Teasing Apart Tone and Stress in Athabaskan

Raimundo Cox-Casals

Department of Linguistics, McGill University

LING 411: Structure of an Indigenous Language

Professor James Crippen

Abstract

This paper undertakes an exploration of the relationship between tone and stress, drawing from evidence gathered from three Athabaskan languages. Rather than viewing tone and stress as homologous systems, I challenge the notion that they represent different applications of a singular phonological phenomenon. Through analysis of three understudied languages—Witsuwit'en, Upper Tanana, and Navajo—I posit that the conventional tone continuum inadequately captures languages that incorporate elements from multiple points across the conventional spectrum. The study finds that tone and stress are often conflated because they are indicated with similar acoustic cues but are in fact distinct processes in the phonology. Phonetic evidence, phonological analysis, and diachronic observations are presented to analyze Witsuwit'en's stress system, Upper Tanana's tonal system, and the interaction between tone and stress in Navajo. My goal is to disentangle tone and stress as separate mechanisms within language, as they tend to be misanalyzed and conflated for having pitch in common as a phonetic correlate. Exposing their differences in isolation and interactions in co-occurrence will contribute to a deeper understanding of linguistic typologies and help to avoid overlooking diverging entities in linguistic analyses.

1 Introduction

1.1 Preliminaries

Principally, work on Athabaskan linguistics has been comparative historical research by scholars such as Sapir, Krauss, and Leer. Leer and Kingston, in particular, focused on tonogenesis: the emergence of tone in language. This is a key field for this paper as languages in the Athabaskan-Eyak-Tlingit (AET) family present great variation regarding tonal systems, with only some members having developed tone at all. I will be reviewing past diachronic analyses of tone languages in the AET family (i.e. Upper Tanana) and juxtaposing their developmental characteristics with those of stress languages (i.e. Witsuwit'en).

In addition to a historical comparison, I explore the synchronic phonological behaviours of tone and stress. While tone has a feature-like autosegmental application, stress emerges from inherent rhythmic patterns and culmination. Through an examination of tone distribution, I conclude that tones are targets specified on tone-bearing units (TBUs). On the other hand, by exploring the predictability of stress, I conclude that stress is a prominence-based landmark located within word domains governed by language-specific parameter settings. In this way, I develop the abstract representations of tone and stress to better explain their patterns in surface realizations.

Finally, adopting a phonetic perspective, I investigate the acoustic properties of tone and stress. In AET languages, pitch (referring to the fundamental frequency of the utterance, or f_0) serves as a pivotal indicator of both prosodic devices. The salient nature of a high tone—namely, the spike in pitch—can sound almost identical to a stressed syllable. In other words, a high-low pattern (H-L) sounds similar to a strong-weak pattern ($\acute{\sigma}$-σ). Given these homogenous phonetic cues, supplementary acoustic evidence is needed to distinguish them. I contend that phonologically, tone exhibits specification that is target-like, in that its phonetic realization corresponds to a high versus low pitch excursion. The range of the expected f_0 envelopes of tones can fluctuate if stress is additionally superimposed onto it, as in some languages (i.e. Navajo). Interaction between tone and stress yields noticeable variation in pitch activity.

1.2 Athabaskan Languages

AET (also known as Na-Dene) is a family of languages that spans throughout North America and encompasses approximately 47 languages, though the count varies depending on the classification system employed, and some of them are no longer actively spoken. Three branches of the Athabaskan[1] subgroup are generally acknowledged: Northern Athabaskan, Pacific Coast Athabaskan, and Apachean. Sapir (1915) initially formulated the Na-Dene phylum, which included the Athabaskan languages, Tlingit, and Haida. Lavine (1979) contested Sapir's inclusion of Haida and advocated for an AET genetic linkage which excluded Haida while including Eyak. Subsequent studies, such as Leer (2008), provided further support for this claim. In this paper, all three languages analyzed belong to the Athabaskan subgroup of AET, hence the terms AET and Athabaskan are used interchangeably to denote this family, although the latter term technically excludes Tlingit and Eyak. Figure 1, adapted from Cook & Rice (1989), illustrates the currently established genealogical connections among AET languages and excludes Haida from the family.

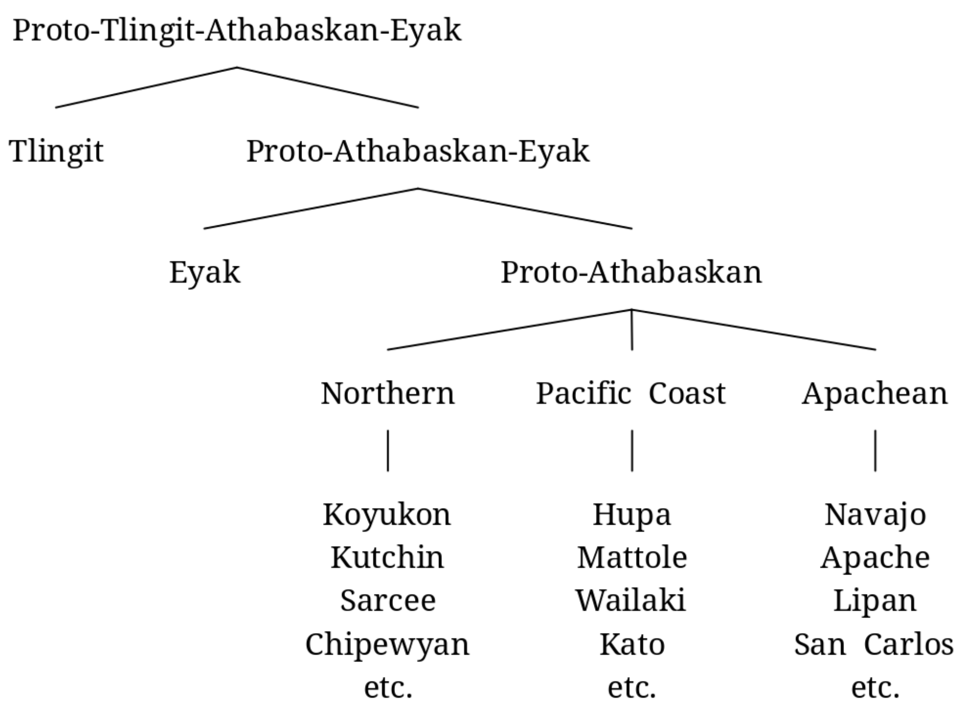

Figure 1: The most widely accepted genealogical family tree of AET (Na-Dene languages).

Upper Tanana (also Nabesna or Nee'aanèegn') belongs to the Northern Athabaskan branch and is spoken in communities situated in eastern Interior Alaska, near the border with the Canadian Yukon. The name refers to the upper reaches of the Tanana River. Despite its cultural richness, the language faces critical endangerment status, with an estimated 30 to 40 fluent speakers remaining (Lovick, 2020).

[1] This paper will use the spelling of *Athabaskan* though other spellings have been used historically to approximate pronunciation such as *Athapaskan, Athabascan,* and *Athapascan.*

Wistuwit'en (also Babine, Witsuwit'en-Babine, or Northern/Western Carrier) is also a member of the Northern Athabaskan branch, spoken along various rivers of west-central British Columbia. Translated as "people of the Wa Dzun Kwuh River" (or the Bulkley River), Witsuwit'en was classified as endangered in 2002 when the estimated speaker count was 185 (Hargus, 2007), predominantly composed of elders.

Navajo (also Navaho, Diné Bizaad) falls within the Apachean branch and is spoken primarily in the Navajo Nation in the southern United States. The etymology of the term Navajo remains uncertain, though it is suggested to be a Spanish loanword denoting a geographical location (Williams, 2009). Unlike the other languages discussed herein, Navajo has a fairly robust speaker base of around 170,000 according to the 2019 American census.

