An Experiment in Computational Parsing of the Navajo Verb

Mans Hulden
Shannon T. Bischoff
University of Arizona

1 Introduction

This paper presents preliminary research on a computational parser for core Navajo morphology where any inflected verb is automatically decomposed, together with the inflectional and derivational structure of the verb. The grammatical implementation largely follows Faltz (1998) and Young and Morgan (1987); Young et al. (1992). We also report some proposals for reducing the amount of allomorphy and phonological rules in the description of Navajo verbal morphology, and potential uses of such a parser.

After an introduction to the elements of the Navajo verb and the notation employed by lexicographers, pedagogists, and other scholars to express those elements, a brief discussion of subjects and objects is included. We also present some problems posed to the novice learner of Navajo attempting to navigate the primary reference material Young and Morgan (1987); Young et al. (1992). Finally, we discuss the parser and the advantages it presents to the Navajo learner. We conclude with remarks for future work on the proposed Navajo parser.

2 The Navajo Verb

In this section we will provide a general overview of how the Navajo verb is presented and discussed in the primary reference material and in the (Faltz, 1998) primer. For notational simplicity, Faltz' formalism will be adopted. However, while it does not follow the traditional sources Young and Morgan (1987); Young et al. (1992) in the letter, it does follow them in spirit.

2.1 General Assumptions: Templates and morphophonological complexity

There are several general assumptions regarding the Navajo verb arrived at by Young and Morgan (henceforth YM) and adopted by Faltz (1998), among others, based on the phonological and morphological facts of the language (cf. esp. Young and Morgan (1987)). The Navajo verb:

1. is morphologically complex
2. it corresponds to phonological word
3. its stem is usually word final and monosyllabic
4. is commonly divided into tripartite structure based on morphotactic and phonological facts: *disjunct*, *conjunct*, and *stem*

The tripartite structure presented in Table 1 illustrates each of these generalizations. It should be noted, as mentioned, for notational simplicity we adopt a simplified structure more in line with Faltz (1998) as opposed to the more articulated structure of YM with 16 positions.

Table 1: *Navajo verb tripartite structure.*

outer prefixes	plural	object	inner prefix	subject	classifier	stem
disjunct prefixes		*conjunct prefixes*		*stem*		

This filling of a template can be illustrated in Table 2 with a concrete instantiation of the verb meaning 'They made a fire.'

Table 2: **dideidíɬjéé'** 'They made a fire' (Faltz, 1998, p. 172)

outer prefixes	plural	object	inner prefix	subject	classifier	stem
di	**da**	**y**	**d**	**íí**	**ɬ**	**jéé'**
disjunct prefixes		*conjunct prefixes*		*stem*		

In Table 2 we see a phonological rule that changes **da+y** to **dei**. We also see each "slot" is filled in our template. In the following example we see that it is not necessary for each prefix position to be filled. Also, a prefix position is not strictly limited to one morpheme. This is the result of our simplified

Table 3: **didideeshjah** 'I will make a fire' (Faltz, 1998, p. 171)

outer prefixes	plural	object	inner prefix	subject	classifier	stem
di	**Ø**	**Ø**	**d+d**	**eesh**	**ł**	**jah**
disjunct prefixes		*conjunct prefixes*		*stem*		

template which departs from YM in that they assume a slot for each distinct morpheme.
In Table 3, **didideeshjah**, we see that a phonological rule inserts **i** between **d+d**.[1] It should be noted that the Navajo verb consists minimally of a subject, classifier (both of which may be null), and a stem. Next we turn to some interesting semantic relationships between the stem and obligatory prefixes.

2.2 Semantics of the Verb

The general meaning of a *verb* is arrived at via the combination of the various lexical morphemes and the stem—importantly the word-final stem does not provide the semantic content alone to derive necessary meaning. Meaning is arrived at via the combination the morphemes in the *disjunct*, *conjunct*, and *stem*.

2.2.1 Stem sets

It appears to be standard practice in the Navajo literature to organize the stem (excluding the subject and the classifier) into stem sets. Following YM, Faltz (1998) also organizes stems into such sets. These are based on five elements, seen here in Table 4:
From these "sets" a general semantics can be derived. That is, from the combination of each element listed in the stem set a semantics is arrived at. Next we will expand on each of these elements individually.

2.2.2 Stem sets: mode

YM organize stems in sets that correspond to the seven *modes*. Faltz adopts this practice, often limiting the sets to the five primary modes, since the two missing modes can always be derived from the others. It will be seen below

[1] **d+d** is a portmanteau morpheme that is further divided in YM87, but not in Faltz.

