not used with a third-person singular subject.

(53)  a. I don’t know where he live.
       b. He don’t know.
       c. It taste like a rock.

The constructions illustrated below in (54) appear to show subject agreement marking on the wh-word what, rather than the expected subject pronoun it or that.

(54)  a. I forgot what’s it’s call.
       b. I forgot what’s that.

In the examples shown in (54), the raised what is marked with an –s. One possible explanation for this pattern might be that the –s on what is the contraction of the inflected auxiliary is from the embedded clause, perhaps showing in (54a) with optional reduplication. 41

Finally, example (55) illustrates the widely observed pattern—occurring in 20% of the clear past-time utterances—in which the regular past tense marking was not overtly marked on an inflected verb. The context of the utterance in (55a) clearly indicated that the action was in the past; the teacher was checking to see if anyone needed help turning on their recorders, and the

---

41 Cross-linguistically, it is rather common for embedded questions to follow the same word order as matrix sentence questions, as shown in the following Spanish examples

(i) ¿Qué  comía Carlos? (ii) Me  pregunto  qué  comía Carlos
What  eat.IMP.3sg Carlos  me  ask.1sg what  eat.IMP.3sg Carlos
“What did Carlos eat?”  “I wonder what Carlos ate.”

Holm (2000) reports that many English-based creoles allow for variable use of subject-verb inversion in questions, both embedded and matrix. For example, in Bahamian English (an English-based Atlantic creole) “one finds ‘I can go?’ varying with ‘Can I go?’, and ‘I don’t know where I can go’ varying with ‘I don’t know where can I go.’” (Holm, 2000: 236) The examples in (54) may suggest a parallel pattern in Navajo English.
speaker (whose recorder had been on for several seconds) was reporting that no help was needed. In (55b), the class was working through a lesson on traditional plants and seasonings used by Navajo cooks, and the speaker was recounting a story about having tasted rock salt before.

(55)  a. I already turn it on.
     b. I know what it is, ‘cause my mom give me some of those.

In the absence of adverbials or other contextual clues such as those described above, such utterances may appear to be standard instances of present tense reporting. Temporal adverbials were somewhat rare in these data (only 13 temporal adverbs were observed); possible factors contributing to the small number include the limited size of the data set, the age of the participants, or the classroom context which did not provide a natural setting for storytelling and other situations promoting the reporting of past events. The majority of unmarked past tense verbs were indicated through contextual clues. Utterances that were not clearly demarcated by context or temporal adverbials were considered to be ambiguous, and were not used to inform the discussion on tense marking.

Returning to example (55a) above, to confirm the absence of a final /-d/ on the verb, the recordings were evaluated using the acoustic analysis software Praat, and the spectrograms were examined for any evidence of a stop consonant on the end of the participle.\footnote{It is reasonable to assume that Navajo phonology has had a strong influence on the phonological patterns of Navajo English. Navajo syllables have the form CV(V)(C), and consonant clusters are only known to occur at the boundary between the conjunct verbal prefixes and the verb stem (e.g. Hoijer, 1945; McDonough, 1990, 2000). Not only are word-final consonant clusters unknown in Navajo, but McDonough (2000: 28) points out that only stems may have codas, with bound morphemes exclusively having the form CV or CVV.} Examples (56) and (57) provide a nice comparison pair, as the same phrase—turn(\textit{ed}) it on—is being uttered in both cases. In (56), the closure is clear during the verb-final \textit{d} stop, and there is a clear burst at the
beginning of the vowel $i$ in $it$. In contract, (57) shows no closure between the $n$ of $turn$ and the $i$ of $it$, nor is there a perceptible burst.

(56) “turned it”

(57) “turn it on”

However, while the phonetic record clearly indicates the absence of stop consonants on the end of numerous verbs, these data should be reviewed with caution. Numerous authors (e.g. Penfield, 1977; Schaengold, 2004) have observed that word-final stops are dropped in Navajo-accented English; Penfield provides examples of pronunciations such as *equipment* : [ikwipmen] or *understand* : [andə-stæn], and (comparing the phonological properties of English spoken by members of the Navajo, Hopi, and Mohave tribes) adds that “The past tense ending –(e)d was
dropped consistently in all the varieties of Indian English examined.” (Penfield, 1977:29)

However, the prevalence of word-final stop deletion suggests that the absence of a final /-d/ in utterances such as those shown in (55) above should not be taken as simple evidence of unmarked participles; this current study collected numerous examples of word-final stop deletion, such as first : [ʃʊs], which are presumably not morphosyntactically driven. A carefully designed elicitation task might be able to tease apart the phonologically driven effect of word-final stop deletion from the morphosyntactic features.

Preliminary evidence from this data set suggests that a phonological constraint promoting word-final stop deletion may in fact be at least partly responsible for the prevalence of apparently unmarked past tense verbs; data such as that shown in (58) show that the vowel /ɪ/ is present in the verb, suggesting that the past-tense form did is being used—with the word-final –d being dropped according to presumed phonological constraints—rather than an unmarked stem do (which would presumably contain the vowel /u/.)

(58) Yeah, I di’.

Lending further support to the idea that the unmarked past tense forms might be phonologically conditioned, it is worth noting that the rates of past tense marking differ starkly for regular verbs—which mark past tense with the –ed morpheme—and irregular verbs which mark past tense with a stem change. Of the 52 clearly past-time utterances containing irregular verbs such as be, have, write, or find, 51 (98% of the utterances containing irregular verbs) showed the use of a past-tense form, in contrast to only 42% of the regular verbs. The extremely high rate of past-tense marking in irregular verbs—where past-tense marking is not dependent on a word-final consonant—supports the idea that speakers of Navajo English are not necessarily omitting...
the past tense marking on regular verbs, but that the word-final –ed is not overtly expressed due to phonological avoidance of word-final stops.

However, it would be hasty to dismiss all unmarked past tense forms as phonological artifacts. Irregular verb forms such as went or found are widely believed to be retrieved complete from the lexicon, in contrast to regular verb forms such as tasted or pressed, which are instead generated via the computational system. (e.g. Pinker, 1999) It is not clear whether it is reasonable to conclude that the absence of –ed on apparently semantically past tense verbs should indicate that the speakers are failing to mark past tense on regular verbs—which, unlike irregular verbs, undergo morphological processing to produce the past-tense form—or whether speakers should be thought of as using a past tense morpheme that remains unexpressed due to phonological constraints.

Do-support

The data included 63 utterances containing questions which would typically demonstrate do-support in standard English. Of these, the majority, 63%, showed do-support which appeared to conform to standard English morphosyntactic rules, such as those shown in (59).

(59)  a. How do you work this one?
      b. What does that say?
      c. Where do you live?

However, the remaining 37% of the utterances collected show either absent or non-standard do-support. Questions were frequently formed without the use of do, as shown in (60), or with the apparent use of an alternate auxiliary, as shown in (61).

(60)  a. How much it cost to get them speakers?
b. How this works?
c. What they eat at your house?

(61) How’s you record yours?

Questions that did contain the auxiliary do alternately incorporated it in a non-standard construction, such as that shown in (62a), where the verbal inflection appears on the lexical verb, rather than the auxiliary (as would be expected in standard English), or in (62b-d), where verbal inflection appears on both the auxiliary and the lexical verb.

(62) a. What do your grandma eats?
    b. How did you did it?
    c. Y., did you got yours?
    d. How di’ you pressed it?

The use of do-support showed a high level of variability, even in the speech of individual speakers. For example, four seconds after producing the utterance in (62a), the same speaker repeated the sentence with a slight change, this time dropping the final –s on eat such that neither the auxiliary do nor the lexical verb eat bore inflectional morphology, as shown in (63).

(63) What do your grandma eat?

Overall, the data set suggested that do-support is an unstable construction, and that the inflectional patterns associated with do-support are not regularized.

