

# Tone, Intonation, Stress and Duration in Navajo

*Emily Kidder*  
University of Arizona

**Abstract**

## 1 Introduction

The phonological categories of tone, stress, duration and intonation interact in interesting and complex ways in the world's languages. One reason for this is that they all use the phonetic cues of pitch and duration in different ways in order to be understood as phonologically meaningful. The Navajo language has unique prosodic characteristics that make it particularly valuable for the study of how pitch and duration interact on a phonological level. Navajo is a tonal language, and also has phonemic length, however, the existence of prosodic elements such as intonation and stress have been a matter of debate among scholars (De Jong and McDonough, 1993; McDonough, 1999). Intonation has been assumed to be a universal characteristic, present in tonal and non-tonal languages alike, though evidence to the contrary has been presented (Connell and Ladd, 1990; Laniran, 1992; McDonough, 2002). Stress or accent is similarly thought to be manifested on some level in all languages, even when it is not used contrastively (Hayes, 1995). In this paper I explore the evidence available for whether or not stress and intonation exists in Navajo. I first give an overview of both the phonetics of prosody in section 2, and the phonology of prosody in section 3. In section 4, I go over how these prosodic characteristics surface in Navajo, arguing the hypothesis that a meaningful lack of stress is apparent in Navajo, which is caused by the fact that it contrastively uses the acoustic cues for stress for other purposes. I also hypothesize that this can be considered to be a causal factor in the lack of intonation that has been noted by native speakers.

## 2 The Phonetics of Prosody: Overview

In the study of prosody, or the rhythmic and intonational elements in language, the phonetic properties of pitch and duration play an important role. In this section, an overview of the general properties of prosodic elements is given.

### 2.1 Segmental Duration

The phonetic notion of segmental duration is a complex notion that has a variety of repercussions in many areas of phonetics and phonology. Differences in the duration of segments can be measured along a continuum of length by looking at the physical speech signal. Yet the perception of duration is relative, and can be effected by other phonological phenomena (i.e. vowels tend to lengthen word finally or before a voiced obstruent). Both vowels and consonants can vary in length, and in many languages long vs. short vowels, or geminate vs. non-geminate consonants are different phonemes, and are cued by continuing the speech sound for a relatively shorter or longer space of time. Length crucially distinguishes between phonemes in many languages, an example of which is Finnish. Finnish shows a distinction between both long and short vowels, and long and short consonants, as in the following example:

- (1) a. [tuli] 'fire'  
b. [tuuli] 'wind'  
c. [tulli] 'customs'

(Karlsson, 1999, p.10)

Duration operates in a segmental or syllabic domain, but not in the domain of words or phrases. Duration in the syllabic domain is dealt with in moraic theory, where it plays a key role in the notion of syllable weight. In moraic theory, a mora is a unit of syllable weight: a syllable with one mora is termed a light syllable, while a syllable with more than one mora is heavy (Hayes, 1995). Syllable weight, which is cued by duration, also plays an important role in the assignment of stress, which will be discussed in more detail in Section 3.

## 2.2 Segmental Pitch/Tone

The phonetic phenomenon of pitch, or fundamental frequency, is created by the vibration of the vocal cords during speech. Pitch and intensity are related notions, and higher pitch tends to be accompanied by higher intensity. This is because in order to raise pitch, a higher sub-glottal pressure must occur which causes the vocal cords to vibrate more rapidly. This higher sub-glottal pressure also causes an utterance to be louder (have a higher intensity level). Tone can be defined as the phonologically contrastive use of pitch in the domain of the segment or syllable (Yip, 2002), and when this is present it is termed a tonal language. A tonal language can have a variety of different tonal systems, some simply have a distinction between High tone and Non-High (Low) tone, while others can have many more levels. Some languages also have contour tones, in which the pitch can rise or fall over the course of the segment or syllable. An example of a tonal language with both level and contour tones is Cantonese, which is outlined in (2):