Table 4: Stem sets in Navajo are seen in the left-hand column. The right-hand column shows the information necessary for constructing a complete "lexical entry" of a Navajo verb—these include the classifier, knowledge about lexical prefixes in different positions, the verb's transitivity, and a conjugation pattern. The classifier and some lexical prefixes also make it highly likely that a verb carries a certain telicity, or transitivity, but these two patterns are not completely regular.

Stem set	Classifier:
I:	
P:	Lexical prefixes:
F:	
R:	Transitivity:
O:	
	Perfective: s-P
	...

that the phonological realization of a given stem may vary in the different modes.

- *Imperfective*: I-Mode
  - indicates ongoing action in the present, or sometimes in the immediate future
- *Perfective*: P-Mode
  - indicates a completed event in the past - this mode is primarily divided into the y-P and s-P sub-modes
- *Future*: F-Mode
  - indicates an event in the future
- *Iterative*: R-Mode
  - indicates generic action, emphasizing the repetitive nature of the action

- *Usitative*: U-Mode
  - indicates generic action, emphasizing the habitual nature of the action
- *Optative*: O-Mode
  - this mode is used in certain dependent structures whose meaning involves a notion of desire
- *Progressive*: Pg-Mode
  - many Navajo verbs do not have a progressive mode, it is used rather than the imperfective to indicate ongoing action in the present

It should be noted that not every stem set employs each of the modes. We turn now to obligatory morphology, that is, the obligatory prefixes.

2.2.3 Stem sets: Obligatory Prefixes

Faltz (1998) identifies a number of obligatory prefixes—prefixes which must occur with the stem in all modes to arrive at the general semantics of the verb—many of these prefixes do not always avail themselves to a ready semantic reading in isolation. Some examples follow where outer corresponds with disjunct and inner conjunct and no gloss is given:

outer prefixes	inner prefixes
ha	**d**
na	**j**
b'iki	**'**

A challenge posed by these obligatory morphemes will be discussed more at length below. In short, while the general semantics of a given verb in Navajo may be regular across dialects, the morphological parts of that verb may be analyzed in different ways by different speakers. That is, some speakers may analyze the morpheme **ha**, for example, as representing an unspecified object, while others may interpret it as a locative. However, both interpretations may contribute appropriate semantic content to arrive at the same semantics for a given complete verb-form. This will be clarified further below.

2.2.4 Stem sets: Classifier

There are four obligatory classifiers in Navajo—a stem must appear with one:

Navajo Classifiers	
barred-l	**ɬ**
plain-l	**l**
d-classifier	**d**
zero	**Ø**

The exact function and meaning of these morphemes is not clear. The barred-l is generally associated with transitive constructions, while the plain-l is associated with intransitive. These two are collectively referred to as valence markers (McDonough, 2000a). A number of phonological rules often reduce the various classifiers to null. These rules will not be presented here, but it should be kept in mind that a given classifier may not surface due to phonological effects.

2.2.5 Stem sets: telicity

Stem sets may be divided into pairs, with one occurring with the mandatory na prefix the other without:

Table 5: Stem set pairs

- Stem sets with na:
  - atelic in nature—referred to as continuative aspect stem-set
- Stem sets without na:
  - telic in nature—referred to as momentanous aspect stem-set

Not all stem sets come in such pairs. Such a generalization holds true much of the time, but like many other phenomena in the language is not completely regular.

2.2.6 Stem sets: examples

Here are two examples of stem sets which form a pair. The first illustrates the continuative aspect and the second the momentanous aspect. Each is followed by an example of a fully inflected verb. Note that the stem set in each example has different phonological realizations for the various modes in some, but not all cases. In general, there is no apparent regularity between stems in the various modes of a given stem set. Some heuristics for guessing other stems based on knowledge of the stem in one mode is given by Faltz (1998), but such are not expected to work with any consistency.