**Participles**

As illustrated in (64), most of the participial constructions appearing in the data set show apparently unmarked past participles, where the participial forms finish in (64a) and sharpen in (64b) appear without expected –ed morpheme. However, there were an extremely small number
of participial constructions collected—only five in the entire data set—and more examples would be needed to draw any meaningful conclusions.

(64)  a. I’m finish with mine.
      b. O., your pencil’s already sharpen!

As with the recordings of apparently bare past tense verbs, the recordings of apparently unmarked past participles were examined in Praat, and the spectrograms were visually assessed for any indication of a word-final stop consonant. As discussed in the section treating past tense verbs, it seems wise to approach these data with caution even aside from the small sample size, as the possible interaction between the phonological and morphological systems confounds the interpretation of the observed patterns. Again, while outside the scope of this current project, a carefully designed elicitation task would be helpful for understanding the underlying patterns of verbal inflection in this speech community.

4.1.2. Nominal Number Marking

Most utterances containing standard English obligatory contexts for plural number marking showed overt plural marking such as that shown in (65a) and (65b).

(65)  a. Look, I’m putting my pictures right here . . .
      b. Those guys are laughing.

A widely discussed feature of many creole languages is the tendency to mark aspect rather than tense on the verb, frequently through the use of verbal auxiliaries rather than bound morphemes (e.g. Singler, 1990). The data set collected here was not suitable for researching the presence of this pattern in Navajo English; not only was the size small, but data from young child speakers seems poorly suited to investigating non-acquisition related questions about the tense/mood/aspect system of a language due to the potential for incomplete acquisition of relatively complex semantic systems. A valuable direction for future research would be to explore whether—using data from adult speakers—this pattern might occur in Navajo English.
However, a number of collected utterances showed non-standard number marking; after setting aside plural noun tokens that represented repetitions of the teacher’s prompt\(^{44}\), 4 of the remaining 27 nouns occurring in obligatory contexts for plural number marking showed either non-overt plural marking as shown in examples (66-67).

(66) ‘Cause baby, baby aren’t supposed to taste chili beans.

(67) Speaker A: Grandma can’t drink that. Speaker B: Yeah, they can.

In (66), the plural form *aren’t* taken in conjunction with the apparent general semantic interpretation of the nominal *baby* suggests an obligatory context for plural noun marking in Standard English. In (67), the use of a modal obscures the intended number feature associated with the subject *grandma*, but the response given by speaker B supports the interpretation of (67) as a general statement, typically indicated in Standard English with a plural. Further instances of apparently absent plural marking are shown in (68), where plural determiners are used with unmarked nouns.

(68) a. L. likes those kind.

b. Um, those thing, those circle thing?

Other utterances show nouns which appear to exhibit non-standard number marking of some sort, but which lack other clear syntactic indications of number. For example, in utterances such as (69) it is unclear whether nouns such as *smoothie* represent unmarked plurals or whether the speaker might be treating the noun as uncountable.

__________

\(^{44}\) The complete data set contained 45 tokens of plural nouns, but many of these were repetitions of the teacher’s prompt, often repeated by multiple students. For example, 10 tokens of the noun *rocks*, including four separate utterances of the phrase *looks like rocks* were collected immediately after the teacher’s description of a type of bluish clay traditionally used by Navajo people. To avoid artificial weighting of the data, repetitive tokens were left out of the analysis.

108
(69) My grandma likes to drink, drink, smoothie.

As with many of the non-standard verbal features described above, nominal marking in this data set shows significant variation and both marked and unmarked plural nouns were fairly common. Assessing the rates of unmarked plurals, however, is somewhat challenging; many utterances did not provide clear evidence of standard English obligatory contexts for plural marking (such as utterances where the noun appears in object position without a number-marking determiner, or without a clear contextual clue to the intended number).

4.1.3. Negation

As has been reported for a number of other vernacular forms of American English (e.g. African American English (Labov, 1972; Green, 2002), Appalachian English (Wolfram & Christian, 1976), Mohave English (Penfield, 1977) and West Texas English (Foreman, 1999)), so-called double negation or negative concord is found in the collected data; as shown in example (70), the standard English NPIs any/anything are replaced by the negatives no or nothing.

(70) a. I’m not pressing no buttons.
    b. I don’t have nothing here.

Sentential negation in the absence of negative concord items appears to follow the same principles as standard English, as shown in (71).

(71) a. No, I’m not turning, stop it.
    b. Don’t say it!
    c. . . . I can’t see those things.

No examples were found of otherwise non-standard mechanisms for indicating sentential negation.
4.1.4. Code-mixing

Finally, a small handful of utterances—47 of the 619 total utterances—contained examples of simple code-mixing. As illustrated in the examples below, effectively every instance of code-mixing appeared to conform to Muysken’s category of insertional code-mixing, where a single Navajo lexical item was inserted into an otherwise English utterance. No instances of sustained alternational code-mixing or congruent lexicalization were observed.

(72) Áádóó one time when I was a baby . . .
    And then one time when I was a baby
    “And then one time when I was a baby . . . “

(73) shí-sister
    1.sg.POSS-sister
    “my sister”

(74) Ha’ísh your favorite color, red?
    What.INT your favorite color, red?
    “What’s your favorite color, red?”

(75) Yágo. Give me yágo dootl’ízh.
    sky.AS Give me sky.AS it-is-green/blue
    “Sky-ish. Give me blue.” (=Give me the blue crayon.)

(76) It smells like dleesh.
    It smells like clay
    “It smells like clay.”

In a number of instances of observed code-mixing, the class was working on a particular production activity to practice a given set of Navajo vocabulary words which contained the Navajo lexical items that appeared in the code-mixed utterances. In (73), the speaker was listing off family members, and appeared not to recall the Navajo word for ‘sister’, using the English word instead. In other instances, such as (75), the class was discussing color terms, and was
coloring pictures while talking about the color names; the utterance in (76) was uttered during a class lesson about traditional Navajo uses for naturally-occurring minerals in the area, and the teacher had brought in samples to support the vocabulary lesson. Further discussion of these code-mixing data is reserved for chapter 5.

4.1.5 Data Conclusion

To conclude, the aim of this section has been to present the data in a neutral and descriptive manner, avoiding any sustained discussion of potentially contributing influences or likely implications of the data. The following table shows the rates of departure from non-standard English in each of the areas discussed in this section.

(77) Rates of Non-Standard English Constructions

<table>
<thead>
<tr>
<th>Construction</th>
<th>Number of obligatory contexts</th>
<th>Number of non-standard constructions</th>
<th>Rate of non-standard constructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plural Marking</td>
<td>27</td>
<td>4</td>
<td>15%</td>
</tr>
<tr>
<td>Do-Support</td>
<td>63</td>
<td>23</td>
<td>37%</td>
</tr>
<tr>
<td>Past Tense Marking</td>
<td>97</td>
<td>19</td>
<td>20%</td>
</tr>
<tr>
<td>Past participle constructions*</td>
<td>5</td>
<td>3</td>
<td>60%</td>
</tr>
<tr>
<td>Third-Person Singular Verbal Agreement</td>
<td>143</td>
<td>30</td>
<td>21%</td>
</tr>
</tbody>
</table>

*The extremely small number of past participles does not support robust conclusions; these data are included solely because they inform the discussion of the more widely attested use of overt/covert past tense marking.

As mentioned above, the data in this chapter comes from recordings of twelve participants over a two-month period. It was unfortunately not possible to provide a meaningful assessment of
interspeaker variation in this data set. Due to the close proximity of the participants during recording sessions, some utterances could not be definitively assigned to a speaker, hampering any rigorous investigation into interspeaker variation. Personal observation during recording sessions suggested that the majority of the non-standard English utterances were produced by just two participants; future research, taking this into consideration, would benefit greatly from implementing methodologies allowing for more careful attribution of individual utterances to specific speakers\textsuperscript{45}. The next section discusses the potential impact of developmental factors in the observed data; a discussion of the role of linguistic influences is reserved for chapter 5.