2 Background

2.1 What is Tone?

Tone is commonly regarded as a phonological primitive akin to phonological features such as [±nasal] or [±voice], primarily due to its role in distinguishing meaning in "tonal languages." In essence, phonological representations of tone allow for the distinction between two utterances when the difference between them corresponds to a distinction of signs in the system of the language (Anderson, 1978). We must note the difference between the phonetic specification of a given feature, which may involve a considerable range of potentially distinct values determined by language-particular factors, and the categorical interpretation of a feature, which entails a binary choice of whether a segment possesses the property associated with the feature or not. In the case of tone, f_0 will be the principal phonetic correlate responsible for the acoustic cue intended for contrastive interpretation, albeit subject to variation when applied in conjunction with intonation, accentual patterns, speakers' vocal ranges, etc. (Goldsmith, 1980). These extraneous factors must be abstracted away when searching for a theoretical representation of tone as a linguistic construct. Consequently, tone extraction from the acoustic signal parallels the parsing of linguistically relevant phonetic elements as features while filtering out acoustic interference.

Once phonetic observations are confined to factors systematically and discretely manipulated by a language's grammar, the groundwork is laid for formalizing a phonological theory of tone. With the advent of autosegmental phonology, phonological representation has evolved from a singular sequence of entities to multiple parallel sequences of entities. Tones, however, are conceptualized as constituents of a separate tier of autonomous entities equal to the string of consonants and vowels comprising the melody (Halle & Vergnaud, 1982). This framework posits that segments and tones operate independently, yet are somehow synchronized. The principles listed below delineate the conditions for tone-segment synchronization as outlined by Halle & Vergnaud (1982), taken from Goldsmith's work.

i. All tones must be associated with (at least) one syllabic element.
ii. All syllabic elements must be associated with (at least) one tone.
iii. Association lines do not cross.

Additional parameters regarding rules on the tonal tier may vary depending on the analysis. However, this paper emphasizes the autonomy of the tonal tier from the melody, rather than delving into the derivation process of surface realizations. Under this perspective, tones are stored in the lexicon and are associated with TBUs (syllables or moras) at the time of utterance. Rules concerning tone refrain from long-distance assimilation or dissimilation; they are solely conditioned by adjacent tones, such as the interactions of tone sandhi in Mandarin (Woo, 1969). Woo argues that all tones serve as specifications for "levels" of high (H) and low (L) pitch that form part of lexical entries and undergo phonological rules independently. Thus, though tones are ultimately associated with TBUs and are densely specified for nearly all syllables in a language, their interaction remains detached from segments, although they act like any other feature in distinguishing minimal pairs. Finally, tones uniformly associate with long and short syllables without the distinction observed in stress systems; all syllables are equal TBUs and are paired with tones in a one-to-one fashion (McCawley, 1978).

2.2 What is Stress?

Stress is a system of metrically assigned prominence where a stressed syllable is marked with primary (or secondary) prominence relative to its surrounding syllables. Stress assignment consists of a word-level metrical structure, which locates a head akin to a landmark or anchor guiding one's orientation within its domain. Like tone, stress is a perceptual notion, but unlike tone, stress manifests through various phonetic cues such as excursions in f_0, syllable duration, intensity/loudness, and vowel quality (Hargus, 2007). While these acoustic correlates contribute to the perception of prominence, they are not all always present in the signal of a given token. There are two main criteria definitional to a stress accent system (Hyman, 2006):

i. OBLIGATORINESS: every lexical word has *at least* one syllable marked for the highest degree of metrical prominence (primary stress).

ii. CULMINATIVITY: every lexical word has *at most* one syllable marked for the highest degree of metrical prominence.

Obligatoriness is the more important of the two, constituting an absolute universal in stress systems for there to be an obligatorily headed metrical constituent built at the word level. We encapsulate these two properties under the Optimality Theory constraint HEAD(PWd), asserting that every phonological word must have a unique head and therefore one solitary accent (McCarthy, 2002). This constraint targets syllables and not other constituents like moras. Additional properties of a stress system include 1) privativity: stress is either present or absent on a given syllable, 2) demarcation: stress fixed on initial, final, etc. syllables can indicate where word boundaries are, and 3) rhythmicity: echo stresses (secondary prominence) frequently occur on every other syllable.

While suprasegmentals like tone (which corresponds to features) or length (which corresponds to timing slots or moras) are lacking in universality, stress exhibits promising universality. Hyman (2008) argues that most studies asserting stress absence in languages are "not looking hard enough." It is easy to see how this might happen, particularly in languages with concurrent tone, where stress identification becomes challenging. Even in extreme cases of stress absence, rhythmic groupings resembling Morse code persist, such as that observed in Pohnpei (Rehg & Sohl, 1981). Here, syllables are organized into word-based groups (potentially feet) devoid of obligatory prominence, thus violating HEAD(PWd), yet still showcasing some rhythmic properties. By capturing the predictable rhythmic and structural distribution of stress on a metrical grid (Hayes, 1986), we envision the syntagmatic assignment of prominence and its reiterations along the rest of the word like a constant beating drum.

2.3 Summary

Tones are featural, distinctive, and paradigmatic; in a tone system, one identifies the tones of each morpheme as they are specified as sequences of pitches in the lexical entry of each and are retained throughout derivation except when modified by phonological rules. Rules affecting pitch are conditioned by immediately adjacent tones and have no long-distance interaction. The relevant units can be syllables or moras of any segmental composition with no short/long distinction.

Stress is structural, contrastive, and syntagmatic; in a stress system, one locates the peak in prominence within each lexical word through metrical computation, accentuating points with greater salience amidst their surroundings. Primary stress assignment may have long-distance effects of rippling secondary stresses across alternating syllables. Stress, exclusive to syllables, discriminates light versus heavy types based on their segmental material, asymmetrically attracting stress.

Tone and stress share a common phonetic tool in f_0 for their phonetic realization. The following sections will discuss how U. Tanana helps to distinguish tone (§4), how Witsuwit'en helps distinguish stress (§5), and how Navajo makes use of both systems (§6). These findings demonstrate the distinct phonological behaviours of tone and stress despite their near-identical phonetic outputs.

3 Tone in Upper Tanana

3.1 Tone Incorporation

In Lovick's grammar of U. Tanana (2020), we see that the language distinguishes between unmarked, low-marked, and high-marked syllables. Unlike its closely related neighbour Tanacross, U. Tanana has no contour tones at all, even as a result of suffixation. To prevent ambiguity, I will take "marked" to mean cross-linguistically rare as per OT, and "unspecified/underspecified" shall signify a lack of tonal specification; a departure from Lovick's terminology. In the following wordlist, many syllables in U. Tanana lack tonal specification, articulated instead at the average frequency of their locale in the utterance.

a.	gah /kah/	[kah]	'rabbit'
b.	neetsǫǫ /neːtsʰõː/	[neːtsʰõː]	'our grandmother'
c.	ihhaał /ihhaːɬ/	[ihhaːɬ]	'I'm walking'
d.	neljit /neltʃitʰ/	[neltʃitʰ]	'she is scared'

Figure 2: Syllables underspecified for tone.

L-tones in U. Tanana are lexical since minimal pairs are distinguished by their presence versus their absence, comparable with the presence/absence of a phoneme or feature, as illustrated below.

	Low-marked	*Unspecified*
a.	na- /nà-/ 'CONT'	na- /na-/ 'IT'
b.	ih- /ìh-/ 'NEG.PFV:1SG.S:Ø'	ih- /ih-/ 'AA.PFV:1SG.S:Ø-'
c.	deeł /tèːɬ/ 'PL go IPV'	deeł /teːɬ/ 'PL go PFV'
d.	nän' /nʌ̀nʔ/ 'land'	nän' /nʌnʔ/ 'you SG'

Figure 3: Minimal pairs for L-tone.

