PHONOLOGY

ARTICLE

An acoustic study of Tetsót'ıné stress: Iambic stress in a quantity-sensitive tone language

Alessandro Jaker[1] and Phil J. Howson[2]

[1]Department of Linguistics, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada;
E-mail: alessandro.jaker@utoronto.ca
[2]Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Schützenstraße 18 10117 Berlin, Germany;
E-mail: howson@leibniz-zas.de.

Abstract

This paper presents both distributional and acoustic phonetic evidence for iambic stress in Tetsót'ıné (ISO: CHP), a Dene (Athapaskan) language with contrastive vowel length and four contrastive tones. In our acoustic study, we find that the primary correlate of stress in Tetsót'ıné is duration, whereas intensity plays a secondary but statistically significant role. There was no statistically significant effect on F0 in our results. We discuss our results in relation to several proposals regarding the typology of stress systems. Based on the Functional Load Hypothesis (Berinstein 1979) and Dispersion Theory (Flemming 1995, 2001), we find that our results are to some extent unexpected. We suggest that our results are most consistent with the Iambic–Trochaic Law (Hayes 1995), which predicts that iambic stress systems prefer to use duration as their primary stress correlate.

Contents

CrossMark

1. Introduction: Stress in a quantity-sensitive tone language

In his now classic cross-linguistic study of the typology of stress systems, Hayes suggests that stress may be unique among phonological categories, in that its phonetic realisation depends almost entirely upon properties that are also used (even within the same language) for other purposes:

(1) 'The multiple phonetic cues for stress and the subordinate role of loudness are particularly interesting when one considers that languages use duration and pitch in their phonological systems for entirely different purposes. Duration is the phonetic cue for vowel length, which is phonemic in many languages…. Further, pitch is the phonetic cue for tone in languages with phonemic tone systems and is also the basis of intonation. The basic point is this: aside from the marginal role of loudness, stress is parasitic in the sense that it invokes phonetic resources that serve other phonological ends' (Hayes 1995: 7).

Hayes (1995: 7) goes on to suggest that languages with a phonemic vowel length contrast will avoid using duration as a correlate of stress. Tuttle (1998: 18) applies this same line of reasoning to tone, and provides evidence that pitch is not used as a correlate of stress in Lower Tanana Dene (Athabaskan), a language with contrastive tone. Along these same lines, all other things being equal, we might expect that a language with phonemic vowel length might rely more on pitch as a cue to stress, whereas a language with contrastive tone might rely more on duration. This typological generalisation has been termed the Functional Load Hypothesis (Berinstein 1979), as stated in (2).

(2) 'Change in F0, increased duration, and increased intensity, in that order, constitute the unmarked universal hierarchy for perception of stress in languages with no phonetic contrasts in tone or vowel length; in languages with such contrasts the perceptual cue correlated with that contrast (i.e. F0 with tone and duration with length) will be superseded by the other cues in the hierarchy' (Berinstein 1979: 2, cited in Lunden *et al.* 2017: 565).

However, in the literature, the Functional Load Hypothesis has been seen as controversial. For example, Lunden *et al.* (2017), based on a typological survey of 140 languages, found that there is no correlation between whether a language has contrastive vowel length and whether it uses vowel duration as a phonetic cue for stress. In this context, our paper poses the following empirical question: What would we expect stress to look like in a language with both contrastive tone and contrastive vowel length? In this paper, we will present an acoustic study of stress in Tetsǫ́t'ıné, a Dene (Athabaskan) language variety with four contrastive tones (high, low, rising, and falling) and a surface vowel length contrast (short *vs.* long) in both stems and prefixes. We will provide both acoustic phonetic and distributional evidence

to argue that this language exhibits a quantity-sensitive, left-to-right iambic stress system, in addition to its tone and vowel length contrasts.

If the Functional Load Hypothesis were taken at face value, we might expect that a language such as Tetsǫ́t'ıné would employ intensity as its main stress correlate. Indeed, a very similar prediction is made by Dispersion Theory (Flemming 1995, 2001), as we will discuss in §7. At the same time, the statements in (1) and (2) both seem to assume a universal dispreference for using intensity as a primary stress correlate. If this dispreference for intensity were strong enough, it might be expected that a language such as we describe could not actually exist. Indeed, it has been claimed that no language can exhibit stress, tone and phonemic vowel length simultaneously. For example, Spahr (2016: 206) states that 'in a language with both stress and autosegmental tone, stress must be represented as a quantity contrast…leaving no room for an independent length contrast.' Since languages exhibiting stress, phonemic vowel length and tone are indeed typologically rare (though see Potisuk *et al*. 1996 for one example), we believe that an acoustic study of Tetsǫ́t'ıné stress will be an important contribution to our understanding of the phonetic typology of stress systems, since the language exhibits a combination of prosodic features which is generally assumed to be typologically unusual.

Our study also adds to a growing body of work on metrical structure in tone languages. The earliest such study that we are aware of is that of Rice (1990) on metrical structure in Hare. Rice uses trochaic footing to explain several phonological processes in Hare, including movement of high tones and vowel syncope. In her analysis of the Neo-Štokavian dialect of Serbian or Croatian, Zec (1999) presents evidence for a bidirectional interaction between tones and metrical structure whereby the foot inventory is constrained by tones, and the distribution of tones is constrained by feet. Finally, DeLacy (2002) appeals to the foot to account for the preferential attraction of stress by the falling HL tone over the simple H tone in Ayutla Mixtec. However, we believe that our study stands apart from these previous studies for several reasons: we provide a combination of distributional and acoustic phonetic evidence for iambic stress in Tetsǫ́t'ıné; the distributional evidence includes several different processes (tone movement, vowel length adjustment and consonant deletion); and we provide evidence for iterative stress (more than one stress per word). Therefore, it is our opinion that when one considers the evidence as a whole, Tetsǫ́t'ıné may provide the most compelling case for metrical structure in a tone language among all the examples just cited.

The remainder of this paper is organised as follows. In §2 we provide background on previous work on stress in Dene languages. In §3 we provide background information on Tetsǫ́t'ıné, including distributional evidence for contrastive vowel length, tone and stress. Sections 4–6 describe our acoustic study. In §4 we present our hypothesis, where we list four possible tone patterns for two-syllable words and eight for three-syllable words, and the hypothesised locations of stresses in each of these tone patterns. In §5 we describe our experimental methods, and in §6 we present our results. In §7, we discuss our results in terms of several theoretical proposals relating to stress: the Functional Load Hypothesis, which is closely related to Dispersion Theory (Flemming 1995, 2001), and the Iambic–Trochaic Law (Kager 1993; Hayes 1995). Broadly speaking, our results suggest that duration is the primary cue for stress

in this language. This is predicted by the Iambic–Trochaic Law, for iambic languages, but not by Dispersion Theory or the Functional Load Hypothesis, which predict, all other things being equal, that intensity ought to be the primary stress cue in a language with both contrastive vowel length and tone. Therefore, we interpret our results as providing phonetic evidence in support of the Iambic–Trochaic Law. §8 provides a summary and conclusion.

2. Previous work on stress in Dene languages

In early work on Dene languages, at least for those Dene languages that are tonal, it was often assumed that stress played little or no role in the phonological system (Rice and Hargus 2005b: 34). For example, Hoijer (1946, cited in Rice and Hargus 2005b: 34) describes Chiricahua Apache as 'a succession of evenly stressed syllables', and Everett (1998) provides some useful historical background which may explain this state of affairs. Whereas in early work (e.g. Bloomfield 1933) it was assumed that stress was realised primarily by increased amplitude, by the late 1950s stress came to be directly associated with fundamental frequency. This gave rise to the assumption held by many linguists then as now that stress and tone are somehow mutually exclusive: a 'tone language' cannot also be a 'stress language'.

More recently, several linguists have found evidence for stress in Dene languages, including metrical stress, from a variety of perspectives. Rice (1990) presents morphophonemic evidence for trochaic footing in Hare and Tuttle (1998) presents phonetic evidence for trochaic footing in Tanana. Hargus (2005a) provides evidence that high tone attracts stress in Sekani, whereas Hargus (2005b) finds near-minimal pairs for stress in both Sekani and Deg Xinag. One difference between Hargus's (2005b) study and ours is that Hargus specifically excluded verbs from the data set, whereas our study focuses entirely on verbs. This is because we believe that stress plays a significant role in shaping verbal morphophonemics (see §3.4). Finally, Leer (2005) argues that stress has played a significant role in the history of Dene languages, specifically in causing a shortening of suffixes with full (long) vowels after the stem.

Thus, whereas several previous authors have provided evidence for stress in Dene languages, our study is novel in that we present converging evidence from morphophonemics (§3.4) and acoustic phonetics (§6) for iambic stress in Tetsǫ́t'ıné. In addition, with the exception of Tuttle (1998), our study is unique in that we situate our results within the phonetic typology of stress systems (Gordon and Roettger 2017; Lunden *et al.* 2017): given the exceptionally crowded prosodic space of Tetsǫ́t'ıné – with a vowel length contrast and four contrastive tones – how is it possible to phonetically realise stress without neutralising the other prosodic contrasts in the language?

3. Background on Tetsǫ́t'ıné

3.1 Segmental inventory

Tetsǫ́t'ıné is a dialect of Dëne Sųłıné (Ethnologue: CHP) spoken in Canada's Northwest Territories. It is a member of the Dene (Athabaskan) language family. In

this section we will present the segmental inventory of Tetsǫ́t'ıné, as well as the transcription system which will be used in this paper. The consonant inventory in (3) is similar to the Dëne Sųłıné consonant inventory proposed by Li (1946: 398), the main difference being that Tetsǫ́t'ıné does not seem to possess a labial-velar series (cf. Cook 2004: 22–23), whereas Tetsǫ́t'ıné does seem to distinguish alveolar /n/ from palatal /ɲ/ underlyingly (Jaker and Kiparsky 2020). Our choice of symbols to represent the consonant inventory in (3) also differs from the standard orthography of the language, as well as from convention among Deneists. In the Dene linguistics literature, the 'plain' stops and affricates, which are phonetically weakly voiced or voiceless, are customarily transcribed as voiced ⟨d, dz, dl⟩, whereas the aspirate series is transcribed as voiceless ⟨t, ts, tɬ⟩ (e.g. Li 1946; Ackroyd 1982; Rice 1989; Cook 2004). In this paper, following IPA convention, we will transcribe the plain series as voiceless [t, ts, tɬ], and the aspirate series as aspirated [tʰ, tsʰ, tɬʰ], following, for example, Lovick (2020). The rhotic is sometimes realised as a trill [r], though more commonly as an approximant [ɹ], as in North American English. Here, we represent this segment in broad transcription as ⟨r⟩.