4.2. Origins of Non-standard Features: Developmental Effects?

Before launching into a sustained discussion of potential contact language phenomena, it is important to address the obvious possibility that the observed data are attributable to simple developmental factors. Section 4.2. turns to this question of the potential developmental origins of the variation observed between standard English and the vernacular speech produced by the study participants. As noted above, there is no reason to believe that any departures from standard English are attributable to the speakers’ status as English language learners; all the participants in the project were English-dominant, and had very low levels of Navajo fluency. However, child speakers do not always have adult-like command of their target languages; it is reasonable to wonder whether observed departures from standard English might be attributable

\textsuperscript{45} While the lavalier microphones were intended to limit ambiguities in speaker identification, their efficacy was reduced by the physical realities of the classroom environment. In particular, the participants frequently engaged in activities which involved moving around the classroom and sitting very close to one another—either on the carpeted floor or at their child-sized tables—reducing the perceptual differences between the voice of the designated speaker for that recording device and the participant’s classmate(s). A possible improvement for future classroom-based research might be to employ a visual recording component to help assess the identity of individual speakers.
to incomplete language acquisition. However, this straightforward explanation does not seem to be supported by an examination of the prevalence of the observed non-standard features.

The data presented in the preceding section were all produced by first- and second-grade students. Previous research has found that typically developing children of that age produce morphosyntactic speech errors at relatively low rates, which vary according to the particular morphosyntactic feature being assessed; Rice et al. (2008) reported, for example, that typically developing 7-year-old children in their study produced non-standard constructions associated with do-support in 9% of utterances. In contrast, the data set collected for this current project show that 36% of relevant utterances contain non-standard do-support constructions. Other features in this data set show similarly high rates of non-standard expression; after removing repetitive tokens, 15% of the clearly obligatory contexts for plural noun marking had no overt plural morpheme (and due to the prevalence of ambiguous cases especially for nouns in object position, the true rate of unmarked plurals is likely higher). These rates are in contrast to the findings of Matthews & Theakston (2006), who reported that standard-English acquiring 6 year old children in their study produced plural marking on 98% of high-frequency regular nouns, 99% of low-frequency regular nouns, 90% of high-frequency irregular nouns, and 93% of low-frequency irregular nouns. While developmental factors may certainly impact the prevalence of certain types of non-standard features observed in these data, comparisons to age-matched speakers exposed primarily to standard English suggest that the data collected for this project should not be attributed solely to the incomplete acquisition of morphosyntactic rules.

In fact, while incomplete acquisition of standard English morphosyntax seems unlikely to account for the observed rates of non-standard features, developmental factors may play a role in
a slightly less straightforward manner. Research on other English vernaculars suggests that sociolinguistic factors affect the expression of non-standard vernacular features, and some research on African American English suggests an age-related effect where children who are speakers of non-standard English may be less likely to employ non-standard features in early elementary school, with rates being higher during the preschool years and climbing again in later elementary school. A study by Van Hofwegen & Wolfram (2010) has reported that age was significantly correlated with the rate of expression of non-standard features in African American English (AAE); taking data from a 17-year longitudinal study of 32 children being raised in AAE-speaking communities, they found that “many of the children at 48 months exhibit higher relative vernacularity, and that this vernacularity diminishes in the early grades (1 and 4), while sometimes increasing again in the later grades” (Van Hofwegen & Wolfram, 2010: 436). If the non-standard features observed in this data set are in fact indications of a true vernacular, then the findings of Van Hofwegen & Wolfram suggest that the six- and seven-year old participants in this current project might in fact be producing non-standard vernacular constructions at lower rates than would be found in the speech of other speakers in their communities. If this were the case—and this is of course purely speculative—it would suggest that a more fruitful investigation of Navajo English might work with participants who do not fall into the “early elementary school” age group.

4.3. Language Data: Conclusion

This current chapter has presented the data collected for the current project, and has discussed potential developmental origins for the presence of the observed non-standard English features. The chapter concludes by arguing that the non-standard features observed in the data set cannot
be solely attributed to the incomplete acquisition of standard English morphosyntax. While all of these constructions exhibited some degree of variation, the discussion pointed out that the rates of non-standard usage exceed levels that could plausibly be attributed to speech error or incomplete acquisition of standard English, and argued that the rates of non-standard usage suggest a target language differing materially from standard English. Having ruled out potential developmental origins for the observed non-standard features, chapter 5 turns to the discussion of the factors which might have influenced the departures from standard English grammar.
Chapter 5: Discussion

This chapter turns to the remaining two research questions guiding this project:

- RQ2: To what extent do the levels of morphosyntactic regularity support classification of the children’s dialect of English as a contact language?
- RQ3: What are the implications of the observed language patterns for bilingual families and educators?

The chapter is divided into three sections; section 5.1. discusses the theoretical implications of the data and argues for the suitability of a contact language analysis for the vernacular speech documented in this project, while section 5.2. discusses the implications for families and educators working with Navajo language revitalization efforts. Finally, section 5.3. describes directions for future research and provides a conclusion for the chapter and the dissertation.

5.1. Theoretical Implications: Contact Language Phenomena

This section is structured around the second research question, namely discussing the merits of a contact language analysis for the documented vernacular language. As all the participants in this project were highly English-dominant—or even effectively monolingual speakers of English—this discussion does not entertain the notion that any contact language effects might be innovations on the part of the children; rather this discussion assumes that any non-standard language is representative of the language used by adult speakers in the children’s speech community. Many of the non-standard features documented in this data set have been argued to be characteristic of creole and semi-creole languages, and the following discussion argues that the vernacular speech documented in this project meets the argued linguistic profile for a semi-creole as described in Holm (2004) or McWhorter (1998). While this discussion focuses primarily on the linguistic properties of the vernacular speech, the historical prevalence
of diachronic factors in identifying creole and semi-creole languages motivates a review of the sociohistorical profile of Navajo English in comparison to those of other argued semi-creole languages.

5.1.1. The Case for Navajo Origins

As mentioned above, all participants in this project were considered to be highly English-dominant, and had low levels of productive Navajo fluency; there is no reason to believe that patterns of non-standard constructions in the data stem from the speakers’ status as English language learners or from innovative linguistic transfer from a home language. Given the linguistic profile of the participants, the following discussion assumes that non-standard features documented in the children’s speech are representative of the speech used in the Navajo community in this area. As discussed in section 1.2.3, researchers such as Penfield (1977), Leap (1982), Holm (2004), and Schaengold (2004) have reported that effectively monolingual speakers in many Navajo communities use a Navajo-influenced English vernacular. This discussion works from the assumption that the data collected for this project are examples of that same Navajo-influenced speech; not only do the data correspond to other descriptions of Navajo-influenced English, and not only are the children all members of a Navajo community, but the data collection site—a Navajo language and culture classroom—is a setting in which speakers might be particularly likely to employ speech styles with stronger Navajo cultural associations.

46Penfield (1977) reports that some speakers exhibit register-shifting behavior, switching between Navajo English and standard English. While this current project did not explore contextualized differences in speech styles, Schaengold (2004: 32) reports that “Teenagers in particular may use [Navajo English] to establish their identity as Indians”; younger children, such the participants in this project, are presumably aware of this linguistic tactic of asserting ethnic identity. As this current data set was collected during a Navajo immersion class—a setting in which the use of the Navajo language was strongly promoted, despite the students’ limited command of that language—the use of Navajo-influenced English vernacular might be expected.
As outlined in chapter 4, the data collected for this project include constructions which do not appear to follow the morphosyntactic rules of standard English. The following discussion compares the four non-standard constructions highlighted in chapter 4 with comparable constructions in standard Navajo. Several parallels emerge, as reviewed below.

In the verbal system, the variable use of do-support in interrogatives is plausibly traceable to Navajo influence. Navajo, like most languages worldwide, does not employ a dummy auxiliary for question formation.

(78) a. Ha’át’íísh biniiyé díí shéínishóód?
   What.INT 3.sg.for-the-purpose-of DEM 1sg.2sg.drag-over
   “Why did you drag this thing over to me?”

   b. Na’álchíníísh hóló?
   2sg.child.INT 3sg.exist
   “Do you have children?”