H-tones in U. Tanana are phonetically salient, exhibiting signs of phonemic status, notably exemplified by its incorporation in the negative suffix -v́.

	Affirmative		*Negative*		
a.	shyah	/ʃʲah/	shyay	/ʃʲáj/	'SG go PFV' (N, T)
b.	'į̨h	/ʔĩh/	'ąy	/ʔã́j/	'see IPV' (N, T)
c.	deeł	/teːɬ/	deel	/téːl/	'PL go IPV' (T)
d.	deeł	/tèːɬ/	deel	/téːl/	'PL go IPV' (N)
e.	got	/kòtʰ/	godn	/kót/	'punch PFV' (N)

Figure 4: Minimal pairs for H-tone.

These examples from Lovick (2020) illustrate the distinctiveness of H- and L-tones in U. Tanana and how they are identified on each TBU. The functional load of lexical tone like that of U. Tanana resonates with phonological features, meaning even though it operates on its tier independent from segmental information, its overall contribution to the utterance is comparable. Cross-linguistic research has shown that tonal information can exist in the lexicon without segmental information just as how segmental information can exist without tonal information (see Broadwell & Zhang, 2000 for an example from Zapotec), exemplified by the tonal negative morphemes in Figure 4. To further support the featural properties of tone I will look at Lovick's instrumental study on the phonetic realization of tones in U. Tanana.

3.2 Phonetic Evidence for Underspecification

In the previous section, I mentioned that U. Tanana has underspecification concerning tones, a property reminiscent of features. This section will portray how this phenomenon works analogously with feature geometry, as elucidated by Lovick's findings. In phonological studies, the representation of segments may initially be underspecified but becomes fully specified. Various fill-in rule processes are proposed to ensure that all features attain values across segments in output forms. Keating (1988) argues that underspecified values persist in phonetic forms, where phonetic data reflect the presence or absence of feature values in surface forms.

While most Russian consonants contrast as either palatalized or not, /x/ operates differently. Instead, it relies on context for its value of [±back]. When a segment acquires features from a neighbour, it displays the relevant property of that feature to the same extent as the neighbour does. Thus, we anticipate /x/ before back vowels and /ç/ before front vowels. The fricative /x/ allows examination of the time course of assimilation effects. Below, (a.) shows the alleged generalized rule, (b.) shows the fronted form, and (c.) shows the unfronted form. /X/ represents the dorsal fricative unspecified for [±back].

a. [+cont, +cons, –son, DOR] → [αback] / __ [αback]

b. /X/ → [ç] / __ i

c. /X/ → [x] / __a

Keating's phonetic data contradicts this analysis when /X/ is placed in mirror-image contexts before and after /i/ as well as between two /a/ vowels.

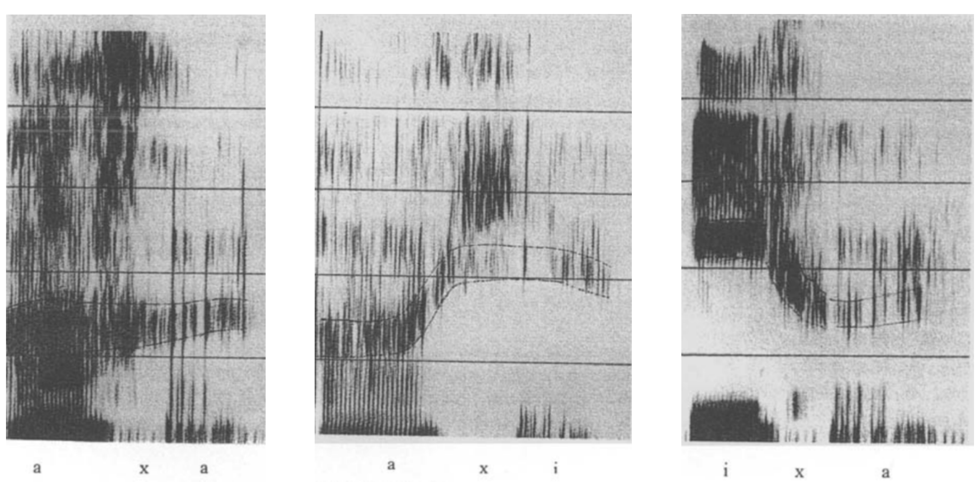

Figure 5: Fronting of Russian /x/ before /i/. Left, steady low f_2 frequency component in fricative. Middle, before /i/, steady higher f_2 in fricative. Right, after /i/ transitional f_2 in fricative between the two

qualities of the other two examples.

For the token /ixa/, the second formant, reflective of backness, shows a continuous transition from front to back. By contrast, for /axi/, the fricative is extremely fronted throughout its duration. This provides support for the idea that the fricative is [-back] for its entire duration rather than having only a transitional front quality. The /x/ preceding /i/ acquires a value for [-back] by spreading from /i/ whereas the other instances of/x/ remain unspecified for [±back], thereby displaying interpolatory effects.

This leaves three possible output representations for any given segment that may emerge from the phonology, each with different acoustic reflexes. These schematic representations are illustrated below using VCV sequences.

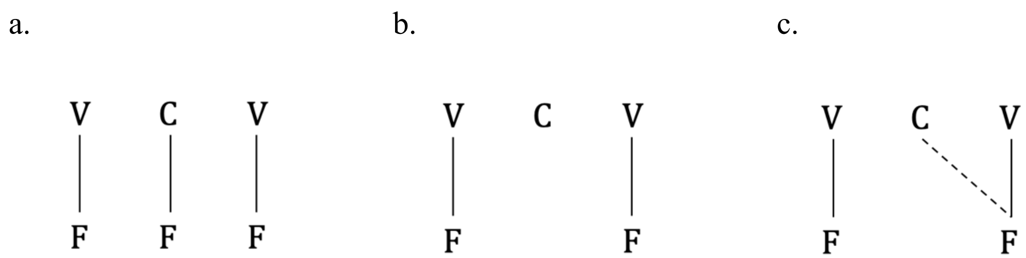

Figure 6: VCV sequences with the three possible representations for specification and underspecification that may be outputted by the phonology.

In (a.), each segment has a value for feature F, resulting in distinct phonetic qualities for each, with no vowel-to-vowel effect in either direction due to the consonant's feature value blocking such interactions. In (b.), a transition occurs between the two vowel targets and the consonant lacks its own phonetic quality, leading to a gradual shift. In (c.), the consonant acquires a value for F from V_2 via a spreading rule, yielding no coarticulation effect of V_1 on V_2 because the consonant's new value for F again blocks all interactions. However, there is an effect of V_2 on V_1 because the consonant's new feature will affect V_1. Lovick's analysis displays this precise behaviour for tone in U. Tanana.

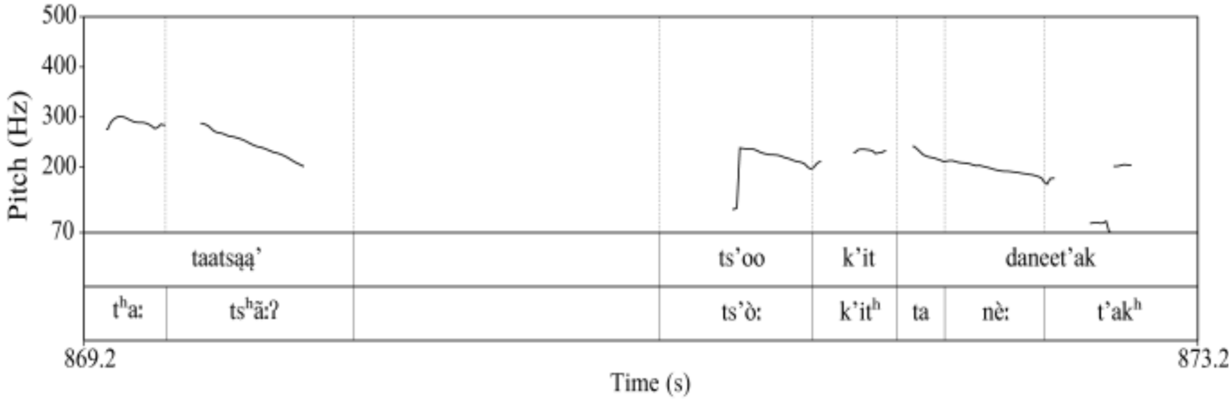

Figure 7: Pitch tracks for L and unspecified tones, taken from Lovick (2020).