- (2) [yau]
- a. high level       ‘worry’
  - b. high rising     ‘paint’
  - c. mid level       ‘thin’
  - d. low level       ‘again’
  - e. very low level ‘oil’
  - f. low rising       ‘have’

(Yip, 2002, p.2)

Tone can be measured phonetically by looking at the pitch track of an utterance, which is a measure of the fundamental frequency of the speech sound. Though pitch is the primary cue for phonological tone, scholars have noted that high tones also have the correlates of high intensity and longer duration. Intensity, however, has been found to be less perceptible to hearers, and as was discussed above, there are anatomical reasons behind why pitch and intensity may often co-occur. Tonal languages also have a tendency to assign stress to higher toned syllables, as is discussed in De Lacy (2002), which also makes sense because higher tones are likely to have longer durations, and stress is often correlated with heavy syllables. Tone is the manifestation of pitch in the domain of the segment or syllable, and also plays an important role in intonation.

### 2.3 Phrasal Pitch/Intonation

Another place where pitch and duration play important roles is in the domain of intonation. Intonation has been described as being very similar to tone, only realized in the domain of the utterance rather than the domain of the word or syllable. Stress or accent also plays a crucial role in intonation, as the peaks and valleys in an intonation contour are very likely to co-occur with the stressed syllables in a word (Beckman and Pierrehumbert, 1986). Intonation and its internal structure are fundamentally related to the structuring of discourse topics and the conveyance of pragmatic information in human speech. The various factors that make up intonation, primarily pitch contours that vary across a phrase, have been a subject of interest for scholars who wish to uncover what kinds of pragmatic and syntactic meaning they add to speech. The primary method of intonation analysis involves the use of the Tones and Breaks Indices (ToBI) system of transcription, developed by Beckman and Elam (1993). The ToBI transcription system, based on the work of Pierrehumbert (1980), provides a model for transcribing the intonation contour by breaking up the intonational phrase into discrete units of High and Low accents, which are hierarchically structured. This system utilizes both the physical speech signal analyzed by a computer, and the transcriber's perception of the contours. The physical speech signal for intonation is the pitch track, or the measure of the fundamental frequency over the course of the utterance.

## 3 The Phonology of Prosody: An Overview

The phonetic aspects of pitch and duration are manifested in many ways on the phonological level of language. Acoustically, the roles of pitch and duration are different from one language to another, and how languages treat them phonologically is affected by what else is phonologically salient in the language. The following overview looks at how stress and intonation are manifested in language, and how they relate to the phonetic correlates of pitch and duration.

### 3.1 Lexical Stress

Stress is the 'linguistic manifestation of rhythmic structure' (Hayes, 1995, p.1), which is cued both by pitch as well as longer duration and higher inten-

sity (though intensity has been shown to play a minor role in the perception of stress (Hayes, 1995), so will not be dealt with here). Stress is parasitic; it has no particular physical realization of its own, but instead it relies on the phonetic correlates of other phenomena, including pitch and duration.

Stress is often assigned based on the foot structure of a particular language. In metrical theory, words are broken up into smaller units called feet, which normally consist of two syllables, or two moras. Stress is then assigned to either the first syllable in the foot (trochaic), or the second syllable (iambic), and can be assigned hierarchically as primary, secondary, etc. Many languages, such as Latin, have stress systems that are affected by syllable weight also, so regardless of where the default stress would occur in a word, if a heavy syllable is present, it will acquire stress.

### 3.1.1 Relation of Stress to Contrastive Length

Acoustically, stress is perceived differently depending on the language. This is often determined by what else is present in the language that might be perceptually effected by it. Berinstein (1979) shows that languages that have phonemic distinctions in vowel length avoid using duration as a cue for stress, and instead they perceive high pitch as a more salient cue. Hayes (1995, p.7) analyzes this as sensible, because “using duration to mark stress in these languages would obscure the phonemic vowel length contrast.” The necessity of conserving phonemic contrasts in language is an important point, and will play an important role in the following analysis of Navajo.