As shown in the examples in (78), the lexical verb remains in-situ and the interrogative sentence function is indicated through the use of question words and interrogative particles. The English dummy auxiliary do has no parallel construction in Navajo, and the variable and innovative uses of the do-support constructions observed in these data might be influenced by the absence of do-support in Navajo.

The prevalence of apparently unmarked past tense verbs and past participles is somewhat more surprising, given the extremely robust system of Navajo verbal inflection. However, a closer look at the differences between the Navajo and English systems of verbal inflection shows only minimal overlap in the specific types of information encoded in the overt verbal inflection. Schaengold (2004) points out that the Navajo verbal system does not in fact mark past tense (although it does mark future tense), employing instead a system of progressive/perfect aspectual

118
marking where past time is indicated through adverbials; an argument could be made that the absence of the \(-ed\) morpheme on Navajo English past tense verbs might be the result of influence from the Navajo verbal system.

However, this argument should be taken with caution. As discussed in chapter 4, there is some reason to suspect that phonological influences may be implicated in the absence of the \(-ed\) morpheme, and that the variation from standard English may not be solely traceable to morphosyntactic influences. While the argument above might account for the absence of the \(-ed\) morpheme on past tense verbs, it does not necessarily account for the corresponding absence of the \(-ed\) morpheme on past participles, where the morpheme presumably represents perfect aspect and not past tense. While it is certainly plausible that the Navajo system of representing past tense through adverbials plays a role in the reduced rate of overt past tense marking in Navajo English, the expression of the \(-ed\) morpheme is almost certainly also dependent on morphophonological factors.

Non-overt \(-ed\) morphemes were not the only non-standard features observed in the verbal system; a number of utterances with third person singular subjects employed verbal constructions that appeared without the standard English \(-s\) (such as my grandma like or he live). It is challenging to connect this pattern to Navajo morphosyntax; Navajo, as discussed in chapter 2, marks subjects, objects, and even indirect objects through a complex system of preverbal affixes, and it would seem that the English third person \(-s\) (part of a vastly less sophisticated system), would be a natural counterpart to the Navajo system of argument-verb agreement. Like the variably expressed past tense \(-ed\) morpheme, though, the absence of third person \(-s\) might also be the result of morphophonological factors; standard production of the \(-s\) morpheme on verbs such
as *walk* or *taste* would produce a coda consonant cluster, which—as discussed in section 4.1.1.—are not permitted in standard Navajo phonology.

Turning now to the non-standard features observed outside the verbal complex, as noted above, plural number marking in the collected data is sporadic, with numerous utterances containing apparently unmarked plural nouns in both subject and object positions. Recalling back to chapter 2, standard Navajo has no plural marking of nominals—aside from a handful of irregular nouns—that is semantically comparable to English plural marking, and instead relies solely on verbal morphology and/or context to indicate nominal number. The occasional absence of overt plural number marking in the collected data⁴⁷ potentially shows the influence of standard Navajo, where regular nouns are not inflected to show plural number.

Further supporting the notion that the vernacular English observed for this project is influenced by Navajo, a number of lexical items in the data set suggest analogies to Navajo (or, at least, non-English) lexicalization strategies. For example, in (79), the speaker’s use of the phrase “cow meat” over the standard English “beef” suggests an analogy to the standard Navajo term for “beef”, *béégashii bitso’*, which literally translates to “cow meat”.

(79) She likes to eat cow meat with bread?

Another example which clearly shows a non-English nominalization strategy is shown below in (80), where the speaker was asking for the location of a pencil sharpener; the use of the

---

⁴⁷ Analyzing the apparently variable use of plural number marking on English nominals was somewhat challenging, particularly for decontextualized utterances since there were not always clear morphosyntactic cues indicating the semantic number of a given unmarked nominal. If verbal agreement—a cue only available for the identification of nouns in subject position—was not indicative, and semantic number could not be inferred from accompanying articles, determiners, or quantifiers, the unmarked nominal was assumed to be semantically singular. The true rate of semantically plural unmarked nouns is presumably higher than the stated documented rates.
innovative construction *sharpen things* suggests a potential analogy to the typical means of nominal construction in Navajo\(^48\).

(80)  Where’s your sharpen things at?

As discussed in chapter 2, Navajo has only a small handful of prototypical nouns, with the great majority of nominals being formed of (often quite complex) nominalized verbal constructions. For example, the Navajo word for table, *bikáá’adání*, literally means “sitting around it”. A more typical English-speaker’s approach to such a lexical gap might be to employ an alternative nominal phrase such as “that thing that sharpens pencils”; this speaker’s use of the verbal construction *sharpen things* might show the use of Navajo nominalization strategies rather than those typically employed by speakers of standard English.

In conclusion, of the four highlighted non-standard constructions documented in this data set, two (variable *do*-support and non-overt plural number marking) seem be reasonably associated with morphosyntactic influence from Navajo, while the remaining two (non-overt past tense *–ed* and 3\(^{rd}\) person singular *–s* agreement) seem to be more challenging to assign to Navajo morphosyntactic influence, but may show the influence of the Navajo phonological system.

\(^48\) Another, perhaps more straightforward explanation, is that the nominal *sharpen things* can be traced to the influence of Spanish. While none of the participants in this project identified as Hispanic, the Navajo community in this area has had regular exposure for many years to Spanish-speaking communities. Spanish agentive deverbal nouns show a strikingly similar pattern, as shown in the examples in (i)

(i)  a. lava-platos  
     wash,3sg-dish.PL  
     “dishwasher”  
   b. gira-discos  
     turn,3sg-disc.PL  
     “record player”  
   c. abre-latas  
     open,3sg-can.PL  
     “can-opener”

The Navajo English phrase *sharpen things* appears quite similar to these constructions; while there is no apparent 3\(^{rd}\) person inflection on the verb *sharpen*—as would be expected in the comparable Spanish construction—as pointed out in chapter 4 roughly 20\% of the present-tense utterances with 3\(^{rd}\) person singular subjects showed verbs without the standard English *–s* agreement morpheme.
5.1.2. The Case for Origins in Contact Language Universals

Assuming that the vernacular English documented for this project shows influences from Navajo, of course, does not rule out the likelihood of it also exhibiting parallels to other contact languages. As discussed in chapter 1, many researchers working on contact language phenomena have argued for the existence of a typological class of creole and semi-creole languages whose members share some number of linguistic properties; proponents of this idea (e.g. Holm, 2000, 2001, 2004; McWhorter, 1998, 2005; Krämer, 2013) suggest that creole and semi-creole languages are characterized not only by their genetic ancestry but also by a set of synchronically identifiable creole characteristics. This is not to suggest that the genetic influence of the superstrate and/or substrate language(s) is dismissed, but simply that there are a set of prototypical creole features which serve as a sort of default setting for the creole grammar. Under this view, for example, a prototypical creole property that is present both in a contact language’s superstrate and/or substrate language(s) would be almost certain to appear in that contact language, while a prototypical creole property that appears in neither the superstrate and/or substrate languages would be less likely to occur in the given contact language. The preceding subsection discussed the possibility of Navajo language influences; this current section, however, turns to the discussion of parallels to contact language universals.

As discussed in chapter 1, there is little reason—both for linguistic and historical reasons—to believe that the English currently spoken in Navajo communities has undergone a dramatic process of pidginization, creolization, decreolization, or any other such process of fundamental linguistic restructuring. The vast majority of the morphosyntactic system represented in this data set suggests the presence of the same morphosyntactic system as is found
in standard English, with only minimal modification. Looking briefly at the criteria for a full creole language proposed by McWhorter (2011)—the criteria that a creole language shows little to no use of lexical tone, inflectional morphology, or noncompositional combinations of derivational markers—it is clear that it would be challenging to argue that Navajo English is a true creole language; while lexical tone appears to be absent in the Navajo English examples (an unsurprising finding, given the minimal role of lexical tone in standard English), and there are few examples of compositionally opaque morphology (also perhaps unsurprising, given the small size of the data set and the age of the speakers), it would be difficult to argue that the data set (which includes utterances such as D., she said no more sharpening or it looks like those round balls) shows little to no use of inflectional morphology.