The data presented demonstrates that syllables[2] unspecified for tone adopt the same pitch range as their neighbours when positioned between two mirroring L-tones. The L-tones of the words /ts’òː/ and /nèː/ enclose the unspecified words /k’itʰ/ and /ta/, giving a steady f_0 contained inside of the L-tone envelope. This occurs because T_1 and T_2 are both L and necessitate no gradual transition between their frequency targets. Therefore, since unspecified syllables have no pitch target, f_0 seamlessly bypasses them on its journey from T_1 to T_2, yielding a constant pitch throughout.

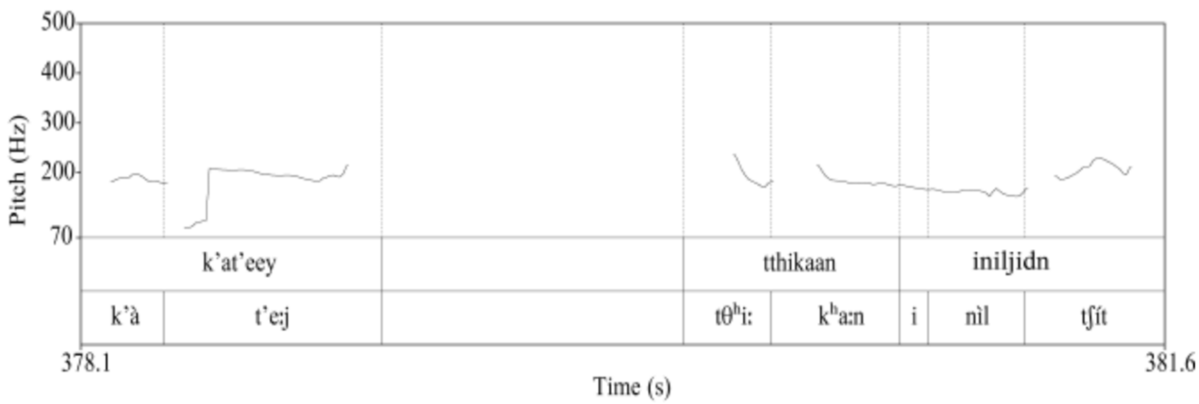

Figure 8: Pitch tracks for H and unspecified tones, taken from Lovick (2020).

Comparing the pitch change between an unspecified, an L-tone, and an H-tone, we observe the target-like properties of tones. From the unspecified tokens /tθʰiː/, /kʰaːn/, and /i/, f_0 descends to hit the L target on /nìl/, then sharply rises to hit the H target on /tʃít/. These examples indicate that for the Russian case, where a fricative unspecified for backness does not account for a point in the

[2] Syllables, moras, vowels, etc. can associate with tones, but I have chosen the syllable as the TBU here in accordance with the data from U. Tanana though other languages may differ.

trajectory between back and front place, it behaves transitionally. In the case of U. Tanana, syllables unspecified for tone do not establish a target and are thus transparent to the pitch trajectory between two specified tones. See the following diagram for the three kinds of tonal outputs that result in different pitch reflexes.

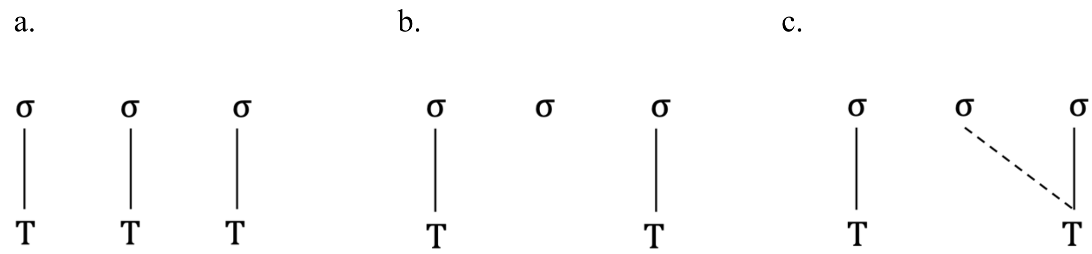

Figure 8: Syllable sequences with the three possible representations for specification and underspecification of tone that may be outputted by the phonology.

In (a.), each syllable has a value for tone T, so each will have its own pitch target. In (b.), there will be a transition between the two pitch targets and the medial token lacks its specification for tone, resulting in a gradual shift in frequency. In (c.), the medial token acquires its surface tone value from a neighbour by a spreading rule.[3] Toneless syllables in U. Tanana correspond to option (b.) where they are fully unspecified for T, functioning as a sort of all purpose wildcard that takes the pitch of its adjacent tones. With this insight, tones are revealed to behave exactly like features in that they can be specified/underspecified, spread through rules, and interact with their neighbours once outputted. Tone, while manifesting as pitch cues and forming part of a language's prosody like stress, appears to share a closer connection with features and segments which are stored in the lexicon and undergo phonological derivation.

3.3 Tonogenesis

An alternative approach to isolating tone from stress is a diachronic perspective that looks at the birth and evolution of tone and how it is set apart from other features. From the data available we see that some languages in the family have no tone at all, some have a simple two-tone system, and others are reported to have up to four tones. Intriguingly, H-tones in some languages correspond to L-tones in others, and vice versa. This inconsistency in the existence and attributes of tone can be explained through a historical lens and proto-language reconstruction. I argue that the evolution

[3] Tones can spread in either direction depending on the language, not just right-to-left as shown here.

of tone is parallel to the path of historical change taken by any segment, suggesting that tone, just like a segment, comprises featural information.

The rise of tone, known within the field as "tonogenesis," likely happened multiple times in the AET family. Typically, the process involves the suprasegmentalization of segmental information, as it transitions from the realm of the melody into the tonal tier. A noteworthy contemporary example of tonogenesis is unfolding in the Seoul dialect of modern Korean (Kim, 2012). Here, oral stop consonants exhibit a three-way contrast between voiceless, aspirated, and tense/ejective. The feature [+spread glottis] of aspirated stops underwent strengthening, affecting any following vocalic segments to produce an added "exhale" effect. The intensified release exerted an upward pitch modulation on the succeeding vowel, now being reinterpreted as an H-tone, whereas the originally unaspirated stop assumes an L-tone to maintain contrast. Eventually, the [±SG] contrast between aspirated and voiceless stops neutralizes syllable-initially, leaving behind a purely tonal contrast.

Similarly, languages in the AET family,[4] acquired tones through the assimilation of a [+glottal] feature (also termed [+constricted glottis]). Vowels assimilated with final glottalic consonants because of a [+glottal] spread to which vowels respond by constricting. The original contrast between glottalic and non-glottalic final consonants from Proto

Athabaskan was retained in non-tonal languages while being lost in tonal languages (Kingston, 2005). From the constriction in some languages, the vowels tightened the vocal folds and prompted a higher f_0, later reinterpreted as an H-tone. In other languages, the increased tension and anticipation of the final glottal induced creaky voice, where the vocal folds come into contact in a way that is perceived as a low pitch. Whether constriction became an H- or L-tone designates if a language is high-marked or low-marked. In U. Tanana, glottalization gave rise to the L-tone, making it low-marked (Leer, 1999). The asymmetries in the development of constriction from Proto-Athabaskan are summarized by Rice & Hargus (2005):

[4] Tonogenesis probably happened multiple times in AETat different points in time, but for simplification of discussion, I will group the birth of tone from glottalization in the Athabaskan (Dene) languages.