(3) *Tetsǫ́t'ıné consonant inventory*

		Labial	Inter-dental	Alveolar	Lateral	Palato-alveolar	Velar / Uvular	Glottal
Stops and affricates	Plain	p	tθ	t, ts	tɬ	tʃ	k	
	Aspirate		tθʰ	tʰ, tsʰ	tɬʰ	tʃʰ	kʰ	
	Ejective		tθ'	t', ts'	tɬ'	tʃ'	k'	ʔ
Fricatives	Voiced		ð	z	l		ɣ	
	Voiceless		θ	s	ɬ	ʃ	x	h
Sonorants	Oral	w		r		j		
	Nasal	m		n		ɲ		

Dene languages are polysynthetic, prefixing languages, with a stem preceded by a number of prefix positions (Rice 1989). In Tetsǫ́t'ıné, a contrast between long and short vowels is found in both stems and prefixes. In stems, this contrast is largely a reflex of the historical (Proto-Dene) contrast between full and reduced vowels (cf. Krauss 1964, 1983). The contrast between long and short vowels in stems is fundamentally a length contrast, and realised phonetically as a combination of duration and vowel quality differences (Jaker 2019), not unlike the contrast between 'tense' and 'lax' vowels in English. This is reflected in our transcription of vowel length in stems (e.g. /ʌ/ *vs.* /aː/, as shown in (4)). In prefixes, long vowels are of more recent historical origin, resulting from the deletion of intervocalic consonants (Cook 2004: 41), which may still be a synchronically active process (Jaker to appear). That is, long vowels in prefixes originate historically as sequences of two adjacent vowels which were previously separated by a consonant. Although we lack phonetic data on the correlation between vowel length and vowel quality in prefixes in Tetsǫ́t'ıné, Jaker's subjective impression is that it is phonetically more of a pure length difference.

This is reflected in our transcription of prefix vowel length, where we write, for example, ⟨a⟩ *vs.* ⟨aː⟩, ⟨i⟩ *vs.* ⟨iː⟩, ⟨e⟩ *vs.* ⟨eː⟩, etc. Like many Dene languages, Tetsǫ́t'ıné also has contrastive nasality, which will be written as a tilde over the vowel: ⟨ã⟩.

(4) *Tetsǫ́t'ıné vowel inventory*

	STEMS			PREFIXES		
	Front	Central	Round	Front	Central	Round
High	iː		ʊ, uː	i, iː		u, uː
Mid	eː, ə		oː	e, eː		o, oː
Low		ʌ, aː			a, aː	

3.2 Distributional evidence for vowel length contrast

In previous work on Tetsǫ́t'ıné and related Dëne Sųłıné dialects, nearly all studies agree, for the dialects they examine, that Dëne Sųłıné exhibits some sort of vowel length contrast on the surface (Li 1933, 1946: 399; Haas 1968; Krauss 1983; Cook 1983, 2004: 28–30, 41). Sources differ, however, as to whether this vowel length contrast is the same in stems as in prefixes, and as to whether the contrast is underlying or derived in the synchronic phonology. Cook (1983: 424), for example, explicitly denies the existence of an underlying vowel length contrast in Dëne Sųłıné.

In this paper, we will assume that long vowels in stems (i.e. full vowels) are underlying, whereas long vowels in prefixes are synchronically derived (Jaker to appear). However, nothing in our acoustic study (§4–6) or the interpretation of its results depends upon this assumption.

3.2.1 Stems

Stems in Tetsǫ́t'ıné exhibit an inventory of five long (i.e. full) vowels /iː/, /eː/, /aː/, /oː/, /uː/, and three short (i.e. reduced) vowels, /ə/, /ʌ/, /ʊ/ (in IPA broad transcription; see Jaker 2019 for a discussion of the precise phonetic quality of these vowels). The relationship between long and short vowels in stems is primarily a historical one, which has been mediated by a vowel reduction process that applied in Pre-Proto-Dene (Krauss 1964). A relationship between long and short stem vowels synchronically exists only in a few fossilised stem alternations (cf. Li 1933, 1946). The vowel system in (4) could be thought of as a standard five-vowel system for the long vowels, plus a three-vowel system for the short vowels. In terms of their phonological features (Jaker 2020), it is expected that /iː/ and /eː/ will reduce to /ə/, /aː/ will reduce to /ʌ/, and /uː/ and /oː/ will reduce to /ʊ/, wherever stem alternations involving vowel length are still found in the language.

For ease of exposition, we will compare only pairs of vowels which differ from each other solely in the property of being long or short, without additional vowel height differences. Thus, in the examples below we will compare /ʌ/ *vs.* /aː/, /ə/ *vs.* /eː/, and /ʊ/ *vs.* /uː/. Since it is often difficult to find true minimal pairs in Dene languages, due to the large phoneme inventory of these languages, some near-

minimal pairs are also included. The vowels being compared are highlighted in bold.

(5) *Short /ʌ/* vs. *long* /aː/ (Cardinal *et al.* 2021)

	a.	*Short /ʌ/*			b.	*Long* /aː/	
	i.	xʌ́s	'boil, sore'		i.	xá:s	'tree stump'
	ii.	ts'ásk'ʌðe	'across'		ii.	k'aːði	'health, healing'
	iii.	təne ɣʌjé	'person's age'		iii.	ɬue ɣaːjé	'fish's fins'
	iv.	hatsʰʌ́l	'nail'		iv.	tʰezá:l	'anchor'
	v.	jʌθ	'snow'		v.	horeljaːθ	'flames'

(6) *Short /ə/* vs. *long* /eː/ (Cardinal *et al.* 2021)

	a.	Short /ə/			b.	*Long* /eː/	
	i.	təl	'blood'		i.	te:l	'crane'
	ii.	tsʰəs	'firewood'		ii.	tɬe:s	'grease'
	iii.	tʰuːðəle	'broth'		iii.	tʰuːðé:lé	'rubber boots'
	iv.	tʃʰəθ	'duck'		iv.	hetʃʰe:s	'it is boiling'
	v.	ðə́θ	'hide, pelt'		v.	tʃé:θ	'hook'

(7) *Short /ʊ/* vs. *Long* /uː/ (Cardinal *et al.* 2021)

	a.	*Short /ʊ/*			b.	*Long* /uː/	
	i.	ʔeɬk'iːθkʰʊ́né	'gunpowder'		i.	ɬuːek'uːné	'fish eggs'
	ii.	ts'ʊs	'fly' (n.)		ii.	ts'ú:s	'elbow'
	iii.	helts'ʊ́n	'kiss' (3sg.)		iii.	setsʰuːné	'my grandmother'
	iv.	hutʰʊ́n	'hold' (3sg.)		iv.	ʔerílt'uːs	'glue' (3sg.)

As mentioned above, many of the examples in (5)–(7) are not perfect minimal pairs, often involving additional differences in consonant place of articulation and tone. However, we are not aware of any evidence that consonant place of articulation, such as the difference between /tsʰ/ and /tɬʰ/, could condition a vowel length difference, as in (6a.ii) and (6b.ii). The role of tone, however, does require some comment. In prefixes, as we will see in §3.4, there is some interaction between tone and vowel length, which is mediated by the fact that high tone attracts stress. In stems, however, there is no evidence that stem tone has any effect on whether the stem vowel is long or short. For example, in motion verbs, stems generally have a high tone in the imperfective, and a low tone in the perfective. But if the stem contains a long vowel, that vowel will remain long whether the stem tone is high or low (e.g. *heeté:l* 'they (pl.) leave', *heéte:l* 'they (pl.) left'). Similarly, a short vowel will remain short, whether its stem tone is high or low (e.g. *heeʔʌs* 'they (two) leave', *heéʔʌs* 'they (two) left'). Therefore, tone differences in the examples in (5)–(7) do not constitute a confound in relation to the contrast between long and short vowels.

3.2.2 *Prefixes*

In prefixes, Tetsǫ́t'ıné has an inventory of five vowels, /i/, /e/, /a/, /o/ and /u/, all of which can be made long in certain morphophonemic environments. Briefly, all

long vowels in prefixes are the result of intervocalic consonant deletion, although it is not always the case that intervocalic consonant deletion results in a long vowel. (See Jaker (to appear) for a detailed study of consonant deletion and prefix vowel length in Tetsǫ́t'ıné). Since all surface long vowels are derived via morphophonemic processes, it could be said that prefix vowel length is contrastive on the surface in Tetsǫ́t'ıné, in that it realises morphosyntactic distinctions on the surface, even though it is not contrastive underlyingly. Some examples will be provided below.

There are two situations in which vowel length can be used minimally to signal a grammatical contrast: number marking in 'bare verbs' and aspect marking in ɣe conjugation, d/l-classifier verbs. We will examine each of these in turn below.

In verbs which lack a thematic disjunct prefix or any other thematic material – Jaker and Cardinal (2020: 74) call these 'bare verbs' – the third-person singular imperfective forms contain a short vowel, whereas the corresponding third-person plural forms contain a long vowel. Some examples are given in (8) below. The vowels being compared are highlighted in bold.