A full creole analysis for Navajo English, of course, has not been the subject of much (if any) serious academic consideration. As discussed in section 1.2.3., previous researchers (e.g. Bartelt, 1982; Schaengold, 2004) have made arguments which support a classification of Navajo English as a semi-creole or partially restructured language, a category of contact languages suggested to be typographically far more similar to their superstrate languages than a full creole. Indeed, many (although not all) of the specific features suggested by Holm (2004) to be characteristic of this proposed class of contact languages are in fact found in this data set, as illustrated in table (81)\textsuperscript{49}.

\textsuperscript{49} As pointed out in section 1.2.3., a troubling observation about Holm’s semi-creole checklist is the fact that many (if not all) of the features on this list are in fact quite cross-linguistically common, and a decontextualized diagnostic application of this checklist results in a high rate of false semi-creole identifications. The discussion returns to this point at the end of section 5.1.2.
Following Holm (2004), the table in (81) uses the + symbol to indicate the presence of a characteristic in the language in question and the 0 symbol to indicate the absence of that characteristic; for example, the first line shows that standard Navajo does not have the feature ‘zero 3s PRES inflection’, i.e. that Navajo does use verbal agreement morphology on present-tense verbs with third-person singular subjects. As (81) shows, the Navajo English data collected for this project demonstrate 8 of the 18 proposed features (in comparison, for example, to the 15

<table>
<thead>
<tr>
<th>Verb Phrase</th>
<th>NE</th>
<th>NAV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Morphology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zero 3s PRES inflection</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Zero 1p PRES inflection</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Zero PAST inflection</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Aux./preverbal marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semantic influence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Negation          |    |     |
| Negative concord  | +  | +   |
| Discontinuous double |    |     |

| Noun Phrase       |    |     |
| Number            |    |     |
| Zero plural inflection | +  | +   |
| Unbound pluralizer | 0  | +   |
| Associative plural |    |     |
| Gender            |    |     |
| No agreement in NP | +  | +   |

| Possession        |    |     |
| [possessor 0 possessed] | 0  | 0   |
| [possessed 0 possessor] | 0  | 0   |

| Pronouns          |    |     |
| Reduced case marking |    |     |
| Zero reflexive pronoun | +  |     |

| Clauses           |    |     |
| Word Order        |    |     |
| QW S-V/Aux (direct) | 0  | +   |
| Dependent clauses |    |     |
| Zero subject REL  |    |     |
| Zero subordinator “that” | +  | +   |

| Total number of +’s | 8  | 11  |

+ = attested presence of feature  
0 = attested absence of feature  
= not applicable or unknown

NE = Navajo English  
NAV = Navajo
found in African American English or the 9 in Afrikaans). Several caveats should be mentioned here, though; due to the limited size of the data set, it is possible that some “semi-creole” features were simply not observed. For example, almost no examples of possessive constructions were collected; the absence of the Holm’s proposed semi-creole possessive construction is only strictly indicative of the near-absence of possessive constructions in the data. Furthermore, for several of the unchecked categories above, the data set did include qualifying examples, but in such small numbers that no meaningful conclusions should be drawn from them; a larger sample size would allow for more robust conclusions to be drawn, and might well indicate that this variety of English demonstrates more of the features shown in the table. Even allowing for these limitations, though, the checklist’s synchronic linguistic evidence suggesting partial restructuring for Navajo English—with the caveat that this small data set provides only preliminary evidence—appears to be as strong as that supporting a partial restructuring analysis for several of Holm’s exemplified semi-creoles, such as Afrikaans and nonstandard Caribbean Spanish.

Of course, as noted in chapter 1, applying a checklist such as this is not necessarily an effective diagnostic method for identifying partially restructured languages. Similarly to many of the other substrate languages presented in the original table in section 1.2.3., Navajo itself shows 11 of the 18 features; uncontroversial examples of creoles such as Gullah or Jamaican Patois do not differ significantly from the so-called semi-creoles in their numbers of these features. As suggested in chapter 1, the linguistic criteria outlined in this chart appear to be limited in their ability to distinguish semi-creoles from languages with other diachronic origins; at the very least, it seems reasonable to suggest that these linguistic criteria should be considered necessary but not sufficient grounds for inclusion in the class of semi-creole languages.
While Holm’s checklist is perhaps crippled by its very specificity, McWhorter (2005, 2011) adopts a more general approach in instead proposing that semi-creoles are distinguished by “significant” simplification from the superstrate language in one or more of the three grammatical areas proposed to comprise the Creole Prototype. As mentioned in section 1.2.2., McWhorter argues that “if a grammar is new . . . distinguishable from older grammars in terms of particular grammatical features which are known to arise only over time” (McWhorter, 2005: 10); these three particular ‘features’ are argued to be inflectional affixation, tone, and derivational noncompositionality. Arguing against the claim of other researchers (such as Mufwene, 1997) that the semi-creole/creole distinction is invalid, McWhorter writes:

The Creole Prototype . . . contains exactly the criteria in question [to justify the category of semi-creoles]. Afrikaans (sic), for example, has significantly reduced the inflectional affix paradigms of Dutch, but it retains Dutch’s derivational morphology in much of its evolved idiosyncrasy as well as a wealth of other elaborifications typical of Germanic . . . In contrast, Negerhollands Dutch Creole had no inflectional affixes, and what derivational morphology it had was semantically regular. For precisely this reason, we can confidently classify Afrikaans as a semi-creole, in contrast to the Dutch creole Negerhollands.

McWhorter, 2005: 157

McWhorter does not, however, offer a clear diagnostic for distinguishing between significant and slight simplifications, and it is not clear how one might use the Creole Prototype alone to distinguish between the tonal reduction found in the (semi-)creole Lingala—a Swahili & Bobangi based putative semi-creole which McWhorter argues shows ‘slightly reduced’ tonal distinctions from the tonal system of the lexifier Bobangi—and modern-day Mandarin, which has four lexical tones rather than the eight argued to have existed in Late Middle Chinese (e.g.
Wang, 1986). However, momentarily leaving aside this concern over over-identification of semi-creoles, the Navajo English documented in this project—which includes utterances showing no verbal agreement morphology for 3rd person singular subjects, no verbal inflection for past tense verbs, no nominal inflection for plural nouns, and questions without standard English-like do-support—does indeed show a slight simplification of English inflectional morphology, which under the criteria outlined by McWhorter should presumably be sufficient to ‘confidently classify [Navajo English] as a semi-creole’.

While the purely linguistic evidence does seem to provide support for classifying Navajo English as a contact language, the distinction between the specific subcategory of semi-creoles from full creoles based only on linguistic features seems somewhat fraught; both the proposed diagnostics discussed here seem to result in a potentially unacceptable rate of false positive identifications. The vulnerability of these purely linguistic diagnostic tools reinforces the argument that linguistic criteria are necessary but not sufficient for accurate identification of semi-creole languages; it seems that reliable identification of a semi-creole language might require sociohistorical contextualization.

Looking at the sociohistorical background of Navajo English, the case for a semi-creole analysis is in fact strengthened. Holm suggests that a full creole emerges only in a language contact environment where native speakers of the superstrate language represent less than 20-25% of the population; in circumstances where that proportion is exceeded, Holm suggests that a semi-creole is likely to emerge. Historically, the Navajo-speaking population has represented far less than 75% of the population in most speech communities where English is the language of communication; as discussed in chapter 2, while communities within the reservation have always
had predominantly Navajo populations, as recently as the 1980s the language of communication in most Navajo-majority communities has typically been Navajo, with English being reserved for communication with non-Navajos. Boarding schools were the exception to this generalization; however, there is no evidence that suggests that any fully developed English-based creole language was ever in use in Navajo boarding school communities. While admittedly not impossible, it would be surprising for a fully developed creole language to have arisen in Navajo boarding schools, have spread to non-boarding school communities, and have then undergone a process of decreolization, all without undergoing any sort of documentation.