Pre-Proto-Ath	*Proto-Ath*	*High-marked*	*Low-marked*
VV	VV	V̀V̀	V́V́
VVʔ	Vʔ	V̀ʔ	Véʔ
vR	vR	v̀R	v́R
vR’	v’R’	v́R’	v̀R’
VVR	VVR	V̀VR	V́VR
VVR’	VVR’	V́VR’	V̀VR’
vT	vT	v̀T	v́T
vT’	v’T’	v́T	v̀T
VVT-R	VVT	V̀VT	V́VT
VVT(-T/S)	VVS	V̀VS	V́VS
VVT’-R	VVT’	V̀VT	V́VT
VVT’(-T/S)	VV’S	V́VS	V̀VS
VV’T(’)-R	V’T(’)	V́VT	V̀VT
VV’T(’)(-T/S)	VV’S	V́VS	V̀VS

Figure 9: The evolution of tone from Pre-Proto-Athabaskan to Proto-Athabaskan and to high marked and low-marked languages. T = stop/affricate, S = fricative, R = sonorant, V = short full vowel, VV = long full vowel, v = reduced vowel.

This generalization reveals the emergence of tone as a divergence from the laryngeal feature [+glottal], splitting off and becoming an individual distinctive entity. By drawing parallels with the lenition of Spanish intervocalic stops, a resemblance arises between tonogenesis and spirantization, both involving the spread of a feature to surrounding segments, precipitating a historical shift. Below are examples from Penny (2002).

Latin	*Spanish*	*Example*	*Gloss*
/-pp-/	/p/	kuppa > kopa	‘wine glass’
/-p-/	/b/(= β)	ku:pa > kuba	‘wine vat’
/-b-/	/β/	kibu > seβo	‘food’ > ‘bait’
/-tt-/	/t/	gutta > gota	‘drop’
/-t-/	/d/(= ð)	kate:na > kadena	‘chain’
/-d-/	/Ø/	sede:re > se.er	‘to sit, be’

/-kk-/	/k/	sikku > seko	‘dry’
/-k-/	/g/(= ɣ)	sekuru > seguro	‘safe’
/-g-/	/Ø/	lega:le: > le.al	‘loyal’

All geminates simplify to singletons, intervocalic voiceless stops acquire a [+voice] feature from the neighbouring vowels, and intervocalic voiced stops acquire a [+continuant] feature becoming fricatives. Tonogenesis happens analogously with lenition as they both derive from historical feature spreading. However, we do not identify a [±tone] feature because of the independence that tone gains from the melody upon suprasegmentalization and its separate behaviour regarding rules and derivation. Nonetheless, tone is born out of features and is featural at its core. In the same way as some languages do not contrast [±voice] or [±nasal], the lack of contrasting tones in languages can be attributed to the featural essence of tone. I conclude that tonogenesis mirrors segmental change in its reliance on feature interaction, elucidating the featural behaviour of tone, albeit distinct from features themselves, owing to the newfound autonomy they possess.

4 Stress in Witsuwit’en

4.1 Morphology and Weight Sensitivity

In past analyses, Witsuwit’en has been erroneously categorized as tonal (Krauss, 2005). However, after closer inspection, it becomes evident that the positioning of “H-tones” within a word is dictated by a combination of morphological factors (stem/root vs. prefix syllable), syllable weight (closed vs. open syllable), position of syllable within a word, and vowel category (full vs. reduced vowel). This predictability points towards the language adhering to a system of prominence marking, not simply a tonal one. Early transcriptions of data, such as Barbeau’s, focused on pitch rather than tone, reflecting the fact that high pitch is one of the ways Witsuwit’en manifests stress (Hargus, 2007). Additionally, in words containing three or more syllables and prefix vowels that are all short, stress falls on the penult and preceding alternate syllables (Story, 1984). Though Story does not designate a single stress within the word as the strongest, the rhythmic alternation of echo stresses is compelling evidence for a primary and secondary stress system. This section will discuss the parameters governing stress assignment, exposing its predictability vis-à-vis tone.

Stress is attracted to stem syllables in Witsuwit’en as in many other Athabaskan languages (Rice & Hargus, 2005). Each stem receives some degree of stress, with compound words allocating primary stress to the leftmost stem and secondary stress to the subsequent stems.

a. [c'əɬtə́y] ‘gun’ [sə́s-c'əɬtə̀y] ‘bear gun’

b. [ɬəyə́l] ‘goat’ [ɬəyə́l-tsə̀y] ‘goat meat’

The notion that stress gravitates towards stems suggests that phonology is informed by morphosyntax, indicating that prominence is not purely phonologically decided (Crippen et al. 2023).

When all other factors are equal, stress is attracted to word-initial syllables. This is best illustrated with words containing vowels of the same quality and syllables of the same weight.

a. [líli] ‘bed’

b. [déde] ‘sickness’

c. [dóso] ‘gunny sack’

d. [búzu] ‘hello’ (handshake)

In trisyllabic words containing full vowels, secondary stress is apparent on alternating syllables from the left edge of the word.

a. [níninɢì] ‘it’s dired out’

b. [nédibìs] ‘nighthawk’

c. [déwesyèl] ‘I didn’t walk inside’

d. [wéc’ezùʔ] ‘it’s not goodd

Stress displays a preference for long vowels and full vowels, with primary stress falling on the long full vowel even when it is not the leftmost full vowel in the word. This gives the ranking VV > V.

a. [də́ɬ-neyéːldə̀c] ‘they are talking to themselves again’

cf. [də́ɬ-néyeldə̀c] ‘he’s talking to himself again’

b. [nqaʔáːninzə̀n] ‘they want you’

cf. [nqáʔninzə̀n] ‘he wants you’

c. [newec'óːɬyìts] ‘they shouldn’t rest’

cf. [néwec'oɬyìts] ‘she shouldn’t rest’

d. [ts'ɛnéːbətaɬdzə̀ɬ] ‘he’s going to wake them up’

cf. [ts'ɛ́netadzə̀ɬ] ‘he’s going to wake up’

Words containing both full and reduced vowels have stress on the leftmost syllable with a full vowel. Secondary stresses occur on sequences of syllables with reduced vowels.

a. [wédəzə̀gəlkʷə̀ts] 'I'm not coughing'

b. [ʔə̀ndəníst gèɣ] 'I made a mistake (speaking)'

Lastly, stress is attracted to closed syllables. If the first syllable is light and the second is heavy, stress skips the initial syllable and lands on the following heavy syllable.

a. [nə.gə́l.wə̀s] 'I'm hot'

b. [nə.c'ə́z.nə.qə̀y] 'We're sewing'

In totality, stress assignment in Wistuwit'en is quite complex and relies on the coming together of competing parameters in what may initially appear as a "random" distribution of high pitches that mimic a tonal pattern. Upon more careful inspection, we can retrieve the set of parameters that govern stress and predict its assignment.

4.2 "Quality" Sensitivity

On top of weight sensitivity, where long and short vowels are contrasted for having one and two moras respectively, some linguists have suggested a weight scale (Kenstowicz, 1996). We can observe in languages like Witsuwit'en that stress is indeed sensitive to vowel length as expected, but seemingly vowel quality as well. The introduction of a scale allows for the categorization of how strongly different syllables draw stress without the binarity of a monomoraic versus bimoraic analysis. Below is a summary of Hargus' (2007) findings.