### 3.1.2 Relation of Stress to Contrastive Tone

Just as Berinstein (1979) showed that languages with phonemic vowel length tend to avoid using duration as a cue for stress, tonal languages can also avoid using pitch as a cue for stress, and instead utilize duration to signal a stressed syllable. An example of this is the language Maimand (Nambiquará family, Brazil), which is both tonal and has a stress assignment system. In this language, the most obvious features of a stressed syllable are length and intensity (Eberhard, 1995, p.28). In languages which utilize the cue of pitch for stress, it would be more difficult to perceive a stressed syllable which has a lower tone, and it would be easier to perceive a stressed syllable when it co-occurs with a higher tone. Therefore, this avoidance of using a particular phonetic characteristic as a cue for stress can be seen as a strategy to avoid

obscuring phonemic contrasts.

## 3.2 Phrasal Intonation

Intonation is a prosodic representation of not only pragmatic and syntactic information, but it also has an internal syntax of its own. Intonational phrases are broken up into hierarchical units, which really on both the phrasal accents (which correlate to stressed syllables in the individual words), and boundary tones (or pitch targets that create a rising or falling contour).

### 3.2.1 Structure of Intonation

The foundation of intonational structure includes the units of Intonational Phrase (IP) and intermediate phrase (ip). Every IP consists of at least one ip and a boundary tone (%), and every ip consists of at least one pitch accent (\*) (where the intonation peak occurs). Accents in a phrase are determined by looking at the pitch track of the speech file, to determine where the pitch levels are high or low (Gussenhoven, 2004). An example of what a transcribed intonation contour looks like is seen in figure 1.

### 3.2.2 Boundary Tones

Intonation contours are produced by an interpolation of the pitch level between tonal accents, with boundary tones at the end of an intonational phrase that determine rising or falling interpolation. In figure 1, you can see that it ends with the symbol H%, which represents a high boundary tone or a rising target at the end of the phrase. Though tonal accents vary, they correspond to the lexical accents present in words. In this way, the structure of intonation is crucially based on both lexical stress (often termed ‘accent’), and boundary tones. So both stress and tone play a key role in the formation of intonation contours. Stress provides the high and low points of the contour, and intonation itself is realized as the interpolation of pitch between the accents and the boundary tones.

## 4 Navajo Prosody

Navajo has contrastive duration, creating a distinction between long and short vowels. Table 1 shows the Navajo vowel inventory.

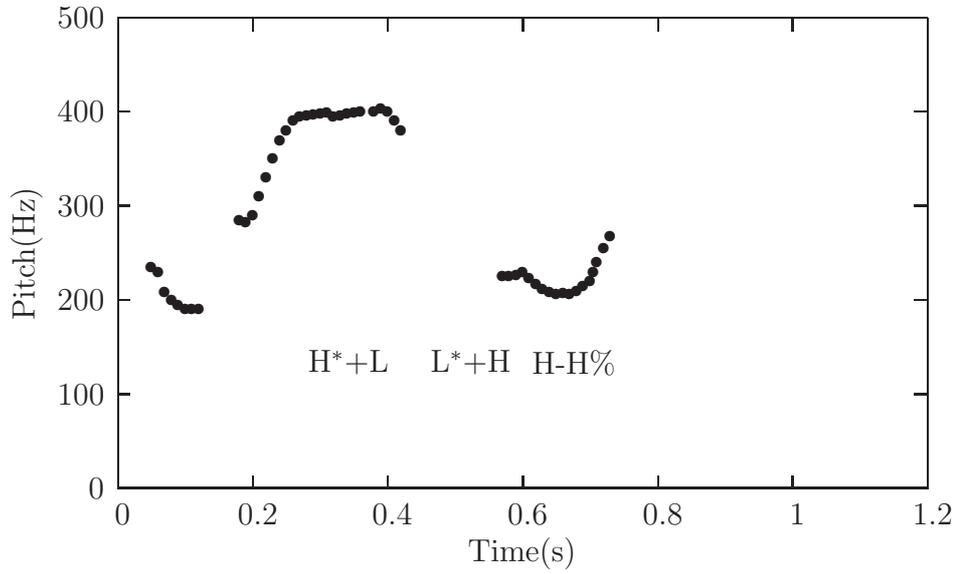


Figure 1: *Intonation contour: 'I've invited John Peck as well' (after Gussenhoven (1999, p.286))*