(8) *Plurality marked by vowel length in bare verbs* (Jaker & Cardinal 2020: 75–77)

	Singular: short vowel		*Plural: long vowel*	
a.	he**t**ʃən	'he/she sings'	he:tʃən	'they sing'
b.	he**l**té:l	'he/she eats many things'	he:lté:l	'they eat many things'
c.	he**t**ã:	'he/she drinks'	he:tã:	'they drink'
d.	he**ɬ**pe:s	'he/she boils [obj.]'	he:ɬpe:s	'they boil [obj.]'

In *d/l*-classifier, ɣe conjugation verbs (Jaker and Cardinal 2020: 88), the perfective forms are distinguished from their corresponding imperfective forms by vowel lengthening. Short vowels with high tone in the imperfective become long vowels with falling tone in the perfective, whereas short vowels with low tone in the imperfective become long vowels with low tone in the perfective. This pattern is best illustrated using the verb *ʃétʰi:* 'eat', as shown in (9).

(9) *Perfectivity marked by vowel lengthening* (Jaker & Cardinal 2020: 128–129)

	Imperfective: short vowel		*Perfective: long vowel*	
a.	ʃé**s**tʰĩ:	'I eat'	ʃê:stʰĩ:	'I ate'
b.	ʃé**t**ʰĩ:	'he/she eats'	ʃê:tʰĩ:	'he/she ate'
c.	ʃéhe**t**ʰĩ:	'they (du.) eat'	ʃéhe:tʰĩ:	'they (du.) ate'
d.	ʃéhe**l**ji:	'they (pl.) eat'	ʃéhe:lji:	'they (pl.) ate'

To summarise, in this section we have seen that vowel length is used contrastively in Tetsǫ́t'ıné. In stems, it is used to distinguish different lexical items, as we saw in (5)–(7). In prefixes, vowel length is used to realise morphosyntactic distinctions, as we saw in (8)–(9). This has implications for the phonetic realisation of stress. Apart from those few cases where stress does seem to alter phonological vowel length (see §3.4), to the extent to which vowel duration is used as a phonetic correlate of

stress, we would expect that it would be used in such a way as to not neutralise contrastive length oppositions. That is, all other things being equal, we would expect a lesser degree of vowel lengthening under stress in Tetsǫ́t'ıné as compared with a language which lacks contrastive vowel length.

3.3 Distributional evidence for tone contrast

In this section we will provide an overview of the tone contrasts that exist in Tetsǫ́t'ıné. The first general fact to establish is that the inventory of tonal contrasts depends both on position (stems versus prefixes) and vowel length (short versus long). Specifically, in prefixes, short vowels contrast two tones (high and low), whereas long vowels contrast four tones (high, low, rising and falling). In stems, on the other hand, there are only two tones, high or low, regardless of whether the vowel is long (full) or short (reduced). This is illustrated in (10).

(10) *Tone contrasts in different positions* (Jaker & Cardinal 2020: 116, 122, 124, 146, 173; Cardinal *et al.* 2021)

	Prefixes		Stems	
	Short V	Long V	Short V	Long V
High	héja: 'he/she left'	níːteːl 'we (pl.) arrived'	ðɔ́θ 'hide, pelt'	het'éːθ 'it is cooking'
Low	hetsʰaːɣ 'he/she cries'	nániːɬt'uːs 'I punched you'	ʧʰəθ 'duck'	heʧʰeːs 'it is boiling'
Falling	—	nîːja: 'I arrived'	—	—
Rising	—	hĭːɬk'íːθ 'they shoot'	—	—

The generalisation that stem vowels contrast only two tones is true only when the nucleus of the stem syllable consists of a single vocalic root node. When the stem contains a diphthong, both level and contour tones are possible, as in *θai* 'sand' (low tone), *θái* 'long ago' (high tone), *ʔeɬdðái* 'dry fish' (falling tone), *teʧənðai* 'sawdust' (rising tone). Contour tones can also arise in stems due to postlexical consonant–vowel metathesis and vowel coalescence in a phrasal context: thus compare *ʧíːze* 'whiskeyjack' and *ʧîːz ʧʰoː* 'hawk' (lit. 'big whiskeyjack'). The broader generalisation which emerges from these examples is that in both stems and prefixes, the number of possible tonal specifications is equal to the number of vocalic root nodes. Thus, in prefixes, all long vowels consist of two root nodes and are therefore able to host contour tones. In stems, long vowels can host only level tones, because they consist of a single vocalic root node, except in the cases of consonant–vowel metathesis or diphthongs, as described above.

Given this context, in the remainder of this section we will focus our presentation on long vowels in prefixes, since this is where the full set of tonal contrasts can be most easily illustrated. Although as part of our experimental design we excluded contour tones from the target words in our stimulus set (§4), it is important to consider the full set of tonal contrasts that exist in the language, since these define the overall 'tone space' and can potentially have an effect on how F0 is used as a stress correlate.

In the following examples, we will provide sets of morphologically related words, all with long vowels in the penultimate syllable, which contrast high versus falling tone, and low versus rising tone. In the terminative paradigms of motion and handling verbs, we find falling tones in the singular forms, and high tones in the first-person dual and plural forms. This is illustrated in (11).

(11) *Terminative motion verbs: falling tone in singular, high tone in first-person dual/ plural* (Jaker & Cardinal 2020: 147–148, 184–185)
    a. *Singular: falling tone*
        i. nî:le: 'he/she puts down (pl. objects)' (IPFV)
        ii. nî:la: 'I put down (pl. objects)' (PFV)
        iii. nî:kʰĩ: 'I arrived by boat'
        iv. nî:t'a:ɣ   'I arrived by airplane'
    b. *First-person plural: high tone*
        i. ní:lje: 'we put down (pl. objects)' (IPFV)
        ii. ní:lja: 'we put down (pl. objects)' (PFV)
        iii. ní:kʰĩ: 'we (du.) arrived by boat'
        iv. ní:t'a:ɣ   'we (du.) arrived by airplane'

In the third-person dual and plural forms of the inceptive paradigms of motion verbs, we find a low tone in the imperfective and a rising tone in the perfective. This is illustrated in (12).

(12) *Inceptive motion verbs: low tone in imperfective, rising tone in perfective* (Jaker & Cardinal 2020: 171–172; Jaker's field notes 2 November 2020)
    a. *Imperfective: low tone*
        i. he:ʔʌs   'they (du.) leave (on land)'
        ii. he:kʰĩ: 'they (du.) leave (by boat)'
        iii. he:t'a:ɣ   'they (du.) leave (by airplane)'
        iv. he:té:l   'they (pl.) leave'
    b. *Perfective: rising tone*
        i. hě:ʔʌs   'they (du.) left (on land)'
        ii. hě:kʰĩ: 'they (du.) left (by boat)'
        iii. hě:t'a:ɣ   'they (du.) left (by airplane)'
        iv. hě:te:l   'they (pl.) left'

As part of our experimental design, we excluded any words containing contour tones from the set of target words (see §5). However, in this section we have provided an illustration of the contrastive status of contour tones for the following reason: in a language with contrastive contour tones, it has been proposed that for every tone, the initial pitch target, the final pitch target, the magnitude of the pitch rise or fall, and the slope of the rise or fall are all phonetically specified in the phonetic component

of the grammar (Flemming and Cho 2017). To the extent to which lexical tonal contrasts are not neutralised, this severely constrains the extent to which F0 can be used as a phonetic correlate of stress. Therefore, we would expect, all other things being equal, that Tetsǫ́t'ıné would exhibit a lesser magnitude of higher F0 under stress when compared with languages without contrastive tone or even languages with only two contrastive tones.

3.4 Distributional evidence for stress

In Tetsǫ́t'ıné, in the great majority of cases, stress is non-neutralising. That is, if the position of long vowels or high tones conflicts with a left-to-right iambic stress pattern, it is the position of stress which is adjusted rather than the tone or vowel length (Jaker and Kiparsky 2020). This is true of the set of target words which we will examine in §4–6. However, there is a minority of cases in which conflict between the stress pattern and the tone or vowel length patterns is repaired by altering tone and/or vowel length. These cases are important, because they provide evidence that stress is an integral component of the lexical phonology of the language and not a feature which can be restricted to the postlexical and/or phonetic component of the grammar. Rather, stress plays an active role in the phonological computation.

Tetsǫ́t'ıné is a quantity-sensitive, left-to-right iambic system in which high tone attracts stress. If we put aside the role of tone, the interaction between quantity and iambic feet in Tetsǫ́t'ıné is typologically normal for an iambic system (cf. Hayes 1995). Thus, every heavy syllable constitutes the head of a foot. Light syllables can form a foot together with a heavy syllable to their immediate right, or else two light syllables can also form a foot. This is illustrated in the examples in (13). Here, H and L refer to heavy and light syllables, respectively, rather than tones. The coda consonant ł does not add coda weight in (13c) and (13d). Note also that we assume that word-final light syllables are extrametrical, as in (13e).

(13) *Syllable weight and foot parsing in Tetsǫ́t'ıné* (Jaker & Cardinal 2020: 40–41, 76)

	Weight pattern	*Example*	
a.	(L.ˈH)	(he.ˈtã:)	'he/she drinks'
b.	(ˈH)(ˈH)	(ˈhe:)(ˈtã:)	'they drink'
c.	(L.ˈL)(ˈH)	(ja.ˈneł)(ˈtʰi:)	'you (sg.) speak'
d.	(L.ˈH)(ˈH)	(ja.ˈhĩːł)(ˈtʰi:)	'they spoke'
e.	(ˈH)L	(ˈtʰa:)ɣe	'three'

The stress pattern in (13) holds true for all words with level tone patterns (i.e. all-low tones or all-high tones). Where words contain a mix of low and high tones, high tone attracts stress, sometimes to a different position than we would expect solely based on the weight patterns in (13). For example, the word (ˈʃé.ne)(ˈtʰ ĩ̀:) 'you eat' has trochaic stress, due to the initial high tone, despite having the same weight pattern as in (13c). The relationship between tone and stress will be discussed in greater detail in §4.