Parameter	*Stress is attracted to:*
Morphological category	Stems (vs. affixes)
Vowel quality	Peripheral: /a, e, i, o, u/ (vs. central: /ə/)
Vowel quantity	Long: μμ (vs. short: μ)
Position of syllable in word	Left edge of word (vs. word medial)
Syllable type	Closed CVC (vs. open CV) syllables

Figure 10: A descriptive account for the parameter settings for stress assignment in Witsuwit'en.

Hargus (2001) argues that Witsuwit'en stress, and by extension stress in other languages, prefers peripheral vowels over central vowels. This is similar to how Kwak'wala codas are only moraic if they are non-glottalized sonorants (Zec, 1989). Remarkably, the phonology can check not just the prosodic environment but also the feature information of a segment to determine whether it is stress-attractive or not. Hargus (2007) however, reconsiders this notion, proposing an alternative analysis wherein short central vowels are phonetically shorter than short full vowels, thus attracting stress to a lesser extent. This is motivated because short vowels make worse peaks than long vowels. She concludes that quality sensitivity is simply an extension of length sensitivity, proposing the following ranking for syllables.

VV > V(C) > əC > ə

Long full vowels are the heaviest, followed by full vowels with an optional coda, then schwa (central vowel) with a coda, and finally open syllables with schwa as the lightest possibility. She explains this with the ranking of two constraints: *PEAK/v̆ >> * PEAK/V. This means that the avoidance of stressing "extra" short vowels outweighs the avoidance of stressing short vowels. Although her revised analysis aligns with the observed surface forms, it predicts a three-way phonological contrast, diverging from moraic theory. I suggest a traditional weight distinction between short vowels (V) corresponding to monomoraic syllables (σ_{μ}), and long vowels (VV) and closed syllables ((C)VC) corresponding to bimoraic syllables ($\sigma_{\mu\mu}$). This light/heavy division works in conjunction with an OT constraint, * PEAK(ə) that is violated by stress on reduced vowels, schwa in the context of Witsuwit'en.

The markedness of reduced vowels in stress systems motivates an approach utilizing *PEAK(ə) that is separate from binary weight distinctions but supplements weight distinctions during stress computation to output the observed patterns.

Thus, it appears stress is not contingent on vowel quality, but rather on a combination of sensitivity to syllable weight and the satisfaction of constraints. Nonetheless, the set of parameter settings of Witsuwit'en employed to locate prominence peaks is the mark of a true stress system. While these stresses may acoustically resemble H-tones, their distribution differentiates them as stresses.

4.3 The Nature of Stress Assignment

As previously noted, an H-L tone pattern is phonetically not very different from a σ́-σ pattern (Hyman, 2006), but it distinguishes itself through the rhythmic distribution of stress. While OCP constraints control underlying tones like they would any feature, something like stress adjacency is constrained by *CLASH, for example (Alber, 2005). To show how stress is regulated by a different collection of constraints than tone, I propose the following constraints, drawing from the patterns observed in Hargus (2001).

STRESS-STEM >> * PEAK(ə) >> STRESS-TO-WEIGHT >> ALIGNFOOTLEFT

Note that the precise ranking of these constraints may need to be refined for comprehensive data capture; my immediate aim is to identify the broad mechanisms in charge of stress assignment. STRESS-STEM refers to the priority stems receive when giving prominence. * PEAK(ə), as discussed above, serves as a markedness constraint that is violated by stressed reduced vowels. STRESS-TO-WEIGHT consolidates long vowels and closed syllables into heavy syllables, assigning them stress. Finally, ALIGNFOOTLEFT expresses the preference for assigning stress to word-initial syllables, likely congruous with a left-to-right directionality in foot-parsing, which competes with STRESS-STEM, given that stems are most often word-final. The nature of stress assignment operates under unrelated constraints to tonal ones, such as OCP, *FLOAT, etc. (Meyers, 1997), presenting challenges in integrating both devices under a unified system with affiliated rules. Moreover, stress seems to emerge from foot structure, which, in turn, arises from the inherent rhythmic organization of speech into binary alternating beats. Rather than forming part of lexical entries and participating in derivation, stress is rooted in the necessity of words to culminate and the disposition to locate landmarks within words, helping listeners to detect prosodic chunks and parse language input more efficiently. All these rudimentary qualities are intrinsic to stress, but by contrast, cannot be attributed to tone. Though tones can serve as boundary markers, their constraints and rhythmic properties differ significantly.

5 Tone-Stress Interaction in Navajo

5.1 Marking Prominence

Early grammars of Navajo make no mention of tone (Reichard, 1951; Morice, 1932) and some merely acknowledge a basic H-L contrast (Young and Morgan, 1987). However, analyzing the phonetic patterns of more complex words such as Navajo verbs, we can observe pitch effects over longer utterances and conclude that tone and stress are both active. Previous sections discussed tone and stress from a phonological perspective, conceptualizing them as abstract systems realized as detectable surface phenomena and attempting to explain how abstractions become real. This section takes a different path and focuses on concrete phenomena to derive their abstract representation from their tangible properties.

Declination (the gradual dropping in f_0 throughout an utterance) is a universal linguistic phenomenon. Visualized, the voicing bar of a spectrogram of an utterance tilts downward along the speech sample. Conventional phonological tone analyses strip the signal of intonation interventions and rescale pitch contours to horizontally flatten the average pitch throughout the utterance. This is a practice conducive to departing from pitch as a concrete entity and towards tone as an abstraction with polar values. However, eliminating declination and intonation also erases any traces of stress interactions. Retaining declination as a tilted baseline, tones and stresses are depicted as pitch deflections away from the average f_0 at any given moment. Taking declination into consideration makes McDonough's (1999) finding about phrase-final stem syllables stand out. She found that the f_0 range from H- to L-tones expands noticeably during stem syllables, which are known to attract prominence.

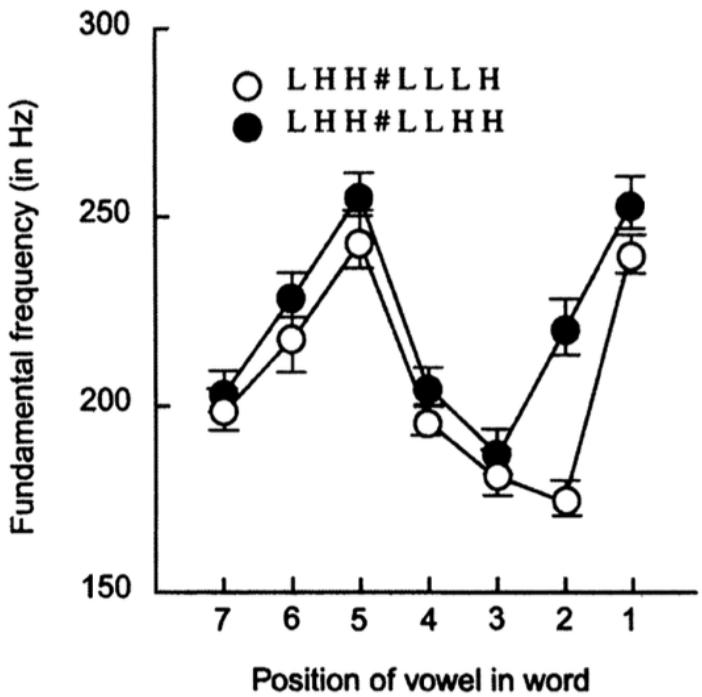

Figure 11: Average range of tones in words with LHH#LLLH and LHH#LLHH tonal contours.