Table 6: Continuative aspect stem-set—atelic 'carry him/her around'

Stem set	Classifier: barred-l
I: té **P**: tį́ **F**: teeł **R**: teeh **O**: teeł	Lexical prefixes: **na** (outer)
	Transitivity: transitive
	Perfective: s-P

Table 7: **naníshté** 'I am carrying you around' (Faltz, 1998, p. 118)

outer prefixes	plural	object	inner prefix	subject	cl	stem	mode
na		**n**		**sh**	**ł**	**té**	**I**
disjunct prefixes		*conjunct prefixes*		*stem*			

2.3 Interim Summary

The verbal word in Navajo comprises a stem and a variety of obligatory morphemes which in turn derive the semantics of the verbal 'word.' The verbal word in Navajo minimally consists of a stem, classifier and subject (the latter two of which may be null). Further, it has been demonstrated that the *templatic* notion employed captures the facts of the Navajo verb composition

Table 8: Continuative aspect stem-set—telic 'carry him/her up out of something'

Stem set	Classifier: barred-l
I: teeh	
P: tį́	Lexical prefixes: **ha** (outer)
F: téél	
R: téeh	Transitivity: transitive
O: téél	
	Perfective: s-P

Table 9: **hashíłteeh** 'You are carrying me up' (Faltz, 1998, p. 113)

outer prefixes	plural	object	inner prefix	subject	classifier	stem
ha	**Ø**	**sh**	**Ø**	**ni**	**ł**	**teeh**
disjunct prefixes		*conjunct prefixes*		*stem*		

rather straightforwardly. To illustrate two points—that obligatory morphology contributes necessary meaning and that speakers may interpret the verb word similiarly and its parts differently—the following example is provided. Faltz provides the following stem list for the verb word 'sing.'

Faltz further provides the following paradigm for the Perfective mode.

Note the I-mode forms in the below paradigm, in particular the first singular with the **ha** initial morpheme.

We can illustrate the derivation of **hashtaał** as follows if we assume Faltz' analysis of **hw** as an obligatory object morpheme:

Note here that a phonological rule in *this* particular form does change **hw** into **ha**.

Navajo speakers indicate that you cannot simply say **taał** on its own. That is, the prefix must accompany the stem. What is even more interesting is that some speakers analyze the prefix morpheme not as an obligatory unspecified or fourth person object, but rather as a locative meaning 'up and out.' Despite the different interpretations of the morpheme the semantics of the verb remains the same, adding strength to the claim that it is not the stem alone that derives meaning, but a composition of the lexical morphemes and

Table 10: 'sing' (lit. sing it) (Faltz, 1998, p. 132)

Stem set	Classifier: barred-l
I: taał	
P: táát	Lexical prefixes: **hw** (4th P object)
F: tał	
R: tał	Transitivity: transitive
O: taał	
	Perfective: s-P

Table 11: P(y)-mode paradigm 'sing (it)' (Faltz, 1998, p. 133)

	sg	dpl	distr dpl
1	hóótáál	hwiitáál	dahwiitáál
2	hwíínítáál	hootáál	dahootáál
3	hóótáál		dahóótáál
4	hojíítáál		dahojíítáál

the stem. We will return to this form and the potential ambiguity between the outer **ha**-prefix and the 4th person object prefix **hw**. Next we turn to a brief discussion regarding the phonological realization of subjects and objects.

2.4 Subjects

Faltz (1998) identifies 28 subject paradigms, a different sub-set for each of the seven possible modes. Some modes have several conjugation patterns. The P mode, for example, is divided into the P(s), P(y), and P(n) sub-modes, or si-perfective, yi-perfective, and ni-perfective, respectively (the latter being YM's terminology). The subject allomorphy is usually determined by the combination of the classifier, mode, and whether it is preceded by a disjunct prefix, conjunct prefix, or no prefix.

Each mode has a set of allomorphs for a variety of phonological realizations depending on various phonological conditions. A potential complication

Table 12: I-mode paradigm 'sing (it)'

	sg	dpl	distr dpl
1	hashtaał	hwiitaał	dahwiitaał
2	hótaał	hohtaał	dahohtaał
3	hataał		dahohtaał
4	hojitaał		dahojitaał

Table 13: **hashtaał** 'You are carrying me up' (Faltz, 1998, p. 113)

outer prefixes	plural	object	inner prefix	subject	classifier	stem
Ø	**Ø**	**hw**	**Ø**	**sh**	**Ø**	**taał**
disjunct prefixes		*conjunct prefixes*		*stem*		

for the novice learner. In addition to the 1st through 3rd person subjects, there is also a 4th person, or "alternate third person" form which will be briefly discussed.

Some allomorphs are not fully specified: for example, in the yi-perfective, the hallmark third person subject form **yi** is sometimes represented as just two underspecified vowels **vv**—vowels which are copied over from earlier phonological material in the disjunct.

In addition, the fourth person subject form is primarily marked both in the object and the subject positions—i.e. the template is not entirely rigid.