In this section, we will examine three types of cases which provide distributional evidence for iambic stress. In §3.4.1, we will examine cases where stress conditions vowel length adjustment (i.e. the transfer of vowel length from its expected location

onto a neighbouring vowel). In §3.4.2, we will examine cases of movement of tone, where high tone surfaces one syllable farther left than expected to align with the iambic stress pattern. Finally, in §3.4.3, we will examine cases where consonants delete intervocalically to make the word more easily parsable into iambic feet.

3.4.1 Stress conditions vowel length adjustment

In Tetsǫ́t'ıné, when the consonant ɣ deletes intervocalically, it often leaves behind a long vowel. In the repetitive paradigm of reflexive verbs (meaning 'do to oneself repeatedly'), in the singular forms, this long vowel surfaces in its expected position, the place from which ɣ was deleted, which also happens to be the second syllable from the left edge and, therefore, the strong position of an iambic foot. This is illustrated in (14). In (14) and elsewhere, we assume that word-final codas (which are also stem-final) are moraic, whereas most codas in prefixes are not moraic, with two exceptions to be described in §3.4.3.

(14) *Long vowel surfaces in expected position in reflexive repetitive singular forms* (Jaker & Cardinal 2020: 85, 218)

	Underlying form	*Surface form*	*Expected form*
a.	/ʔete-ɣe-s-t-ts'ə́r/	(ʔe.ˈteːs)(ˈts'ə́r)	(same)
	'I scratched myself (repeatedly)'		
b.	/ʔete-ɣe-t-ts'ə́r/	(ʔe.ˈteːs)(ˈts'ə́r)	(same)
	'he scratched himself (repeatedly)'		
c.	/ʔete-ɣe-s-t-kʰáːr/	(ʔe.ˈteːs)(ˈkʰáːr)	(same)
	'I slapped myself (repeatedly)'		
d.	/ʔete-ɣe-t-kʰáːr/	(ʔe.ˈteː)(ˈkʰáːr)	(same)
	'he slapped himself (repeatedly)'		

By contrast, in (15), based on the segmental phonology of the language, we would expect to find a long vowel in the third syllable of the word, following the prefix *he*, since this is where the consonant ɣ has been deleted intervocalically. Instead, we find a long vowel in the second syllable, where there is no reason for a long vowel to arise based on segmental processes. However, if Tetsǫ́t'ıné is an iambic language, then these forms have a straightforward explanation. In a left-to-right iambic system, a sequence of alternating light and heavy syllables can be parsed more harmonically than a light–light–heavy–heavy sequence of syllables. Specifically, (light–heavy) iambs are most harmonic according to the constraint GROUPINGHARMONY(IAMB) (Prince 1990, 1991). Therefore, in the examples in (15), it appears that a mora has floated one syllable leftwards from its expected location to allow for a more harmonic iambic parse.[1]

---

[1] A reviewer suggests an alternative analysis, in which a metathesis of the prefixes *he* and *ɣe* occurs at the morphological level. Variable morpheme order is attested in Dene languages (Rice 2000; Hargus and Tuttle 2004), of which this could be an instance. In either case, the result is the same: an improved iambic parse on the surface.

(15) *Long vowel surfaces in unexpected position in reflexive repetitive third-person plural forms* (Jaker & Cardinal 2020: 85, 218)

	Underlying form	Surface form	Expected form
a.	/ʔete-he-ɣe-t-ts'ə́r/	(ʔe.ˈteː)(he.ˈts'ə́r)	*(ʔe.ˈte)(ˈheː)(ˈts'ə́r)
	'they scratched themselves (repeatedly)'		
b.	/ʔete-he-ɣe-t-kʰá:r/	(ʔe.ˈteː)(he.ˈkʰá:r)	*(ʔe.ˈte)(ˈheː)(ˈkʰá:r)
	'they slapped themselves (repeatedly)'		

3.4.2 Stress conditions movement of tone

It has been previously observed that in some languages high tone attracts stress, an observation which DeLacy (2002, 2007) has formalised as the Tone-to-Stress Principle, or TSP. In Tetsǫ́t'ıné, most cases of potential mismatches between stress and tone are repaired by moving the position of stress, as we will see in §4. However, there is a minority of cases where stress–tone mismatches are repaired by moving a high tone one syllable leftwards from its otherwise expected position. This seems to occur only in reflexive semelfactive verbs (meaning 'to do to oneself once') (Jaker and Kiparsky 2020). Reflexive semelfactive forms contain the conjugation marker *hí* in the imperfective and optative (Jaker and Cardinal 2020: 113–114). In singular forms, where this conjugation marker falls on the second syllable from the left edge of the prosodic word (and thus in the strong position of an iambic foot), high tone surfaces in its expected position, as shown in (16).

(16) *High tone surfaces in expected position in reflexive semelfactive singular forms* (Jaker & Cardinal 2020: 100, 110, 217–218)

	Underlying form	Surface form	Expected form
a.	/ʔete-**hí**-t-tθ'í:/	(ʔe.ˈtí)(ˈtθ'í:)	(same)
	'he pinches himself (once)'		
b.	/ʔete-**hí**-ɣu-t-tθ'í:/	(ʔe.ˈtú)(ˈtθ'í:)	(same)
	'he will pinch himself (once)'		
c.	/ʔete-**hí**-t-ɬ-t'u:s/	(ʔe.ˈtíl)(ˈt'u:s)	(same)
	'he punches himself (once)'		
d.	/ʔete-**hí**-ɣu-t-ɬ-t'u:s/	(ʔe.ˈtúl)(ˈt'u:s)	(same)
	'he will punch himself (once)'		

On the other hand, in the corresponding plural forms, high tone occurs one syllable to the left of its expected position, as shown in (17). The result of this, on the surface, is that high tone still falls on the second syllable from the left edge of the prosodic word, a stressed position, by left-to-right iambic foot parsing.[2]

---

[2]The *h* of third-person plural *he* and the *h* of semelfactive *hí* behave differently morphophonemically, in that the former is resistant to deletion intervocalically, whereas the latter is not. Jaker (to appear) suggests that the *h* of *he* is actually /x/ underlyingly. Alternatively, it is possible that the *h* of *he* is underlying, whereas the *h* of *hí* is epenthetic.

(17) *High tone surfaces in unexpected position in reflexive semelfactive plural forms*
    (Jaker & Cardinal 2020: 100, 110, 217–218)

	Underlying form	Surface form	Expected form
a.	/ʔete-he-**hí**-t-tθ'í:/	(ʔe.ˈté)(he.ˈtθ'í:)	*(ʔe.ˈte)(ˈhǐ:)(ˈtθ'í:)
	'they pinch themselves (once)'		
b.	/ʔete-he-**hí**-ɣu-t-tθ'í:/	(ʔe.ˈté)(hu.ˈtθ'í:)	*(ʔe.ˈte)(ˈhǔ:)(ˈtθ'í:)
	'they will pinch themselves (once)'		
c.	/ʔete-he-**hí**-t-ì-t'u:s/	(ʔe.ˈté)(hel.ˈt'u:s)	*(ʔe.ˈte)(ˈhǐ:l)(ˈt'u:s)
	'they punch themselves (once)'		
d.	/ʔete-he-**hí**-ɣu-t-ì-t'u:s/	(ʔe.ˈté)(hul.ˈt'u:s)	*(ʔe.ˈte)(ˈhǔ:l)(ˈt'u:s)
	'they will punch themselves (once)'		

If we compare the expected forms in (17) to the actual forms, we also notice that the actual surface forms also involve vowel length adjustment. That is, in all of the expected forms, the third syllable contains a long vowel, whereas in the actual surface forms the vowel is shortened. This is also expected in an iambic system, since the third syllable is an odd-numbered syllable, which falls in the weak position of a foot by left-to-right iambic foot parsing. Unlike what we saw in §3.4.1, however, the length is not transferred to the preceding syllable; rather, the extra mora is simply deleted.

3.4.3 Stress conditions intervocalic consonant deletion

Finally, we will consider cases where a consonant is deleted intervocalically for prosodic reasons, specifically, as we shall see, to avoid a stress lapse. The initial consonant *n* of the *ne* qualifier prefix is normally preserved following all other prefixes, including the third-person plural subject prefix *he*. An example is the verb *neɬʔĩ:* 'see' shown in (18); we include that both the surface forms and a metrical parse into iambic feet.

(18) *Imperfective paradigm of* neɬʔĩ: *'see', with iambic parse* (Jaker & Cardinal 2020: 133)

	Underlying form	Surface form	Iambic parse	Subject
a.	/ne-s-ɬ-ʔĩ:/	nesʔĩ:	(nes.ˈʔĩ:)	1SG
b.	/ne-ñe-ɬ-ʔĩ:/	nĩɬʔĩ:	(nĩɬ.ˈʔĩ:)	2SG
c.	/ne-ɬ-ʔĩ:/	neɬʔĩ:	(neɬ.ˈʔĩ:)	3SG
d.	/ne-hít-ɬ-ʔĩ:/	nílʔĩ:	(ˈníl)(ˈʔĩ:)	1PL
e.	/ne-uh-ɬ-ʔĩ:/	nuɬʔĩ:	(ˈnuɬ)(ˈʔĩ:)	2PL
f.	/he-ne-ɬ-ʔĩ:/	heneɬʔĩ:	(he.ˈneɬ)(ˈʔĩ:)	3PL

Recall that in Tetsǫ́t'ıné, all consonants are moraic word-finally. Based on evidence from the morphophonemics of optative paradigms (Jaker and Cardinal 2020: 104–111), it appears that the coda consonants of the first-person-plural prefix *hít* and the second-person-plural prefix *uh* also count as moraic, whereas all other coda consonants in prefixes count as non-moraic. For this reason, the coda consonants in (18d) and (18e) are treated as moraic, whereas the other coda consonants are treated as non-moraic in these examples. The main point is that all of the examples in (18) can be parsed into a sequence of well-formed iambic feet. In verbs with additional

syllables, however, this is not always possible. Consider the perfective paradigm of the verb *háútenelt^hən* 'learn', shown in (19).