As discussed in chapters 1 and 2, the linguistic landscape in Navajo communities is somewhat complex—as is perhaps the case in any speech community where multiple languages come into contact—and it is worth reiterating that there are in fact two different contact languages which have been described in the Navajo community: Navajo English and Bilingual Navajo. Reports of the language used in Navajo boarding schools make mention of ‘boarding school Navajo’ and ‘Navajo English’, with the former of these being an English-influenced Navajo (e.g. Harvey, 1974; Schaengold, 2004) and the latter being a Navajo-influenced English (e.g. Penfield, 1977; Bartelt, 1982a,b). As the speech documented in this project cannot reasonably be considered a Navajo-superstrate language, this boarding school Navajo is presumably not a direct ancestor of modern-day Navajo English; boarding school Navajo, is however almost certainly the direct ancestor of modern-day “bilingual Navajo”50 as described in Schaengold (2004).

50 As discussed briefly in sections 1.1.2. and 1.1.3., there is some reason to believe that this modern ‘bilingual Navajo’ could be classified as a mixed lect.
On the other hand, while this dissertation makes no claims of providing a thorough investigation of the diachronic relationship between Navajo boarding-school English and modern-day Navajo English, it seems undeniable that modern-day Navajo English must be a descendent in some fashion of this early Navajo-influenced code. However, this does not imply that early Navajo boarding-school English was a full creole language, nor that the modern-day Navajo English is the result of decreolization. Without further research into the early Navajo boarding-school English, it is not possible to argue meaningfully for any particular contact language classification, but it seems quite plausible that boarding-school Navajo would also have been classified as a semi-creole, perhaps with more basilectal feature than the modern-day Navajo English.

5.1.3. Code-mixing

Although this project initially aimed to explore code-mixing practices in the children’s speech, the code-mixing that was documented in this data set was fairly limited, and the linguistic profile of the participants in this current project was not compatible with the exploration of mainstream theories of modeling code-mixing in the speech of balanced bilingual speakers. The speakers who participated in this project were all highly English-dominant, and did not have fluent command of Navajo; as the two mainstream theories of modeling code-mixing (the null-hypothesis theory and the MLF model) assume basic fluency in both of the speaker’s languages, it would be unreasonable to attempt any comparison of these data to the predications of those two models.

In fact, however, the code-mixing that was observed appears to be somewhat characteristic of the Bilingual Bootstrapping Hypothesis (Gawlitzek-Maiwald & Tracy, 1996;
Gawlitzek-Maiwald (2000, 2003), which assumes that for unbalanced bilingual children, code-mixing is motivated by gap-filling strategies and morphosyntactic transfer from their dominant language (in this case, English). The examples of code-mixing collected here are in line with this hypothesis; effectively all the instances of code-mixing were comprised of inserted lexical items\(^ {51} \), such as in utterances such as “shi-sister” (during an activity where the class had been directed to describe their families in Navajo, this particular utterance followed several apparent attempts to produce the first syllable of Navajo word for “little sister”) or “It looks like a rock but it’s dleesh” (where dleesh, the Navajo word for a specific type of blue-white clay, was used instead of the presumably unknown English term). Some all-Navajo utterances were collected, but these were primarily repetitions of sentences produced by the teacher.

As the speakers who participated in this project were all developing Navajo fluency—but were not currently fluent speakers of Navajo—it is not particularly surprising to note that their code-mixing practices are most closely aligned with the predictions of the Bilingual Bootstrapping Hypothesis. The instructor for the language and culture program reported that students in the higher grades produced more spontaneous Navajo speech and had higher overall levels of Navajo fluency; an interesting avenue for future research would be to compare code-mixing practices in Navajo immersion classes.

\(^ {51} \) The collected examples of rudimentary Navajo-English code-mixing are perhaps most similar to the prestige-driven use of foreign words in an otherwise monolingual discourse. This pattern—examples might be a modern American English speaker using the French phrase allons-y! meaning ‘let’s go’ to encourage his American friends to go somewhere, or a modern Japanese speaker choosing to use the English word ekusukyuuzu ‘excuse’ (Stanlaw, 2004: 183) in a Japanese-language newspaper intended for a monolingual Japanese audience—shows speakers borrowing words from a language of perceived cultural prestige. The prestige of the Navajo language in the Navajo community is somewhat complex, where the language is simultaneously viewed as an old-fashioned language of poverty and yet also a language of cultural richness and authenticity. In the case of this immersion class, students—who do not have high levels of productive Navajo—are strongly encouraged to use Navajo, but are also encouraged to engage in content-based discussions about Navajo cultural practices and history; balancing the demands of communication and the push to use Navajo, the low-Navajo-fluency students appear to have chosen the strategy of incorporating individual Navajo lexical items into their otherwise English sentences.
mixing practices in the higher grades with those observed in these first- and second-grade classrooms. As noted in chapter 1, some researchers (e.g. Paradis et al., 2000) have argued that the code-mixing practices of extremely unbalanced bilinguals is qualitatively different from those of balanced bilinguals, and comparing data from bilingual speakers with each of these profiles would allow for an empirical examination of that hypothesis.

5.2. Practical Implications: Language in the Community

Section 5.2. serves as a discussion of the third, and final research question, namely discussing the implications of the observed language patterns for bilingual families and educators. This section is divided into two parts; subsection 5.2.1. discusses the implications of this project’s findings for families and communities, while subsection 5.2.2. discusses the implications for language educators and schools. Although there is considerable overlap between these two groups, this discussion attempts to highlight the two areas of interest.

5.2.1. Implications for Families and Communities

As suggested above, it is important to note that the language documented in this project was collected in very specific circumstances, and the particular linguistic details described here should not be divorced from the context in which the data was collected. The immersion program which served as the data collection site enrolls a very small subset of the local children, and it would be imprudent to assume that these students are linguistically representative of children in the greater community.

In fact, there is some reason to suspect that the study participants might not be linguistically representative of children in the greater community. While the study participants came from a wide geographic area that was representative of the local Navajo community—with
some commuting from the local reservation and others living in the off-reservation town—their enrollment at the public school was due to parental choice. Little formal research has examined factors influencing Native American parental choice when selecting between on- and off-reservation public schools for their children, but anecdotal evidence provided by community members suggests there is a perception in this community that families who select on-reservation schools place greater importance on the role of Navajo culture in the school, while families who choose off-reservation schooling place greater importance on the availability of college preparatory resources. Platero (2012) points out that these dueling interests—the desire for cultural and linguistic immersion and the desire for access to western educational opportunities—are not always mutually compatible, and that “many Navajo families … immerse themselves in western education and culture while abandoning their own in hopes of becoming successful in a dominant society.” (Platero, 2012:4) The anecdotal observations in this community are supported by the findings of researchers such as Lansing (2011: 175), who—drawing on parent interviews in a different Navajo community in Arizona—reports that “[p]arents and students who seek a more rigorous academic program leave the community” to attend off-reservation schools. Despite the fact that some off-reservation schools—including the one serving as the data collection site for this current project—do offer Navajo language programs, there is a widespread perception that schools located within the Navajo Nation are more closely connected with the Navajo community and that students at these schools experience a more authentic, integrated Navajo education (e.g. Batchelder, 2000, McCarty & Dick, 1996). Families in this community who give the greatest priority to the development of Navajo language abilities might be more likely to choose to send their children to the on-reservation school; there
is some reason, therefore, to suspect that levels of Navajo proficiency might tend to be lower in students attending off-reservation schools. If this is the case, the students who participated in this project might tend to be less fluent in Navajo than their peers who attended the on-reservation school.