Short Oral	Long Oral
i	ii
e      o	ee      oo
a	aa

Table 1: *Vowel Length in Navajo (McDonough, 2003)*

Examples of words in Navajo that contain a long and short vowel distinction are:

- (3) a. [bita'] 'in the middle of'  
b. [bitaa'] '(his) father'  
c. [bito'] 'his fluid, water'  
d. [bitoo'] 'his juice'

(McDonough, 2003, p.196)

Tone in Navajo consists of high and low, with low tone considered the default because it is the most common. The following shows some examples of tone in Navajo:

- (4) **Item**            **Tone**   **Gloss**  
ch'ínísmáás    JJJ    'I roll it out' (I)  
ch'íniikááh    HLH   'they go out (3+ in a group) (I)

(McDonough, 2003, p.196)

The notion of whether tone is purely contrastive or not has been called into question by many scholars, but (McDonough, 1999) gives instrumental evidence that each syllable of a word in Navajo has a tonal target. Tone is only contrastive in the stem domain, which occurs at the rightmost edge of the word. In the conjunct, the domain to the left of the stem, tone is often neutralized, and syllables in the conjunct have been argued to be unspecified for tone. This may be because many of the vowels in the conjunct domain are underlyingly epenthetic. The conjunct domain consists of functional, inflectional morphemes that are characterized cross linguistically by having simpler syllable structure and reduced phonemic inventories than content morphemes. Though the conjunct domain does show a lack of contrast in tonal specification, according to McDonough (2002, p.4), "Navajo has the tonal density of a tone language, if not the contrastive tone function of one," due to the tonal targets which are phonetically present on every syllable.

Navajo has been argued to be a language that lacks intonational distinctions between Yes/No questions, focus statements, and declarative utterances (McDonough, 2002), and native speakers also claim that the language lacks intonation. McDonough (2002) gives phonetic evidence that not only do statements, questions, and focus constructions share almost identical pitch tracks, but also these utterances appear to lack boundary tones at all. This