(19) *Perfective paradigm of* háútenelt^hən *'learn', with iambic parse* (Jaker's field notes, 17 July 2020)

Underlying form	Surface form	Iambic parse	Subject
a. /háú-te-ne-s-l-t^hən/	háútenest^hən	(ˈháú)(te.ˈnes)(ˈt^hən)	1SG
b. /háú-te-ne-ñe-l-t^hən/	háútenĩlt^hən	(ˈháú)(te.ˈnĩl)(ˈt^hən)	2SG
c. /háú-te-ne-Hθ-l-t^hən/	háútenéɬt^hən	(ˈháú)(te.ˈnéɬ)(ˈt^hən)	3SG
d. /háú-te-ne-híd-l-t^hən/	háútenílt^hən	(ˈháú)(te.ˈníl)(ˈt^hən)	1PL
e. /háú-te-ne-uh-l-t^hən/	háútenuɬt^hən	(ˈháú)(te.ˈnuɬ)(ˈt^hən)	2PL
f. /háú-te-he-ne-Hθ-l-t^hən/	háútehě:ɬt^hən	(ˈháú)(te.ˈhě:ɬ)(ˈt^hən)	3PL

In all of the surface forms in (19), the output is parsed into the same, well-formed sequence of iambic feet of the form (heavy)(light–heavy)(heavy). However, in (19f) this is achieved by deletion of *n* intervocalically. If *n* were retained in this form, given that high tone attracts stress in the language, the predicted output would be *(ˈháú)te (he.ˈnéɬ)(ˈt^hən), with a stress lapse between the second and third syllables. This deletion of *n* in the third-person plural form is obligatory in the imperfective, perfective and optative paradigms of the verbs 'teach' and 'learn'. It also occurs variably in verbs which use the PERSISTIVE derivational string (Jaker and Cardinal 2020: 194–204). An example is the verb *níníjú:* 'bring dogs to a place'; the imperfective forms are given in (20).

(20) *Imperfective paradigm of* níníjú: *'bring dogs to a place', with iambic parse* (Jaker & Cardinal 2020: 196)

Underlying form	Surface form	Iambic parse	Subject
a. /ní-ne-hí-s-jú:/	nínífú:	(ní.ˈní)(ˈfú:)	1SG
b. /ní-ne-hí-ɲe-jú:/	nínĩ́jú:	(ní.ˈnĩ́)(ˈjú:)	2SG
c. /ní-ne-hí-jú:/	níníjú:	(ní.ˈní)(ˈjú:)	3SG
d. /ní-ne-hí-hít-jú:/	nínítfú:	(ní.ˈní)(ˈtfú:)	1PL
e. /ní-ne-hí-uh-jú:/	nínúhfú:	(ní.ˈnúh)(ˈfú:)	2PL
f. /ní-he-ne-hí-jú:/	níhǐ:jú:	(ní.ˈhǐ:)(ˈjú:)	3PL

There does not seem to be any likely alternative phonological or morphological explanation for this consonant deletion other than to adhere to a left-to-right iambic prosodic pattern. And if there are cases where consonants are deleted to improve iambic foot parsing, this is one more piece of evidence that iambic stress is phonologically active in the language.

3.4.4 Summary

In this section, we have seen three types of evidence that stress is active in the phonological grammar of Tetsǫ́t'ıné: vowel length adjustments, movement of tone and intervocalic consonant deletion. These types of evidence are significant, because it has been claimed that in a language with contrastive tone, stress can be represented only 'covertly' on the length tier (Spahr 2016: 200). The idea of 'covert' stress implies that stress is somehow read off of other phonological structure but does not operate

upon that structure. In this section, however, we have seen that metrical stress actively manipulates phonological material on the tone, moraic and segmental tiers. This would seem to suggest that in Tetsǫ́t'ıné stress exists on its own representational tier (the metrical grid) and plays a fully active role in the phonological computation.

More generally, we have seen evidence that Tetsǫ́t'ıné has a contrastive vowel length opposition (§3.2) and four contrastive tones (§3.3) in addition to an iambic stress system. At this point, we are left with an odd juxtaposition: in this language, stress must be audible enough to play an active role in the phonology and be acquirable by learners, while at the same time not neutralising the vowel length and tone contrasts that also exist in the language. In §4–6 we will describe an acoustic study we conducted to see what combination of acoustic stress correlates Tetsǫ́t'ıné employs to make this rather unusual combination possible.[3]

4. Hypothesis

The basic hypothesis we sought to investigate was that Tetsǫ́t'ıné exhibits iambic stress realised phonetically as a combination of increased duration, increased amplitude and increased F0 of the stressed vowel when compared with neighbouring unstressed vowels. More technically speaking, we sought to reject the null hypothesis that the position of stress has no effect on the relative duration, amplitude or F0 of adjacent vowels. To formulate this hypothesis precisely, however, it is first necessary to define exactly which syllables were predicted to be stressed under our hypothesis. Specifically, both heavy syllables and high tone attract stress in Tetsǫ́t'ıné; therefore, it is necessary to specify the set of tone patterns and weight patterns used in the stimulus set.

In designing our stimulus set, the first step was to exclude syllable weight from the set of variables under investigation. Since, as discussed previously, some coda consonants in prefixes are moraic, this meant excluding all words with a medial consonant cluster. Thus, all target words were of the form CVCV:(C) for two-syllable words, and CVCVCV:(C) for three-syllable words. Regarding vowel length, we selected words where all of the prefix vowels were short, and the stem vowel, full (long). This resulted in a (light–heavy) weight pattern for all two-syllable words, and a (light–light)(heavy) weight pattern for all three-syllable words. While it would have been more desirable from an experimental design point of view to have all light syllables, this would have been impossible, given the prosodic morphology of the language, since all stems (the final syllable) are obligatorily heavy (Jaker and Cardinal 2020: 73). Since the stem syllable must be heavy regardless, we chose to use all full stem vowels, rather than all reduced stem vowels, since full vowels in stems are much more common. Finally, we allowed for an optional final coda consonant to increase the number of possible lexical items we could use. The end result was that the weight pattern was held constant, with a (light–heavy) pattern for all two-syllable words and a (light–light)(heavy) pattern for all three-syllable words.

---

[3]The raw measurements from our study, as well as a markdown file summarising our statistical results, are available as supplementary material accompanying this article. To protect the privacy of our participants, we have not made the original recordings public; these are archived with the Yellowknives Dene First Nation.

The hypothesised locations of stresses were based primarily on distributional evidence, as illustrated in §3.4, wherever such evidence was applicable. In cases where no distributional evidence was applicable (such as the question of whether the final syllable is stressed in (22g)–(22h)), whether a syllable was hypothesised to be stressed or unstressed was based on Jaker's subjective impression of stress; Jaker is not a native speaker of the language. We propose that Tetsót'ıné stress is iambic by default. Because high tone attracts stress, in accordance with the TSP (DeLacy 2002, 2007), however, a high–low tone pattern will result in a trochee in two-syllable words. This is illustrated in (21), where we provide examples of all four logically possible tone patterns in two-syllable words.

(21) *Tone patterns and resulting stress patterns for two-syllable words* (Jaker & Cardinal 2020: 76, 171–172, 254)

	Tone pattern	Predicted stress pattern	Example	
a.	Low–Low	unstressed–stressed	(he.ˈkʰe:)	'they (du.) sit'
b.	Low–High	unstressed–stressed	(hu.ˈsá:)	'I will leave'
c.	High–Low	stressed–unstressed	(ˈhí.tã:)	'we drink'
d.	High–High	unstressed–stressed	(hí.ˈté:l)	'we (pl.) leave'

As shown in (21), our analysis predicts an unstressed–stressed pattern for the (low–low), (low–high), and (high–high) tone patterns in (21a), (21b) and (21d), and a stressed–unstressed pattern for the (high–low) tone pattern in (21c). For three-syllable words, the stress pattern of the first two syllables will follow the same patterns as in (21). As for the third syllable, it will be unstressed if it has a low tone and the preceding syllable has high tone; otherwise, it will be stressed. This is illustrated in (22).

(22) *Tone patterns and resulting stress patterns for three-syllable words* (Jaker & Cardinal 2020: 127–128, 141, 175–176, 179, 195)

	Tone pattern	Predicted stress	Example	
a.	Low–Low–Low	unstr.-str.-str.	(he.ˈne)(ˈje:)	'they grow'
b.	Low–Low–High	unstr.-str.-str.	(he.ˈne)(ˈkʰé:)	'they are sitting'
c.	Low–High–Low	unstr.-str.-unstr.	(he.ˈné)kʰe:	'they have sat down'
d.	Low–High–High	unstr.-str.-str.	(ni.ˈrĩ́)(ˈʔá:)	'pick up (heavy object)'
e.	High–Low–Low	str.-unstr.-str.	(ˈʃé.he)(ˈkʰĩ:)	'they (du.) eat'
f.	High–Low–High	str.-unstr.-str.	(ˈná.he)(ˈté:)	'they (pl.) live'
g.	High–High–Low	unstr.-str.-unstr.	(ní.ˈní)te:l	'we (pl.) got stuck'
h.	High–High–High	unstr.-str.-str.	(ní.ˈní)(ˈté:l)	'we (pl.) get stuck'

As shown in (22), under our hypothesis, the majority of tone patterns result in a stress clash between the second and third syllables. A stress clash does not occur in (22e) or (22f), where the second syllable is unstressed, or in (22c) or (22g), where the final syllable is unstressed. We suggest that the reason why the final syllable is unstressed in (22c) and (22g) is that a low-toned syllable becomes unstressed following a high-toned syllable (DeLacy 2002), although a formal analysis of this phenomenon is beyond the scope of this paper.