2.4.1 Fourth person or 3a "Alternate third person forms"

The Navajo fourth person has the following properties.

- indicates an unspecified or indefinite subject (like "one" or "they")
- polite way of referring to somebody who is present in conversation
- used for reference tracking, to keep one third person entity separate from another when telling a story
- obligatorily appears with third subject
- may appear adjacent to an object

The fourth person appears in conjunct position, for reasons of notational economy in 'object' position here. YM provide a special position, a second subject position, which this morpheme appears in. An example follows.

Table 14: **jíícha** 'The one cried' (Faltz, 1998, p. 114)

outer prefixes	plural	object	inner prefix	subject	cl	stem	mode
		j	**0**	**íí**	**Ø**	**cha**	**P(y)**
disjunct prefixes		*conjunct prefixes*		*stem*			

Next we see the 4th person **j** with the indefinite or unspecified object, the glottal **'**:

Table 15: **'ajoodląą́** 'The one drank (s.t.)' (Faltz, 1998, p. 127)

outer prefixes	plural	object	inner prefix	subject	cl	stem	mode
		'+j		**oo**	**d**	**dląą́**	**P(y)**
disjunct prefixes		*conjunct prefixes*		*stem*			

NOTE: **'+j** → **'aj** (phonological rule)

2.5 Objects

Unlike the subject prefixes the object prefixes lack a series of allomorphs for the various modes. Faltz provides the following list of objects for Navajo transitive constructions:

Table 16: Object prefixes in Navajo

	sg	dpl
1	sh	nih
2	n	nih
3	y or b or zero	
4	hw	
unspec	'	

Some examples with simple transitive constructions follow.
An example with the fourth person **j**:

Table 17: **hashíłteeh** 'You are carrying me up' (Faltz, 1998, p. 113)

outer prefixes	plural	object	inner prefix	subject	cl	stem	mode
ha	**Ø**	**sh**	**Ø**	**ni**	**ł**	**teeh**	**I**
disjunct prefixes		*conjunct prefixes*		*stem*			

NOTE: **ní** → **i** (phonological rule)

Table 18: **nashijistį** ''One carried me around' (Faltz, 1998, p. 119)

outer prefixes	plural	object	inner prefix	subject	cl	stem	mode
na		**sh+j**	**Ø**	**s**	**ł**	**tį**	**P(s)**
disjunct prefixes		*conjunct prefixes*		*stem*			

NOTE: **sh** → **shi**, **ł** → **Ø** (phonological rules)

With a basic understanding of how the Navajo verb is decomposed and the notational devices (the template and the stem sets) used to illustrate this decomposition, we turn now to a discussion of the Navajo parser.

3 The parser

3.1 Overall design & implementation

The parser has been constructed using standard finite-state methods and toolkits. We shall not give the details of such a treatment here, pointing instead the reader to comprehensive sources such as Beesley and Karttunen (2003) and Sproat (1992). The advantage of such an approach is that the grammar can be described in the direction of generation, while being inherently bidirectional—i.e., very efficient parsing and generation are possible using the same description. The grammar can be specified either using ordered rewrite rules (somewhat akin to SPE-style rules (Chomsky and Halle, 1968), or unordered parallel constraint rules (Koskenniemi, 1983)). We've chosen the former, simply as a matter of taste. Also, we've taken advantage of feature unification available in the so-called Xerox formalism (Beesley and Karttunen, 2003), which allows for a compact description of constraints on morphotactic co-occurrence.

The parser has so far been implemented to handle verbs conjugated in the **I**, **P**, and **F** modes, and has a lexicon of the most common lexical prefixes found in the language. However, no verb stem lexicon is present—verb stems

are defined as **any** monosyllabic unit that does not begin with the unvoiced fricatives **s, sh, ɬ, h**. The motivation for this is simple: we want to focus on capturing the phonology and morphology of the language, while keeping the option open to link the parser to a more comprehensive lexicon.

In the choice of morphological structure we have, as is clear from the above discussion, chosen to follow Faltz (1998). While it may be convicingly argued that YM is more comprehensive in its treatment of morphotactics, this work also postulates a number of allomorphs which are subject to phonological change. The circumstances for these phonological effects are not outlined with the same rigor as is Faltz' treatment, which is why we've chosen to follow the simpler template.