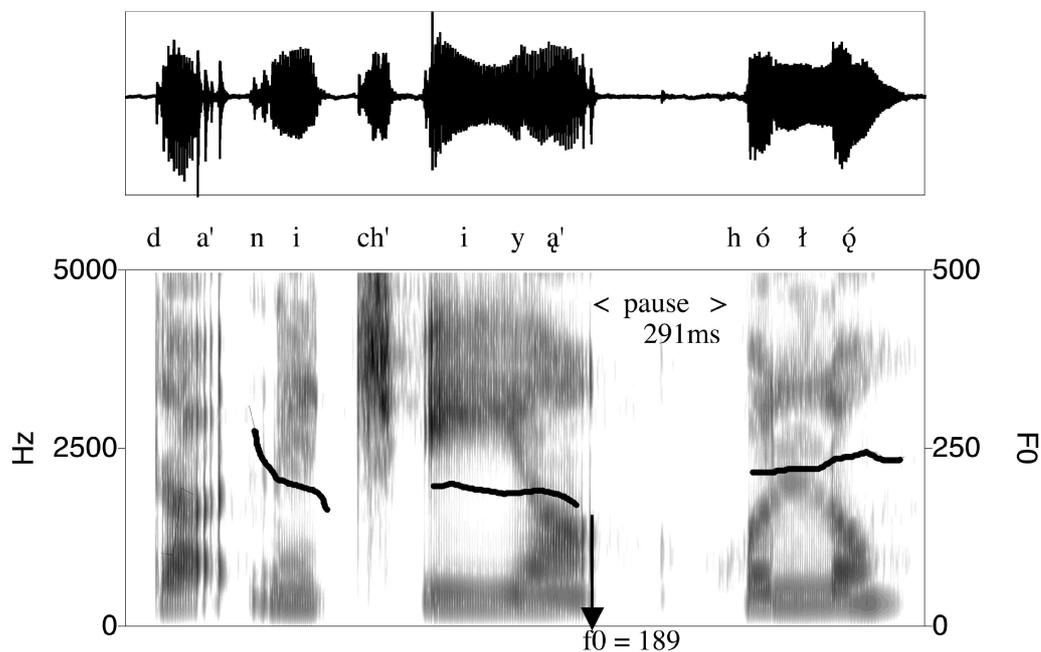


Figure 2: *Pitch track for a Navajo phrase, from McDonough (2002, p.14)*

leads to the argument that Navajo not only does not have contrastive intonation, but that it does not have intonation that compares to that which was thought to be universal in the world's languages. Intonation is basically the interpolation of pitch that occurs between the lexical accents and boundary tones, neither of which seem to be present in Navajo. The pitch tracks only show a relatively level pitch that is distinct for high and low toned syllables. Figure 2 shows the pitch for a phrase in Navajo.

The steady pitch track is apparent in figure 2, with the arrow indicating the morpheme boundary. This is evidence that there is no internal phrasal prosody in Navajo that is marked by an edge or boundary tone at morpheme boundaries.

#### 4.1 Is There Stress in Navajo?

McDonough (1999, 2000b, 2003) has argued that there is some evidence for

the beginnings of a stress-like system that is apparent in Navajo. Though it cannot be said to have a canonical or consistent stress system, she argues that there is a confluence of factors on the Navajo verb stem that seem to make it more prominent. Some cues of stress that are found in other languages with metrical structure include a higher pitch or wider pitch range, a longer duration, a higher amplitude, as well as a greater phonetic inventory in stressed syllables, contrasting with a reduction in pre-stressed syllables. McDonough (2003) describes some aspects of the final syllable in Navajo which have these characteristics. The verb stem is always found at the right edge of the word, and the majority of verb stems are monosyllabic. Verbs tend to be longer, have a wider pitch range and larger phonetic inventory. Also, tonal contrasts are found in the final syllable, but not in the morphemes previous to the stem.

These facts tend to make it appear that Navajo has an incipient stress system in which the final syllable is stressed, however the fact that the final syllable in Navajo is always the verb stem obscures this hypothesis. Verb stems are content words, while the conjunct domain consists of inflectional function morphemes. Content words in languages with metrical structure have similar phonetic cues to stressed syllables. In other words, content words tend to have more prominence in speech than function words, and function words are often reduced. Hayes (1995, p.88) discusses how minimal word requirements typically only apply to content words, while function words and morphemes are often reduced. Therefore it is difficult to say whether or not verb stems in Navajo should be considered to have metrical prominence, because in order to concretely argue this we would need to somehow extricate the cues for stress from the characteristics of content words and morphemes. I argue that in Navajo, the verb stem coincidentally has the same characteristics one would expect to find on a metrically prominent syllable, but that a stress system itself is not present. Verb stems are the most semantically prominent morphemes in the Navajo word, and the fact that they tend to be longer in duration, have more phonetic contrast, and a wider pitch range, can all be attributed to the fact that they are content morphemes.