Of course, this is largely speculative, as the project did not collect comparative data for other children or families in the community. Furthermore, the preceding conversation on Navajo fluency is perhaps somewhat tangential, as the project focused primarily on non-standard English speech; there is no obvious evidence supporting a correlation between fluency in Navajo and the use (or lack thereof) of non-standard “Navajo English”. However, the preceding conversation may have relevance to the use of Navajo English, as well. Schaengold (2004) points out Navajo English holds a stigmatized status amongst many people in both Navajo and non-Na

vajo communities, and that despite its holding a “covert prestige” as a signifier of Navajo cultural identity (particularly among teenagers), the speech style is often the target of jokes and ridicule. There is a widely held perception that Navajo English is the result of laziness, poorly developed language skills (Haskan (2007) reports many Navajos referring to Navajo English as “unstructured English”), or ignorance. It is reasonable to imagine that families with the strongest negative feelings about Navajo English might be more likely to choose the school perceived to be most academically rigorous, where there is a perception that “improper language” would be

52 While the students in this project produced very little Navajo during the recording sessions, it is not clear that the recordings are a strictly accurate representation of the children’s productive Navajo fluency. As discussed, there is significant synchronic variation among the various speech communities that make up the Navajo Nation. This particular speech community is widely claimed to speak a non-standard variety of Navajo which itself shows influences from other indigenous languages such as Apache, and it is quite plausible that children would be hesitant to use non-standard Navajo in the classroom. The potential discrepancy between the defacto standard dialect of Navajo promoted in the classroom and the regional variant which might be spoken in a child’s home is yet another reason to strive to implement home-based data collection methods in future research.
targeted. While a survey of parental attitudes would be needed to provide clear evidence on this point, it is fair to conclude that the children who participated in this project are not necessarily linguistically representative of their peers in the community, and that the language data should be interpreted with reference to the context in which it was collected.

In fact, there is good reason to believe that the context in which these data were collected might have a strong effect on the prevalence of non-standard features in the children’s speech. Penfield (1977:26) notes that “. . . many children are able to switch from S[standard] E[nglish] to I[ndian] E[nglish] on certain occasions,” and it is reasonable to imagine that some—if not all—of the participants in this study might have this ability. While the children who participated in this project were all young—ages six and seven—numerous researchers (e.g. Wagner et al., 2010) have pointed out that children even younger than this are sensitive to linguistic register and can make adult-like inferences about linguistic register choices. There is good reason to believe that the children who participated in this project would be able to engage in register shifting between Navajo English and standard English.

These data were collected during Navajo language and culture classes, where use of the Navajo language was heavily encouraged; the adult instructor spoke predominantly in Navajo, and students frequently urged one another (using English) to speak in Navajo. Although all the students in the class were heritage speakers of Navajo—and had at least some exposure to Navajo at home—their productive levels of Navajo fluency were low, and students demonstrated only very basic conversational abilities in Navajo. Without reliable access to the strongest linguistic means of expressing ethnic solidarity—through the use of standard Navajo—it is
plausible to imagine that the children may have adopted a strategy of employing Navajo English as an available substitute.

The use of an ethnically-linked accent has been widely identified as a tool used to express ethnic solidarity (e.g. Padilla & Borsato, 2010). This context, the Navajo language and culture classroom, was the only setting in the school where Navajo students interacted solely with other Navajo students; as mentioned in chapter 4, the off-reservation school serves a population that is roughly 50% Navajo, with the remainder of the students being largely non-Hispanic white or Hispanic. Looking at the prevalence of non-standard AAVE features in the speech of African-American elementary-school students, Renn (2010) found that speakers used significantly more AAVE features in educational contexts with higher percentages of other African American students, while speakers in educational contexts with lower percentages of African American students tended to use the fewest AAVE features in their speech. In this current project, the prevalence of non-standard features in the children’s speech is arguably affected by ethnically homogeneous educational context within the Navajo language and culture classroom; while no data was collected outside the Navajo language and culture classroom, informal observation suggests that the participants in this study might employ fewer Navajo English features when interacting with their non-Navaajo peers outside the Navajo language and culture classroom.

In many bilingual Navajo families in this community, standard Navajo fluency is limited to the older members of the family, and Schaengold (2004:9) points out that in many Navajo communities, “[standard] Navajo is recognized as an “adult” language”. Even in families where Navajo is spoken, younger children are often English-dominant or effectively monolingual English speakers. Schaengold reports that when speaking with their Navajo-speaking family
members, children and other individuals who do not have fluent command of standard Navajo often employ the mixed lect (referred to in her paper as Bilingual Navajo) as a means of bridging the communication gap; however, as this project found, not all children in the Navajo community have command of Bilingual Navajo. The findings of this paper suggest that in social contexts where the use of the Navajo language is preferred, children who are not functional speakers of Navajo or Bilingual Navajo may employ Navajo English as a means of establishing their Navajo linguistic identity.

Outside the classroom, in bilingual families where a Navajo proficiency gap exists between the older and younger generations, children may adopt a similar strategy when speaking with Navajo-speaking family members. Community concerns about the prevalence of “unstructured English” or “rez speak” (Haskan, 2007) may be tempered by increasing awareness that Navajo English is in fact a “legitimate” vernacular form of English (rather than an imperfectly acquired or lazy attempt to produce standard English). In addition to increasing exposure to standard English input, families who are interested in promoting proficiency in standard English may benefit from encouraging explicit metalinguistic awareness of contexts where standard English or Navajo English would be more appropriate.

However, it is worth noting that some speakers may employ Navajo English as a means of establishing Navajo cultural authenticity; in such a case—where the speaker uses Navajo English as a tool for demonstrating his/her Navajo identity—the use of Navajo English might reasonably be taken as an indication of positive identification with Navajo culture, or positive attitude towards Navajo culture. One of the factors widely argued to affect a language learner’s chances of successful or timely acquisition of a second language is the learner’s attitude towards
the speakers and the culture of the target language (e.g. Dörnyei, 1998; Masgoret & Gardner, 2003, Csizér & Dörnyei, 2005). An interesting direction for future research would be to explore whether habitual use of Navajo English is predictive of improved achievement in Navajo immersion classes.

5.2.2. Educational Implications

The role of non-standard vernaculars in the classroom has been hotly debated amongst researchers, educators, and families; outside Native American communities the discussion in the United States has primarily focused on the role of African American English (AAE) in the classroom. A crucial difference, however, between the role of AAE in the classroom and the role of Navajo English is the target language assumed for each educational context. In the current project, the contextually encouraged target language in the classroom was Navajo, while in the case of AAE, the target language in the (non-foreign language) classroom is typically standard American English. This project did not explore the children’s use of Navajo English in other educational contexts where standard American English would be the target language of instruction; as discussed above there is some reason to suspect that many speakers of Navajo English might engage in register shifting, and that speech outside the Navajo language and

53 A somewhat infamous example of this debate was highlighted in a controversial 1996 resolution passed by the Oakland School Board in Oakland, California. The resolution (No. 9597-0063) stated that African American English (which the resolution termed Nigritarian English) was a unique language showing a “West and Niger-Congo African linguistic structure”, was “not a dialect of English”, and that African American students should receive instruction “both in their primary language and in English”. Criticisms of the resolution included the complaint that it promoted African American children being taught in African American English rather than standard American English, reducing their access to employment and educational opportunities requiring command of standard English. The Oakland School Board later released an amended resolution (No. 9697-0063), clarifying that students would not receive instruction in AAE, but that educators should consider AAE-speaking students to have similar linguistic needs in the classroom as those students coming from non-English speaking homes. Brown (2008) points out that the highly publicized controversy surrounding the Oakland School Board’s resolution “has made curricular reform guided by linguistic research a more politically fraught process” (Brown, 2008:12).
culture classroom might differ from the speech patterns documented in this project. This project restricts itself to the consideration of Navajo English speech within the context of the Navajo language and culture classroom.