Two other issues require comment at this stage: the role of morphology and the role of primary versus secondary stress. Although our experimental design did not

distinguish between primary and secondary stress, Jaker's subjective impression is that where there is more than one stress in a prosodic word, the rightmost stress is the primary stress.[4] Regarding morphology, it should be noted that in all of the conditions in (21) and (22), the final syllable is also the stem syllable, which is, according to (21) and (22), stressed in three-quarters of the conditions. This raises the question of whether, when the final syllable is stressed, this stress can be attributed to its morphological status as the stem syllable rather than to its status as the head of an iambic foot. It is clear that morphology cannot be the sole factor which determines stress, since this would not explain why the final syllable is unstressed in cases of a final high–low tone pattern (as in (21c), (22c) and (22g)). In addition, morphological factors would also not explain why the second syllable of three-syllable words is stressed in all conditions except (22e) and (22f). Nevertheless, we do believe that morphology is relevant to the stress system, in that it is probably not an accident that a language in which stems are word-final has developed an iambic stress pattern, in which, by default, the final syllable is also stressed. Indeed, it has been observed that the more prefixing a language is, the higher is the likelihood of it exhibiting final stress (Gordon 2006: 209–213). The fact that morphological and prosodic dimensions of prominence largely coincide makes the system as a whole more diachronically stable and could be seen as an example of what is referred to as 'harmonic alignment' (Prince and Smolensky 2004).

To summarise, Tetsǫ́t'ıné has contrastive tone and vowel length, as well as a phonologically active stress system. In designing our stimuli, we selected a constant weight pattern for all target words: (light–heavy) for two-syllable words and (light–light) (heavy) for three-syllable words. There is a complex interaction in the language between tone and stress. In some cases, as we saw earlier in §3.4.2, stress conditions movement of tone. In other cases, as we have seen in this section, tone attracts stress onto a different syllable than where it would be expected to appear, based on just left-to-right iambic foot parsing. Specifically, initial (high–low) tone sequences result in trochaic, rather than iambic feet. In sections 5 and 6, we will describe the acoustic study conducted to find acoustic evidence for the stress patterns we proposed in (21) and (22), which, we hypothesise, result from the different tone patterns shown.

5. Methods

5.1 Experimental design and stimuli

The target words for the experimental stimuli were all two- and three-syllable verbs, taken from Jaker and Cardinal's (2020) *Tetsǫ́t'ıné Verb Grammar*. As described

---

[4]We compared primary and secondary stress on the second syllable of words with only primary stress in the word against words with secondary stress that falls on the second syllable to determine any acoustic differences between the two types of stress. We used a linear model to compare Duration and F0, and a GAMM to investigate intensity. We did not observe a significant effect of F0 ($p = 0.297$), duration ($p = 0.931$), or intensity (interval: primary stress, $p = 0.641$; interval: secondary stress, $p = 0.997$). Thus while it is still possible that there is an acoustic difference between primary and secondary stress, we were not able to find evidence for it in the present study.

previously, when selecting the target words we excluded any words with word-medial consonant clusters, words with contour tones and words with non-final long vowels. The result is that all of the target words have only level tones (high or low) on all of the syllables, and all of the words have a constant weight pattern (light–heavy) for two-syllable words, and (light–light)(heavy) for three-syllable words.

With these factors controlled for, the independent variable was stress, whereas the dependent variables were the set of stress correlates (amplitude, intensity and F0). However, given that tone attracts stress, the only way to manipulate the position of stress was to manipulate tone, as shown in (21) and (22). Therefore, our goal was to find phonetic evidence for the stress patterns resulting from each possible tone pattern, in two- and three-syllable words. In two-syllable words, there are 2^2 or four possible tone patterns, as shown in (21), whereas in three-syllable words, there are 2^3 or eight possible tone patterns, as shown in (22). We selected a total of 36 two-syllable words and 28 three-syllable words as stimuli, which are listed in the online supplementary material.

The stimuli were shown to subjects in the form of a PowerPoint presentation, the word being written out in the standard orthography as part of a carrier sentence, along with a picture to illustrate the sentence. The investigator first read the sentence aloud, then asked participants to repeat the sentence twice, and then repeat the target word by itself three times, for a total of five repetitions of the target word (the target word was highlighted in bold in the written stimuli). It was necessary for the investigator to read the sentence out loud because not all participants were proficient in the standard orthography. We acknowledge the potential risk of this procedure; namely, that the pronunciation of the investigator (who is not a native speaker) might influence the participants. However, in Jaker's experience, speakers do not hesitate to correct perceived errors on the part of the investigator; for example, if a sentence is either ungrammatical or not culturally appropriate (see below). It is thus reasonable to expect that native speakers would correct any mispronunciations on the part of the investigator to make them conform to a more native-like pronunciation, in particular, a native-like stress pattern.

In designing these stimuli, and especially the carrier sentences, we gave priority to creating sentences that were culturally appropriate and provided a natural context in which the target word could be used. An example is given in Figure 1. The reason for this is that speakers, especially those who are not accustomed to working with linguists, will often reject a word as ungrammatical if it is not provided in such a context (indeed, this happened several times in the current experiment, in spite of our efforts just mentioned). In a number of cases, the speaker chose to change the target word provided to a different word. In these cases, our policy was that if the new word conformed to the conditions outlined above (same weight pattern, no contour tones and no consonant clusters), then the new word was retained. If, on the other hand, the new word provided by the speaker contained a different weight pattern, a contour tone, and/or a word-medial consonant cluster, the word was excluded from our measurement set.

One disadvantage of the approach outlined above is that it does not control for many of the things which carrier sentences are usually used to control for; for example, boundary tones, and the effect of the tone, stress or weight of adjacent

Dëne łą
Beghúldesche
náhedé.

Figure 1. *Example of experimental stimulus. Translation: 'Many people live in Yellowknife.'*

words. However, we have found this approach necessary, based on past experience that subjects require all target words to be provided in a natural context.

We interviewed a total of four subjects in Yellowknife and Łútsëlk'é, Northwest Territories, three female speakers and one male speaker. All four speakers were in their 50s or 60s, and also spoke English fluently. Three of the four speakers had basic familiarity with the standard orthography of the language, whereas one did not. However, this speaker was still able to complete the task by repeating the sentence uttered by the investigator.

Finally, we will briefly mention the question of vowel quality. It is well known that different vowel qualities, ([i], [e], [a], [o], [u]), are associated with different inherent durations (House 1961; Toivonen *et al.* 2015) and different inherent intensities. Ideally, one would want to hold vowel quality constant, just as we did with weight, to avoid this potential confound. Unfortunately, given that verbal prefixes constitute a relatively small, closed set in Dene languages, we would not have been able to find target words to illustrate all of the tone patterns in (21) and (22), if we had, for example, restricted ourselves to only words with *e* (as in *heneje:* 'they grow' or *henét^he:s* 'they are sleeping'). Therefore, it was necessary to include words with different vowel patterns; for example, *náθija:* 'I went'. However, the effect of these different vowel qualities was accounted for post hoc as part of our statistical analysis (see §5.4).

5.2 Recording equipment

Subjects were recorded using a Marantz PMD 671 Compact Flash Recorder, recording at 24 bits and 44 kHz, using two cardioid condenser microphones placed approximately eight inches from the subjects. Recordings took place at the Yellowknives Dene First Nation Land and Environment Office in Yellowknife, and at the Co-op Bed and Breakfast in Łútsëlk'é. The recording environment was not soundproof,

and there was occasional background noise; a small number of tokens needed to be thrown out for this reason.

5.3 *Segmentation and measurements*

Segmentation was done in Praat (Boersma and Weenink 2020), and all segments measured were vowels. When defining the beginning and the end of a vowel, the general principle was that transitions were part of the vowel. Thus, for example, in a sequence such as [aja] or [ana], only the region of relatively stable formant structure was counted as part of the consonants [j] or [n]; the transitions in and out of these consonants were counted as part of the neighbouring vowels. An illustration of segmentation involving sonorant and glide consonants is given in Figure 2, using the form *heneje* 'they grow'.

Our decision to include target words with nasal, liquid and glide consonants was for the same reason that we included words with different vowel qualities: the need to achieve broad descriptive coverage. That is, there would not have been enough stimuli illustrating all the different tone patterns to measure, based on Jaker and Cardinal (2020), if we had restricted ourselves to only words with intervocalic obstruents. Thus, while the inclusion of intervocalic sonorants does introduce some potential for segmentation error, due to the inherent indeterminacy in demarcating sonorants, we feel that this disadvantage is balanced by the possibility of creating an exhaustive phonetic description of the stress system; that is, measurements of all possible tone and stress patterns.

There were also some additional considerations when demarcating stops and affricates. For the transition from a vowel into a stop or affricate, we counted the end of complex formant structure as the end of the vowel. For the transition from a stop or
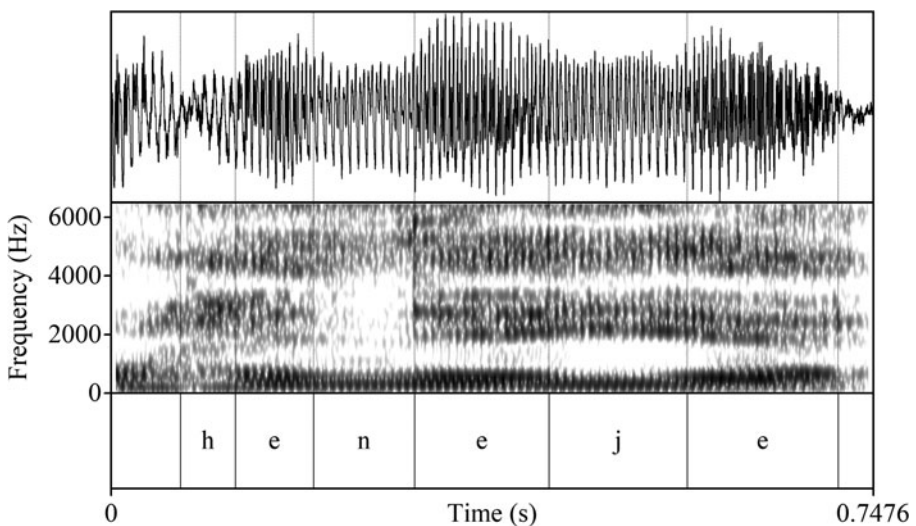
Figure 2. *Segmentation involving intervocalic* n *and* j

affricate into a vowel, it was also necessary to consider laryngeal features. For plain stops [t, k], the vowel begins immediately following the release burst; and for ejective stops [t', k'], immediately following the glottal release. However, it has been noted that aspirate stops in Dene languages are accompanied by a great deal of frication noise, such that they might almost be transcribed as affricates [t^x, k^x] rather than aspirates [t^h, k^h] (McDonough and Wood 2008). For this reason, following both aspirate stops and affricates, we counted the period of aspiration and/or frication noise as belonging to the consonant, and the beginning of complex formant structure as the beginning of the vowel. Similarly, with transitions in and out of fricatives, we counted the end of complex formant structure as the end of the vowel, and the beginning of complex formant structure as the beginning of the vowel.