Normally with declination, the envelopes for the H-tone band and the L-tone band would contract as the pitch depresses towards the end of the utterance, but in Navajo, the range suddenly widens at the end of the word where the verb stem syllable is located. Theoretically, tone is not concerned with how far its pitch excursions depart from the norm as long as H- and L-tones are higher or lower relative to each other. We expect a purely tonal contrast to have the amplitude of its range squeezed at the end of an utterance, but instead, a substantial widening of the range makes H-tones higher and L-tones lower. This is strong evidence for the involvement of another factor affecting the pitch. Since stress is concerned with greater versus lesser excursions (primary vs. secondary stresses) it is a viable candidate for causing heightened highs and lowered lows that are not inherently specified on the tonal tier.

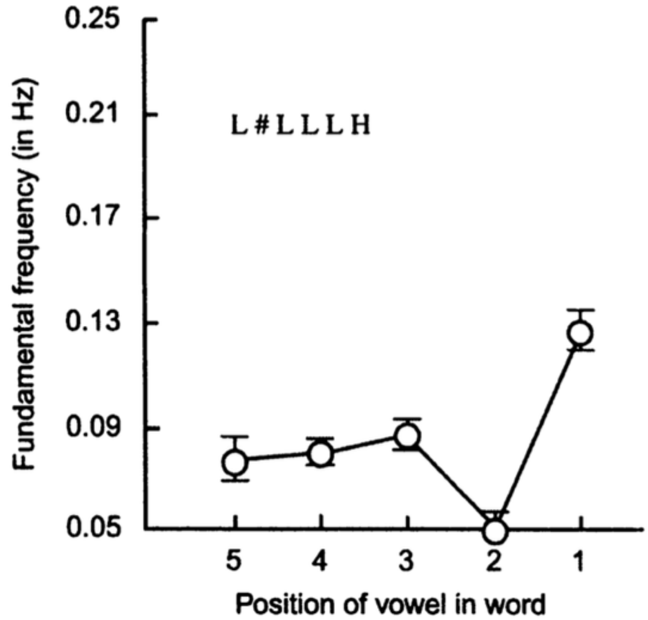

Figure 12: Average range of tones in words with a L#LLLH tonal contour.

In this example, L-tones preceding the final H-tone dip lower than the preceding L tones, defying tonal polarity. This is explained by anti-prominence constraints which dampen any salience as the main stress approaches, priming the stem syllable for prominence (Crippen et al. 2023). The prominent qualities of pre-stem syllables are constrained to intensify the prominence of the stem syllable and maximize the stress relative to its predecessors.

These acoustic data demonstrate that tone alone cannot be responsible for the observed pitch patterns. Unlike languages like Mandarin, where tone sandhi is triggered at any point of the sentence but has phonetically varying realizations due to declination (Goad & Qu, 2011), Navajo tonal indications appear to mostly adhere to declination but bifurcate substantially under stress influence, especially in regions expected to undergo compression if tone were exclusively considered.

5.2 Looking for Foot Structure

The Navajo language shows signs suggestive of both stress and tone, although discerning concrete evidence for their simultaneous existence as separate devices is challenging. When testing for stress, searching for signs of rhythmicity and foot structure is pivotal since they are the most salient surface characteristics that tell stress apart from tone. McDonough (1999) proposes a bipartite analysis where words are made up of three domains based on their prosodic properties. She focuses on verbal morphemes and divides them into a disjunct (D-domain), a conjunct (I-domain), and a stem (V-domain). The stem is normally the final syllable of a verb, preceding it is the conjunct consisting of inflectional morphemes, and at the left edge is the disjunct which consists of adverbial and clitic-like content. Importantly, the disjunct is optional and the conjunct is generally optional, except for one morpheme which McDonough calls the mode-subject morph, which is always the penultimate syllable of the word as it appears immediately before the obligatory monosyllabic stem. This arrangement leaves the mode-subject syllable and the stem syllable as the smallest permissible verb, forming an obligatory paradigm of AUX+STEM, thus constituting a bisyllabic core (McDonough, 2003). Although there is little additional evidence, this enforced word minimality constraint may be due to a metrical requirement based on words necessitating binary feet. Additionally, stems typically exhibit greater prominence, suggesting a right-aligned iambic foot. To better understand how stems mark prominence, I will illustrate certain tendencies observed in Navajo verb structure.

Firstly, stems tend to comprise heavier, more complex syllables (CVV or CVC), whereas conjuncts, barring the mode-subject morph, predominantly consist of lighter, simpler syllables (overwhelmingly CV in shape). Contrary to Witsuwit'en, this contrast hints at a higher ranking of WEIGHT-TO-STRESS than STRESS-TO-WEIGHT, presuming stem syllables attract stress. Following this parameter and variance in constraint ranking, weight will gravitate toward the stem. Secondly, stem syllables are realized as phonetically longer than disjunct or conjunct syllables, though this may be due to final lengthening rather than prominence markers of an iambic foot. Notably, tones, often used to mark prominence, are disproportionately L to the left of the stem while stems demonstrate an on-par distribution of H- and L-tones. These disparities between stem and non-stem syllable types point to the likely construction of a right-headed, right-aligned foot in Navajo verbs, where prominence is realized through complexity, weight, length, and tone. Furthermore, diminished prominence seen in non-stem syllables is attributed to anti-prominence constraints on the conjunct and disjunct, like those documented in Tlingit (Crippen et al., 2023), thereby allowing stems to receive greater relative prominence.

Though foot structure in Navajo is evident from word minimality and prominence constraints, the occurrence of iterative feet supporting a rhythmic stress system remains somewhat elusive. The disjunct, conjunct, and stem combine to form a singular prosodic word because the independence of these domains would entail the culmination of each, which is not attested. As well as this, if iterative iambic feet were being parsed, we would expect secondary stresses on alternating syllables to the right of the stem, or a pattern similar to Witsuwit'en where stress is both attracted to the word-final stem and the left edge because of a left-to-right directionality of computation. Unfortunately, we do not see any signs of secondary stresses, instead, the overwhelming distribution of L tones to the left of the stem resembles that of a pitch accent system. Here, I interpret pitch accent as the rhythmic distribution of tones as markers of accent. Although little attention has been dedicated to the predictability of pitch accent in Navajo, I posit the presence of a simple stress system characterized by a single peak per word, owing to the non-iterative construction of feet in addition to a set of prominence constraints that regulate the distribution of prominent phenomena, such as H-tones, heavy syllables, and long vowels.

6 Discussion

Linguists in the past have proposed a continuum (McCawley, 1978; McDonough, 1999) where "tonal" typologies are defined by three main factors: the independence of the tonal tier, the function of tone, and the density of specification. This conceptual framework predicts that tone, pitch accent, stress, and intonation represent varying manifestations of a single mechanism, with languages occupying distinct points along the continuum. This premise cannot be true if U. Tanana shows that tone is featural and paradigmatic, Witsuwit'en shows that stress is culminative and rhythmic, and Navajo shows signs of having both devices. A spectrum on which languages can shift is hence not the right approach.

Tones in U. Tanana prove to be specified and underspecified in the lexical entries of morphemes. They are subject to rules and derivation akin to other features, undergoing spreading, assimilation, and interaction with immediate neighbours. Conversely, Witsuwit'en stress cannot function this way because its assignment is contingent on a set of parameters and constraints that dictate prominence peaks at a consistent location. The shape of the word and its syllable types influence the locus of stress, not the underlying specification of each word. If tones were assigned like stress and not stored underlyingly, the phonology would require an ordered script in which tones are first located through one set of rules and then interact with each other through a separate set of rules, each at a different time in the derivation. Correspondingly, if stress was stored underlyingly, the lexical entry would need to be informed by the syllabification of the word after all phonological processes post-derivation and this is unfeasible. Additionally, if stresses were fixed underlyingly, it would not be variable after compounding or the incorporation of clitics, contrary to the facts in the data.