The phonetic data in Navajo suggests that there are properties that cause verb stems to be distinct in their phonetic prominence, but these are not consistent enough to say that stress exists on the final syllable. Stem syllables are usually longer in duration than conjunct syllables, but only because the stem is more likely to have a coda consonant or long vowel, and when the conjunct morpheme has a coda consonant or long vowel, the stem is not

noticeably longer (McDonough, 2003). Stems are also more likely to have a wider pitch range, but this goes along with the fact that tonal contrasts are typically found on stems. Stems have a larger phonetic inventory, but content or open class words cross-linguistically have a larger phonetic inventory. All of this evidence seems to show concretely that Navajo does not have a system of metrical prominence. One possible explanation for this apparent lack of metrical prominence in Navajo can be attributed to the fact that stress relies on perceptual cues that are already put to use phonemically. Navajo utilizes both pitch and duration contrastively, as was discussed above. In a language such as this, it is likely that assigning stress to a syllable, which is cued by pitch and duration, could have the effect of obscuring the phonemic contrasts in the language. It has been shown that tonal languages avoid using pitch as a cue for prominence, and instead tend to use duration. But in Navajo, duration is also phonemic. Stress, being parasitic in nature, needs to utilize the perceptual cues from other phenomena, but in Navajo, both of the most prominent cues are in use already. Intensity is the only cue available that has been correlated with stress, but as was mentioned above, it has been shown to play a minimal role in phonetic perception. I argue that the Navajo language shows a meaningful lack of stress, which is due to the fact that the building blocks for the perception of metrical prominence are not available.

### 4.2 Lack of Stress and Intonation

As was discussed in section 4, Navajo has been argued to lack any intonational contours that add pragmatic or syntactic information to speech. This cannot be attributed solely to the fact that it is a tonal language, and pitch is phonemically salient on the syllable level, because tonal languages are known to have intonational systems as well (Yip, 2002). However, if Navajo truly lacks a stress system, and intonation contours are built using both lexical accents and boundary tones, there may also be a causal relationship between the lack of stress and the lack of intonation in Navajo. Lexical accents do not exist in the Navajo phrase that would be the foundation for the peaks and valleys of an intonation contour. The lack of this fundamental building block of intonation may explain why Navajo lacks intonation.

## 5 Conclusion

The apparent lack of intonation in Navajo has many theoretical implications, many of which are summarized in McDonough (2002). Information which is normally carried by intonational means is conveyed in other ways in Navajo. The lack of stress in Navajo is also theoretically interesting, because it calls into question whether or not stress can exist in a language which already uses the prosodic cues of duration and pitch in distinctive ways. Acoustically, the parasitic nature of stress means that it does not have a particular acoustic cue of its own that unequivocally signals it, it instead utilizes other acoustic cues. Yet if these cues are already being used in the system to distinguish meaning, they may simply be unavailable to be used as a cue for metrical structure. Scholars of tonal languages have documented that stress is realized mostly as duration, while scholars of languages with contrastive duration document that pitch is the primary cue for stress. In Navajo, which has both, perhaps stress simply has no cues available, and therefore does not surface at all.

This also has possible implications for the lack of intonation in Navajo, because stress or accent is a crucial part of intonational structure. The lack of lexical accents may be a causal factor in the lack of intonational distinctions documented by McDonough (2002).

In summary, Navajo can be convincingly argued to lack a stress system on the word and phrase level, as well as lacking an intonational system that interacts with the syntactic and semantic components. The features on the stem that have been taken to be a form of stress, can all be plausibly attributed to the fact that the stem is the most salient content morpheme in the Navajo phrase, which carries with it all the same acoustic cues that the stressed syllable does. Looking at instrumental data (as in McDonough (2002), it is evident that not only is stress not acoustically present in any way comparable to other metrical languages, but boundary tones which are crucial to the realization of intonation are not present either. I hypothesize that this is due to the fact that the acoustic cues for stress are pitch and duration, and as discussed above, Navajo utilizes contrastive pitch and contrastive duration. Because of this, the realization of stress on a syllable would have the effect of obscuring its meaning. Future phonetic study on languages with both contrastive pitch and contrastive duration, such as Navajo, will no doubt be crucial in understanding the interaction between duration and pitch in the realization of both a metrical and intonational system, and will shed additional light on the complexity of the interface between phonetic and phonological systems.