Duration and intensity measures were extracted from each vowel in each syllable. Duration measures were taken in seconds. Intensity was extracted at ten equally spaced points over the duration of each vowel to facilitate dynamic comparisons of intensity for stressed and unstressed vowels, and was measured in decibels. F0 was measured in Hertz. For F0, we chose to use mean values rather than dynamic measurements, because there were a substantial number of unidentified pitch values in the output script.

5.4 Statistical methods

Our statistical analysis primarily made use of linear models. The duration of the vowels in each syllable was compared statistically using linear models using the lmer() function (Bates *et al*. 2015) in R (R Core Team 2021), and degrees of freedom were calculated using the Satterthwaite's (1946) method with the lmerTest package (Kuznetsova *et al*. 2017). We performed a total of two linear models, one for two-syllable words and one for three-syllable words. Each model had the fixed effects of stress (two levels: unstressed and stressed), tone (two levels: high and low), vowel quality (seven levels: a, ã, e, i, ĩ, u, ui), syllable (two levels: non-final and final), and an interaction between stress and syllable. We included random effects of participant and item (i.e., which word from the list the data was extracted from). Each of the random effects were coded as a random intercept in the model. Post-hoc tests were performed with the emmeans() package (Length 2020). We generated dynamic intensity plots using generalised additive mixed models (GAMMs). With GAMMs, we can specify random smooth terms to account for differences across participants and words when calculating the group trends. We used the mgcv package (Wood 2011) in R. We included fixed effects of stress (two levels: stressed and unstressed), and we included smoothing terms for interval, and interval by stress. We included factor smooths (i.e. random effects) for interval: stress by participant and interval by word (i.e. which word the intensity was extracted from).

We also performed an F0 analysis for words with flat tone patterns (e.g. low–low or high–high–high) to see if stress had an effect. We compared the mean F0 in each syllable for each of the four word types (two linear models: two syllables, three syllables). Each of the linear models had a main effect of syllable (two or three levels: syllable one, syllable two and syllable three), tone (two levels: low tones and high tones), and vowel quality (six levels: e, i, u, ui, ã, ĩ; five levels: i, u, a, ã, ĩ) and a

random effect of participant and of item (i.e. the word that data was extracted from). The random effects were coded as a random intercept in the model. However, in the three-syllable model, the model produced a singular fit, so we removed the random intercept for item. Linear discriminant analysis (LDA) was performed using flipMultivariates (Displayr 2021). For the LDA, we examined the difference between stressed and unstressed syllables to determine the cue weighting. We performed two LDA analyses with the package flipMultivariates: one for two-syllable words and one for three-syllable words. We coded duration, intensity and F0 into the model. We used the maximum intensity value for each word for each participant, and the mean F0 values. Before performing the LDA, we converted all the values into z-scores to control for inter-speaker differences in speaking rate and intensity. Statistical analysis was done with a two-tailed model, multiple comparisons correction, and false-discovery-rate correction applied to the entire table simultaneously. All plots were generated using the ggplot2 package (Wickham 2016).

6. Results

6.1 Duration results
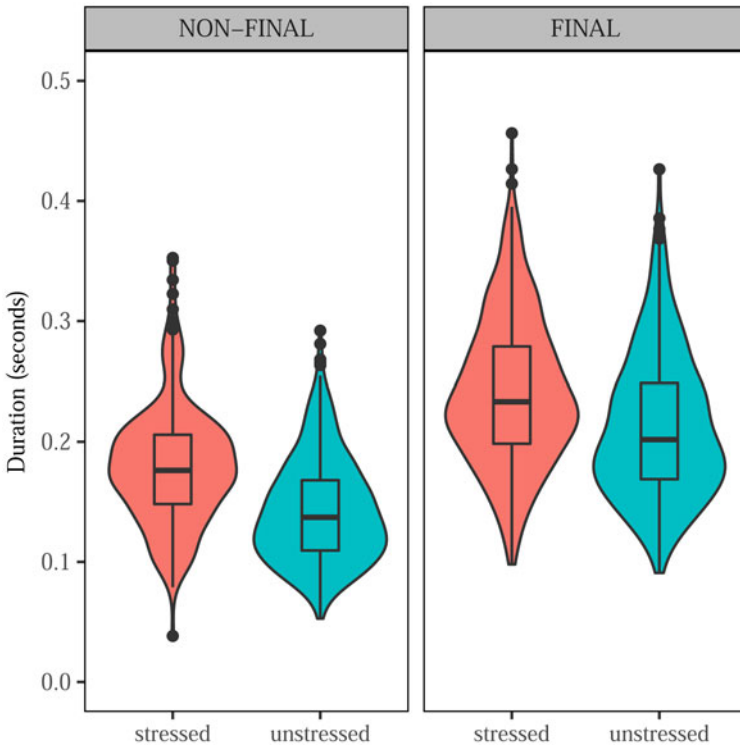
6.1.1 Two-syllable words

The linear model for two-syllable words ($n = 1{,}543$) revealed a main effect of stress ($F_{(1, 554)} = 36.37$; $p < 0.001$), syllable ($F_{(1, 702)} = 171.85$; $p < 0.001$), tone ($F_{(1, 203)} = 19.15$; $p < 0.001$) and vowel ($F_{(7, 932)} = 9.75$; $p < 0.001$), but no interaction between stress and syllable ($F_{(1, 44)} = 0.00$; $p = 0.950$). The R^2 for the model was 0.660. The post-hoc tests revealed that stressed syllables (*M*: 0.220; SD: 0.065) were longer than unstressed syllables (*M*: 0.167; SD: 0.060; $t_{(618)} = 5.984$; $p < 0.001$). See Figure 3 for violin plots of the non-final and final syllables, and Table I for a summary of the fixed effects.

The linear models for two-syllable words revealed that a correlate of stress is related to duration, as stressed syllables are longer than unstressed syllables in both the first and second syllable position.

6.1.2 Three-syllable words

The results of the linear model of three-syllable words ($n = 1{,}526$) revealed a main effect of stress ($F_{(1, 1429)} = 100.56$; $p < 0.001$), syllable ($F_{(1, 1487)} = 546.81$; $p < 0.001$) and tone ($F_{(1, 1353)} = 25.13$; $p < 0.001$), but no main effect of vowel ($F_{(6, 1329)} = 1.27$; $p = 0.265$) and no interaction between stress and syllable ($F_{(1, 1452)} = 0.73$; $p = 0.392$). The R^2 for the model was 0.641. Post-hoc analysis of stress revealed that stressed syllables (*M*: 0.182; SD: 0.062) were longer than unstressed syllables (*M*: 0.126; SD: 0.052; $t_{(1448)} = 10.01$; $p < 0.001$). Table II presents the fixed effects for the linear model for two-syllable words, and Figure 4 presents a violin plot of the stressed and unstressed vowels by syllable position.

The linear models of the three-syllable words revealed that in each case the stressed syllable had a longer duration than unstressed syllables, suggesting this is a correlate of stress in this language.

Figure 3. *Violin plot of the duration for unstressed and stressed vowels for syllable position (Non-Final [left] and Final [right]) in two-syllable words*

Table I. *Summary of fixed effects for the linear model of duration in two-syllable words*

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	0.251	0.018	4	14.094	< 0.001
Stress (unstressed)	−0.025	0.008	66	−3.111	0.003
Syllable (Non-Final)	−0.006	0.008	70	−7.367	< 0.001
Tone (Low)	−0.025	0.006	20	−4.376	< 0.001
Vowel (ã)	0.017	0.006	115	2.787	0.005
Vowel (e)	−0.008	0.005	108	−1.575	0.117
Vowel (i)	−0.014	0.007	88	−2.079	0.038
Vowel (ĩ)	0.001	0.007	101	1.877	0.061
Vowel (u)	−0.001	0.008	118	−1.513	0.131
Vowel (ui)	−0.016	0.023	59	6.945	< 0.001
Vowel (ũ)	−0.019	0.016	86	−1.187	0.236
Stress (unstr.) : Syllable (Non-Fin.)	0.001	0.014	45	0.064	0.950

Table II. *Summary of fixed effects for the linear model of duration in three-syllable words*