Tone functions as a polar (sometimes multipolar) system of distinctive values that correlate with pitch excursions relative to an average level. Importantly, H-tones lift above the baseline into the upper envelope while L-tones fall below. How spaced out their frequency interval is should not be relevant as long as H is relatively higher than L to successfully mark a contrast. Especially toward the end of the utterance, the frequency envelope of H- and L-tones should be fairly compressed, but the data from Navajo suggest differently. Stress is a hierarchical system of contrastive levels of prominence and a sudden expansion of the pitch envelope at the verb stem indicates the presence of stress on top of tone. Because stress utilizes f_0 and needs to be the highest prominence peak, it must amplify the tonal specification already present and decrease any other surrounding salient elements to stick out as much as possible. Tone alone is not capable of achieving this nor does it need to because it solely relies on a comparative difference, whereas stress is the opposite: it relies on marking a single syllable in a domain with maximum saliency.

The origins of tone and stress should also rule out their relationship. Tonogenesis involves the suprasegmentalization of material in the melody in the same way as features can affect each other, spread, or erode over time with the differentiation that tone is taken up to its independent autosegmental tier. By contrast, there is no process of "accentogenesis" denoting the birth of stress, rather stress appears to be a cognitive fact. Rhythm is universal because humans have a tendency towards repetitive patterns. Consonant-vowel alternation is simply the bouncing closure and opening of the speech organ stemming from the elasticity of the articulators. These alternations are grouped into syllables, which are grouped into feet and are assigned beats, reflecting rhythmic binarity at all levels of scope. Heartbeats also come in binary strong-weak patterns because of dual contractions, our walking is binary because of our lateral symmetry, and a 4/4 time signature can be found universally in all music traditions around the world. All these activities are impossible not to rhythmicize because alternating patterns are precognitively programmed into our biology, and this the same phenomenon that gives rise to stress. In other words, stress is simply the phonetic culmination within a rhythmic pattern. The same cannot be said for tone because, unlike stress, tonality is not an intrinsic reaction to rhythm; instead, tonal distinctions are developed language-specifically.

7 Conclusion

The examination of tone and stress as linguistic systems reveals their interaction and distinct characteristics. Tones, as distinctive elements, form part of lexical entries alongside features and are subject to diachronic change and exclusively local phonological rules. Conversely, stress is a symptom of inherent rhythm and operates as a structural and contrastive phenomenon, positioning itself as a peak within the metrical structure of word domains. Despite their fundamental differences, both tone and stress rely on phonetic mechanisms, particularly f_0, for their manifestation in speech, which leaves room for their common categorical association. Understanding the differing mechanisms and interactions of these systems sheds light on how language organizes prosodic constituents and autosegments. Further research into the overlaying of tone and stress will enrich our comprehension of language typologies and their role in communication.

References

Alber, B. (2005). Clash, lapse and directionality. *Natural Language & Linguistic Theory*, *23*(3), 485-542.

Anderson, S. R. (1978). Tone features. In V. A. Fromkin (Ed.) *Tone* (pp. 133-175). Academic Press.

Broadwell, G. A., & Zhang, J. (2000). Macuiltianguis Zapotec tone paradigms. *Ms. SUNY Albany.*

Goad, H. & Qu, C. (2011). On the interaction of tone and stress in contour tone languages: A prosodic account. *Proceedings of the 2008 Annual Conference of the Canadian Linguistic Association,* (pp. 1-39).

Goldsmith, J. (1980). [Review of *Tone: A Linguistic Survey*, by V. A. Fromkin]. *Language*, *56*(2), 413–418. https://doi.org/10.2307/413763

Halle, M., & Vergnaud, J. R. (1982). On the framework of autosegmental phonology. *The structure of phonological representations*, *1*, 65-82.

Hargus, S. (2001). Quality-sensitive stress reconsidered. *Working papers in linguistics, 20,* 25-56. University of Washington.

Hargus, S. (2007). *Witsuwit'en grammar: Phonetics, phonology, morphology*. UBC Press.

Hayes, B. (1986). A revised parametric metrical theory. *North East Linguistics Society, 17*(1), 274-289.

Hyman, L. M. (2006). Word-prosodic typology. *Phonology, 23*(2)*,* 225-257. Cambridge University Press. 10.1017/S0952675706000893

Hyman, L. M. (2008). Universals in phonology. *The linguistic review, 25*, 83-137. De Gruyter Mouton. 10.1515/TLIR.2008.003

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, *5*(2), 275-292. Kenstowicz, M. (1996). Quality-sensitive stress. In John J. McCarthy (Ed.), *Optimality theory in phonology: A reader*. (pp. 191-201). Blackwell. 10.1002/9780470756171

Kim, M. R. (2012). Tonogenesis in Korean: Some recent speculations on the sound change. 언어, *37*(2), 243-283.

Kingston, J. (2005). The phonetics of Athabaskan tonogenesis. In H. Sharon & K. Rice (Eds.) *Athabaskan prosody*, (pp. 137-184). John Benjamins Publishing.

Krauss, M. E. (2005). Athabaskan tone. In S. Hargus & K. Rice (Eds.), *Athabaskan prosody* (pp. 51-126). John Benjamins Publishing.

Levine, R. D. (1979). Haida and Na-Dene: A New Look at the Evidence. *International Journal of American Linguistics*, *45*(2), 157–170. http://www.jstor.org/stable/1264780

Leer, J. (2005). How stress shapes the stem-suffix complex in Athabaskan. In S. Hargus & K. Rice (Eds.), *Athabaskan prosody*, (pp. 277-318). John Benjamins Publishing.

Leer, J. (2008, February). Recent advances in AET comparison. In *Dene-Yeniseic Symposium, Fairbanks*.

Leer, J. (1999). Tonogenesis in Athabaskan. In S. Kaji (Ed.) *Cross-linguistic studies of tonal* *phenomena: Tonogenesis, typology, and related topics*, (pp. 37-66). Tokyo University of Foreign Studies.

Lovick, O. (2020). *A grammar of Upper Tanana, 1: Phonology, lexical classes, morphology.* University of Nebraska Press.

McCarthy, J. J. (2003). OT constraints are categorical. *Phonology*, *20*(1), 75-138.

McCawley, J. D. (1978). What is a tone language? In V. Fromkin (Ed.) *Tone: A Linguistic Survey*, IV. (pp. 113-131). Academic Press. 10.1016/B978-0-12-267350-4.50009-1.

McDonough, J. M. (1999). Tone in Navajo. *Anthropological linguistics 41*(4), 503-540. The Trustees of Indiana University.

McDonough, J. M. (2003). The Navajo sound system. *Studies in natural language and linguistic theory, 55*. Springer Science + Business, Media.

Meyers, S. (1997). OCP effects in Optimality Theory. *Natural Language & Linguistic Theory*, *15*(4), 847-892.

Morice, A. G. (1932). *The Carrier language (Diné family): A grammar and dictionary combined*. Winnipeg.

Penny, R. (2002). Phonology. In *A History of the Spanish Language* (pp. 34–110). Cambridge University Press.

Rehg, K. L., & Sohl, D. G. (1981). *Ponapean reference grammar*. University of Hawaii Press. Reichard, G. A. (1951). *Navajo Grammar*. J.J. Augustin.

Rice, K. & Hargus, S. (Eds.). (2005). *Athabaskan prosody* (Vol. 269). John Benjamins Publishing.

Sapir, E. (1915). The Na-Dene languages, a preliminary report. *American Anthropologist*, *17*(3), 534-558.

Young, R. W. & Morgan, W. (1987). *The Navajo language: A grammar and colloquial dictionary* (Rev. ed.). University of New Mexico Press.

Williams, J. S. (2009). *An intermediate grammar of Navajo: A guide for bilingual educators*. The University of New Mexico.

Woo, N. H. (1969). *Prosody and phonology.* [Doctoral dissertation, MIT]. Zec, D. (1989). *Sonority constraints on prosodic structure*. [Doctoral dissertation, Sanford University]. Routledge.