Support for Navajo language revitalization\textsuperscript{54} is widespread among Navajo educators and communities, and programs similar to the one where this current project took place have been implemented at numerous schools across the reservation and in nearby communities. McCarty & Dick (1996: 4) write that “. . . community-controlled schools have been at the forefront of American Indian bilingual education and a growing movement to stabilize and revitalize indigenous languages”, and this effort has been continued by many non-tribally run schools in and near the Navajo Nation. However, the requirements of the No Child Left Behind (NCLB) act of 2002 have been widely reported to have had a detrimental effect on schools’ ability to provide meaningful instruction in Navajo language and culture, and numerous researchers (e.g. Haskan, 2007; Balter & Grossman, 2009) write that English—although generally not standard English—

\textsuperscript{54}One challenge—one of the many—for language revitalization work is the question of how to define the target language for restoration. A living natural language is rarely monolithic, and the effects of both synchronic and diachronic variation can result in numerous divergent forms of the language with varying levels of documentation. This matter is particularly challenging in the case of dormant languages with no native speakers, where there may be no documentation of various lexical or grammatical properties of the language. For example, Warner et al. (2007) discuss the revitalization work being done with Mutsun, a Yok-Utian language whose last known native speaker died in 1930; Warner et al. point out that although the Mutsun people began a language revitalization project in 1996, the work has been made more challenging by the fact that there is very little documentation of Mutsun suprasegmentals, “lexical semantics is a severe and constant problem” (Warner et al., 2007: 59), and there is effectively no documentation of pragmatics and discourse structure.

Navajo is of course not a dormant language, and modern language revitalization efforts are able to work with native speakers’ rich knowledge of Navajo lexical semantics, articulatory phonetics, and pragmatics. Here, the challenges are slightly different; although Navajo still has many native speakers, has been fairly well documented, and has a generally acknowledged standard form, spoken Navajo shows a good deal of regional variation, and features of the language used by older native speakers are not always present in the speech of fluent younger speakers (such as observance of the Navajo animacy hierarchy in argument order). Such discrepancies between the local colloquial form of the language and the standard form can create additional challenges for children in the immersion classroom.
is the default peer-to-peer language of communication in the elementary school Navajo immersion classroom.

If, as suggested in section 5.2.1., students are in fact using Navajo English in the Navajo immersion classroom as a means of expressing ethnic solidarity, then the prevalence of Navajo English in lower-level Navajo immersion classes might be less alarming than has been suggested. Although the target language—standard Navajo—is not being used, the children’s use of an ethnically-linked non-standard form of English might function as an entry-level means of linguistically performing their Navajo identity. The instructor for the Navajo language and culture classes where the current project was conducted reported that older students used more Navajo in peer-to-peer interactions in the classroom than the younger students; the demonstrated progression of Navajo fluency suggests that the predominance of Navajo English in first- and second-grade Navajo immersion classes should not be interpreted as a death knell for the Navajo language.

5.3. Conclusion

This chapter has discussed the theoretical implications of the data collected for this dissertation, and has commented on the practical implications for families and language educators who are committed to promoting children’s fluency in the Navajo language. In brief, the discussion in section 5.1. reviewed a number of non-standard morphosyntactic phenomena which were consistently present in the collected data: null plural number marking, non-standard patterns of do-support, null agreement marking on present-tense third-person singular verbs, and apparently null marking on past-tense verbs and past participles. Having argued that the rates of non-standard constructions materially exceeded levels that would be expected in the speech of age-
matched children acquiring standard English, the discussion reviewed the suitability of a contact language analysis, working through potential parallels between Navajo morphosyntax and the documented departures from standard English morphosyntax, as well as parallels to other argued semi-creole languages. Section 5.1. concluded by proposing that the vernacular speech used by the children in this project meets the criteria for a partially restructured or semi-creole language, showing influences from Navajo grammar and meeting two different proposed semi-creole profiles.

The discussion in section 5.2. turned to the social and educational implications of the project’s findings, and reviewed several reasons to exercise caution in generalizing the particular linguistic findings to other children in the Navajo community. The discussion went on to argue that the use of Navajo English can plausibly be linked to speakers’ desire for linguistically demonstrated ethnic solidarity, and that such use in the immersion classroom may be a ground-level step towards developing Navajo fluency.

Beyond their own impact, the findings of this project offer a springboard for exploring directions for future research. Several limitations have been discussed earlier in this chapter, most notably the limited size of the data set, the restriction of the data collection site to a single classroom, and the restriction of the participant pool to the children enrolled in the first- and second-grade language and culture programs at this single school. Further limitations include the absence of supporting data such as parent interviews or surveys on family attitudes towards language education and bilingualism.

To address these concerns, a variety of future research projects might be undertaken. One obvious project would be to build to the existing data set to provide a more robust profile of the
morphosyntactic (or phonological) properties of the vernacular Navajo English. An interesting addition to this goal might be to build a data set incorporating data from both children and adults, allowing for the exploration of any qualitative grammatical differences between the speech of adults and children in the community. While naturalistic data—as was collected for this project—is often valuable for the exploration of non-standard language features, a disadvantage is the amount of time required to collect effective quantities of data. Adopting a research design which relied on the collection of elicited speech in the form of short stories or social interviews might allow for more efficient data collection.

Another particularly interesting project would be to examine the relationship of higher (or lower) rates of Navajo English use to students’ later proficiency levels in standard Navajo. Earlier in this chapter it was hypothesized that the use of Navajo English might be considered a signifier of positive attitude towards Navajo language and culture; however, this current project did not provide strong evidence of this claim. To provide such strong evidence, a future project might employ a longitudinal design, and collect initial Navajo proficiency results from students at the beginning of their academic year in the immersion class. In-class recording at the beginning of the academic year could provide data on rates of Navajo English use; year-end Navajo proficiency tests could provide evidence of students’ levels of Navajo acquisition. While a multi-year study design would presumably provide the most conclusive results, it seems likely that reasonably meaningful results could be obtained within a single school year. Participant interviews or surveys could be used to support the understanding of any link between Navajo English use and attitudes towards Navajo language and culture.
An additional direction for future research would be to investigate any differences between the language used in the Navajo language and culture classroom and in the other classes at the same school. As noted above, the students enrolled in the pull-out Navajo language and culture program also participated in standard elementary classes with peers who were not Navajo and did not participate in the Navajo language and culture program. Like the project described in the preceding paragraph, this study might also begin by conducting participant interviews to assess attitudes towards Navajo culture, towards the Navajo language, and towards Navajo English; data collection for such a study would take place in at least two sites, with one data set being collected in the Navajo language and culture classroom, and the other being collected in some other classroom with a roughly comparable structure (language arts or social studies would be likely candidates for consideration, although more research would be needed to determine the most appropriate comparison environment). Either naturalistic data or elicited story-telling could be used to investigate this question; for the latter study design, participants could be interviewed in the classroom setting and could be asked to describe a picture or recount a story. The data collected in these two settings could be compared for specific phonological, morphological, or syntactic features to determine the relative prevalence of Navajo English features in the children’s speech in the two educational contexts.
Bibliography


Auer, Peter. (2000) Mixing is functional (but what are the functions?). *Bilingualism: Language and Cognition* 3(2): 101-102


Bakker, Peter & Mous, Maarten. (Eds.) (1994) *Mixed languages. 15 case studies in language intertwining*. Amsterdam: IFOTT.


Barreiro, Antonio. (1832). *Ojeada sobre Nuevo-México, que da una idea de sus producciones naturals, y de algunas otras cosas que se consideran oportunas para mejorar su estado, é ir proporcionando su futura felicidad.* Puebla, J.M. Campos.


Franciscan Friars. (1912) Vocabulary of the Navajo Language. St. Michaels, AZ.


Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3-15


Hoijer, Harry. (1945). *Navaho phonology*. University of New Mexico publications in anthropology, (No. 1)


151


Lee, T. S. (2007). “If they want Navajo to be learned, then they should require it in all schools”: Navajo teenagers’ experiences, choices, and demands regarding Navajo language. *Wicazo Sa Review*, Spring, 7–33.


162


Young, R. and W. Morgan (1943). *The Navaho Language*. Salt Lake City, Education Division, United States Indian Service