While Faltz' rules appear to cover most of the alternations found in the language, we have made substantial changes to many of the observations, and treated the rules mainly as generalizations about Navajo verb structure and phonology. The reason for this is twofold. Firstly, Faltz' rules mix morphological description and phonological description to a large extent. This leads to a large number of phonological changes that are purely conditioned by morphology. While the results of such an approach are not incorrect, implementation of two constantly interacting levels of description is far more cumbersome than isolating morphotactics from phonology. Naturally, a complete isolation in this respect is not possible—a residue of phonological changes that cannot be detached from morphology seems to be unavoidable.

3.2 Organization

The grammar operates on two levels of abstraction. From the point-of-view of generation, the highest level is a very abstract description of morphotactics. This level is mapped to a more concrete level, at which most phonological rules operate and map a verb to its surface form. This can be illustrated as follows:

Abstract morphotactic and semantic description

⇕

Morpho-phonological description

⇕

Surface form

Maintaining this perspective, the generation of a surface form can be illustrated as follows, where we map an abstract form **ha+...+l+táál+Py+S4** to its surface form, **hajooltáál**. Here the symbols **ha**, **l**, **Py**, and **S4** in the abstraction indicate that the verb has an outer prefix **ha**, the classifier is **l**, the verb is to be conjugated in the yi-perfective, and its subject is to be in the 4th person.

Underlying	[ha+X|X+X+X+lˆtáál]+Py+S4
	⇕
Intermediate form	[ha+0|j+0+oo+lˆtáál]+Py+S4
	⇕
Surface form	hajooltáál

Here the **X**'s in the underlying representation represent *any* morpheme, to be inserted between that level of description and the lower one, in accordance to feature-driven morphotactic constraints. This example derivation is fairly straightforward. There are no phonological rules that apply, and the surface form is effectively a simple concatenation of the morphemes. However, in the mapping of the most abstract form to the concrete phonological form, there are a number of conditions on morphotactic well-formedness that come into play, constrained by features.

3.2.1 Features

Every abstract and concrete morpheme in the lexicon carries one or several features that must be unified to succesfully produce a surface form. In the above example, for instance, **j** carries the the feature [**subj=4**], as does **S4**. The morpheme **oo** carries the features [**subj=4, mode=Py, cl=l, prefix=conj**], while **Py** carries the feature [**Mode=Py**]. This can be illustrated as:

ha	j	oo	l	táál	**Py**	**S4**
	[subj=4]	[subj=4]				[subj=4]
		[mode=Py]			[mode=Py]	
		[cl=l]	[cl=l]			
	[prefix=conj]	[prefix=conj]				

The reason, for instance, **oo** has the features it does is that **oo** is allowed as the 4th person marker in the subject slot only if the mode is **Py**, the classifier of the verb is **l**, and it is preceded by a conjunct prefix. Thus, replacing the **X** markers in the above abstract form with *any* of the known prefixes and their accompanying features, the only form that unifies correctly is the one outlined in the table above.[2]

3.3 Phonological rules

Faltz (1998) gives 21 phonological rules in a loose, verbal format, which have been roughly followed in the construction of the phonological mapping. Many of the rules have several components, and thus the number of rules in the computational implementation is slightly larger. Also, in our implementation, most rules refer directly to a templatic slot and phonemes in their conditioning environments, as opposed to Faltz, who formulates many rules based on morphotactic presence. An example of this difference goes as follows—Faltz formulates a rule as:

> Rule Subj-2: The barred-l classifier (that is the **ł** in the cl position) and the plain-l (that is, then **l** in the cl position) disappear when sandwiched in between any subject prefix that ends in **sh** or **s** and the first consonant of the verb.

This has actually been implemented as (abstracting slightly away from the actual formalism):

define Subj2 l|ł → 0 / sh|s _ ^ C

That is to say **ł** and **l** are rewritten as **zero**, if they occur between **sh** or **s** and the classifier-stem boundary-marker and a consonant. This rule will apply regardless of position; the **sh** and **s** do not have to be in subject position for the rule to trigger—the only requirement is that it immediately precede the classifier.

[2]Strictly speaking, unification of two features is not the only operation that is used in the grammar. A feature can also *require* the precence of another feature, or explicitly disallow another feature as well. In the derivation above, **oo** actually *requires* that the feature [subject=4] be present elsewhere—if this were not the case, 4th person **oo** could surface without the accompanying **j**, which would produce the incorrect form.

Most of our rules contain such reformulations strictly in terms of phonology, and we have yet to find counterexamples where knowledge of morphological function is required for accurate predictions.

3.3.1 Capturing free variation

Some free variation must also be taken into account in the grammar. Free variation in Navajo generally falls into two classes: variation in orthographic practice, and phonological variation, which may be dialectal.

Cases of the former are relatively straightforward to capture. For instance, a high-tone **ń** can alternatively be written as **ní**. Such surface variation is easily captured by the addition of a few trivial rules that are marked as optional.

However, in the case of phonological variation, it appears that (in our formalism) some rules are entirely optional, or some rule orderings may vary.[3] A concrete example will serve here:

Table 19: **1. dayíí'aal/2. deiyíí'aal/3. dayiyíí'aal**

outer	plural	object	inner prefix	subject	cl	stem	mode
	da	**y or yy**		**íí**	**Ø**	**'aal**	**P(y)**
disjunct prefixes		*conjunct prefixes*		*stem*			

There seem to be three alternate surface forms of this verb. First, we have object form variation that can be captured by saying the third person object sometimes optionally surfaces as **yy**, with epenthetic vowels in between. If the object form is **y** only, we get form 1. If we start with **yy**, there is a phonological rule that changes the first **y** to an **i**, and a subsequent rule that changes **ai** into **ei** and yields **deiyíí'aal**. Alternate form 3 can be produced by postulating that the **y**→**i**-change is optional, which in turn would bleed the **ai**→**ei**-rule, yielding **dayiyíí'aal**.

Other cases of free variation operate in an analogous fashion, although the example above is probably among the more intricate ones.

[3]When making this statement, we are not committing ourselves to any paradigm of phonology that includes rule ordering. Rather, we make the observation that rule ordering is one descriptive device among many where some variation can be expressed as alternate orderings in application of the rewrite rules.

4 Potential uses

4.1 Lexicography

The kind of parser we've outlined could be useful for a number of applications. The most directly useful target would probably be that of connecting the parser to perform automated lookup in a wide-coverage dictionary, such as Young and Morgan (1987). Under normal circumstances, a user of existing Navajo dictionaries, in order to find the dictionary entry of a verb-form, must know a) the stem in a particular mode, b) the obligatory lexical prefixes. As seen from the discussion above, both of these presuppose knowledge of the word one wants to look up. The inter-mode variation in stem-sets in Navajo is not regular. When one is faced with, for example, a future mode conjugation of a verb, there is really no reliable way of knowing the stem in any other mode. Also, disentangling the obligatory lexical prefixes from optional inflectional material is not trivial. Many lexical prefixes are homophonous to inflectional ones (an example that has been touched upon is the outer lexical prefix **di**, the inner lexical prefix **d**, the distributive plural marker **da**, and the F-mode marker **d**—all of which can give rise to various surface forms **d+vowel**).

Faltz (1998) in fact provides a clarifying example for the difficulty of knowing a stem in any other mode: lookup of a surface form such as **sodoolzin**.

Our parser in this case gives two feasible outputs:

```
[so+0|0+d+oo+l^zin]+Py+S3+O3
[so+0|0+d+0+oo+l^zin]+F+S3+O3
```

That is, the verb *could* be either in the F-mode, or the yi-perfective. Looking up **zin** in the stem listings of YMM92 will give the semantics of the latter. However, in order to find the F-mode meaning, one really needs to know beforehand that the stem in question is in fact listed under **yį́į́d**. Looking up the entry thus obviously presupposes already existing knowledge of Navajo, or powers of clairvoyance. However, note that the parser has in effect narrowed down the possibilities to two legal parses. If it could further consult an electronic dictionary containing stem-sets, the valid dictionary forms could be automatically retrieved, since there is only one form **yį́į́d**, which is in the same stem-set as F-mode **zin**.

4.2 Spellchecking

When used in tandem with a lexicon, extending the current implementation to work as a spell-checker is a feasible sub-project. In fact, well-designed spell-checkers for morphologically complex languages usually presuppose some non-trivial grammar in an efficient format such as the one outlined here.

5 Conclusion & Future Work

We've presented an outline for a strategy of wide-coverage computational parsing of the Navajo verb. The actually implemented parser currently handles verbs conjugated in the three modes **I**, **P**, and **F**. Most phonological alternations in the language have been captured, and the remaining work will focus on encoding missing conjugation patterns and inflectional allomorphs, as well as linking the parser to a lexicon. In the course of development, we have also tested the parser against some 700 manually conjugated forms, and so far the results indicate near-perfect recognition. Some overgeneration is still present, but most of this is directly attributable to the lack of a stem lexicon. We expect such effects to disappear when the grammar is further constrained with knowledge about stems, their obligatory prefixes, and subcategorization information.