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	0.227	0.014	3.77	15.760	< 0.001
Stress (unstressed)	−0.033	0.005	1351	−5.881	< 0.001
Syllable (Non-Final)	−0.065	0.003	1478	−22.236	< 0.001
Tone (Low)	−0.016	0.003	1353	−5.013	< 0.001
Vowel (ã)	−0.020	0.012	1480	−1.678	0.094
Vowel (e)	−0.006	0.004	1273	−1.579	0.115
Vowel (i)	−0.000	0.005	1351	−0.063	0.950
Vowel (ĩ)	0.001	0.004	1511	0.139	0.889
Vowel (ũ)	0.052	0.044	1090	1.168	0.243
Vowel (u)	−0.009	0.008	1412	−1.193	0.233
Stress (unstr.) : Syllable (Non-Fin.)	−0.004	0.006	1452	−0.856	0.392
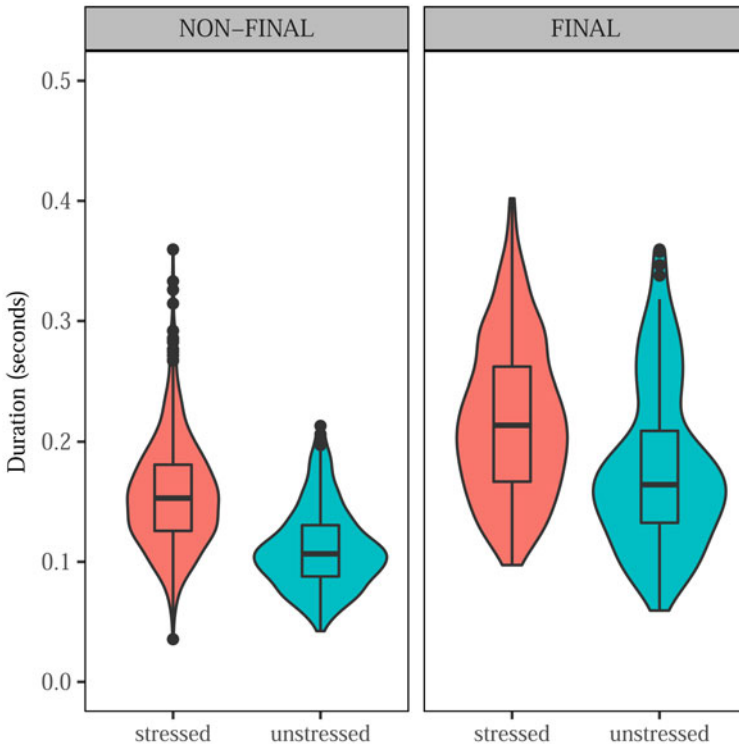
Figure 4. *Violin plot of the duration for unstressed and stressed vowels for syllable position (Non-Final [left] and Final [right]) in three-syllable words*

6.2 Intensity results

The GAMM results for the intensity examination of two-syllable words showed the same trend: in both the first (n = 7,899) and the second syllable (n = 7,569), peak intensity was higher for stressed syllables than unstressed syllables. However, there were subtle differences in the trajectory of the intensity contour. In the first syllable, unstressed and stressed syllables shared a similar sharp increase until approximately the midpoint of the vowel. However, there was a sharper drop in the stressed syllables' intensity, resulting in a margining in the intensity values. In the second syllable, there was also a sharp increase in intensity; however, the increase peaked at only approximately one-quarter of the duration of the vowel. There was then a sharp drop in intensity for unstressed and stressed vowels. The overall trajectories and drop in intensity resulted in maintenance of a significant difference between unstressed and stressed syllables. Figure 5 presents the GAMM results for the intensity analysis of two syllables. Syllable one is on the left, and syllable two is on the right. Table III presents the approximate significance of smooth terms for syllable one and syllable two along with the R^2 values.

The intensity examination also revealed significant differences between unstressed and stressed syllables in three-syllable words, as shown in Figure 6. In the first syllable (n = 5,206), intensity began higher and peaked higher for stressed syllables compared to unstressed syllables. However, after the midpoint of the vowel, the intensity for stressed syllables dropped, while the intensity for unstressed remained level. The drop in intensity resulted in an overall lower value for stressed syllables than unstressed syllables at the very end of the syllable. In the second syllable (n = 5,300), the intensity of stressed syllables was higher than that of unstressed syllables. However, they both shared a similar trajectory. There was a gradual increase in intensity until the midpoint of the vowel, when intensity began to drop. The drop was sharper in the stressed syllable but maintained a significant difference from the unstressed syllables over the entire duration. In the third syllable (n = 4,884), the unstressed and stressed syllables began with no significant difference. However, there was a longer increase in intensity for stressed
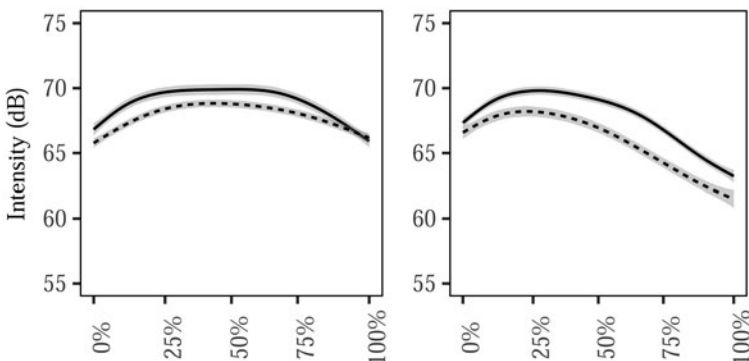
Figure 5. *Dynamic intensity plots for the first (left) and second (right) syllable in two-syllable words. The solid line indicates stressed syllables, and the dashed line indicates unstressed syllables. Grey shading indicates 95% confidence intervals*

Table III. *Approximate significance of smooth terms and R^2 for two-syllable words*

		edf	Ref.df	F	p-value	R^2
	Interval	6.558	7.535	70.658	< 0.001	
Syllable One	Interval: Stressed	1.271	1.727	2.75	0.0862	
	Interval: Unstressed	1.000	1.000	5.995	0.014	0.626
	Interval	7.050	7.821	67.687	< 0.001	
Syllable Two	Interval: Stressed	2.044	2.514	4.848	0.005	
	Interval: Unstressed	1.000	1.001	0.047	0.829	0.651

syllables, resulting in a higher-intensity contour. The peak was just after one-quarter of the vowel and began a decline at the midpoint of the vowel. The stressed syllables had a sharper decline, resulting in no significant difference between unstressed and stressed syllables. However, as with the two-syllable words, the final syllable (here, the third syllable) had much sharper and greater declines in overall intensity for both unstressed and stressed syllables. Table IV presents the approximate significance of smooth terms for syllable one, syllable two and syllable three, along with the R^2 values.

The GAMM analysis revealed differences in overall trajectories and values for intensity in stressed and unstressed syllables in both two- and three-syllable words. Further, those differences also existed for each syllable within each word type. However, the overall tendency was for a higher peak intensity for stressed syllables compared to unstressed syllables.

6.3 F0 results

The results of the linear model for two-syllable words (n = 795) with high–high and low–low tone revealed a main effect of tone (F(1, 33) = 88.43; p < 0.001) and vowel (F(6, 148) = 3.79; p = 0.002), but the effect of syllable did not reach
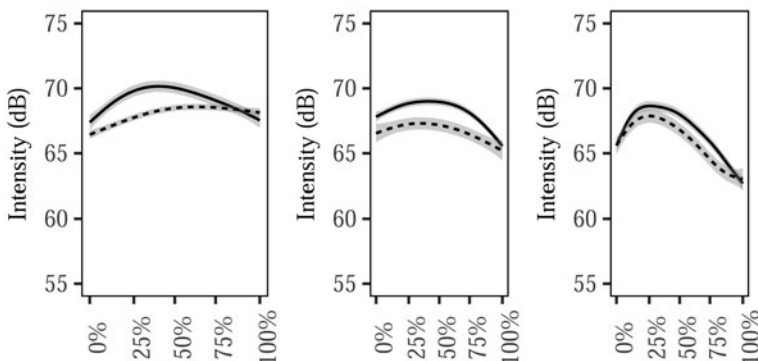
Figure 6. *Dynamic intensity plots for the first (left), second (middle) and third (right) syllables in three-syllable words. Solid lines indicate stressed syllables; dashed lines indicate unstressed syllables. Grey shading indicates 95% confidence intervals*

Table IV. *Approximate significance of smooth terms and R^2 for three-syllable words*

|		edf	Ref.df	F	p-value	R^2
	Interval	4.297	5.176	35.849	< 0.001	
Syll. One	Interval: Stressed	3.251	3.960	14.311	< 0.001	
	Interval: Unstressed	0.000	0.000	0.040	0.999	0.513
	Interval	5.267	6.498	68.312	< 0.001	
Syll. Two	Interval: Stressed	0.303	0.477	0.713	0.560	
	Interval: Unstressed	1.000	1.000	47.580	< 0.001	0.581
	Interval	6.486	7.499	117.770	< 0.001	
Syll. Three	Interval: Stressed	2.669	3.432	8.597	< 0.001	
	Interval: Unstressed	1.000	1.000	15.656	< 0.001	0.524

significance ($F(1, 360) = 3.23$; $p = 0.070$). The R^2 for the model was 0.571. Table V presents a summary of the fixed effects for the linear model for two-syllable words, and Figure 7 presents the violin plot of the mean F0 by syllable for low–low and high–high tone patterns.

The results of the three-syllable analysis ($n = 383$) for words with high–high–high and low–low–low tone, as shown in Table VI, revealed a main effect of syllable ($F(2, 370) = 13.16$; $p < 0.001$), tone ($F(1, 20) = 79.53$; $p < 0.001$) and vowel ($F(5, 68) = 4.44$; $p = 0.001$). The R^2 metric was 0.872. The post-hoc test revealed that syllable one (*M*: 157.95; SD: 39.46) had a higher F0 than syllables two (*M*: 153.30; SD: 37.81; $p = 0.001$) and three (*M*: 148.27; SD: 38.94; $p < 0.001$), but that there was no significant difference between syllable two and syllable three ($p = 0.766$).

The F0 analysis revealed that syllable position did not play a major role in F0. In fact, only the first syllable in the three-syllable context showed an increase in F0 based on position, suggesting that F0 falls as words increase in length. Thus, the data suggests that F0 is not a strong correlate of stress.

Table V. *Summary of fixed effects for the linear model of F0 in two-syllable words*

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	174.503	13.949	4	12.510	< 0.001
Syllable (Syllable Two)	−6.047	3.315	360	−1.824	0.069
Tone (Low)	−31.663	3.367	33	−9.404	< 0.001
Vowel (e)	7.596	4.452	98	1.706	0.091
Vowel (i)	1.153	3.616	127	0.319	0.750
Vowel (u)	7.207	5.107	125	1.411	0.161
Vowel (ui)	20.362	7.318	107	2.783	0.006
Vowel (ã)	16.607	6.691	216	2.482	0.014
Vowel (ĩ)	−19.036	27.604	755	−0.690	0.491
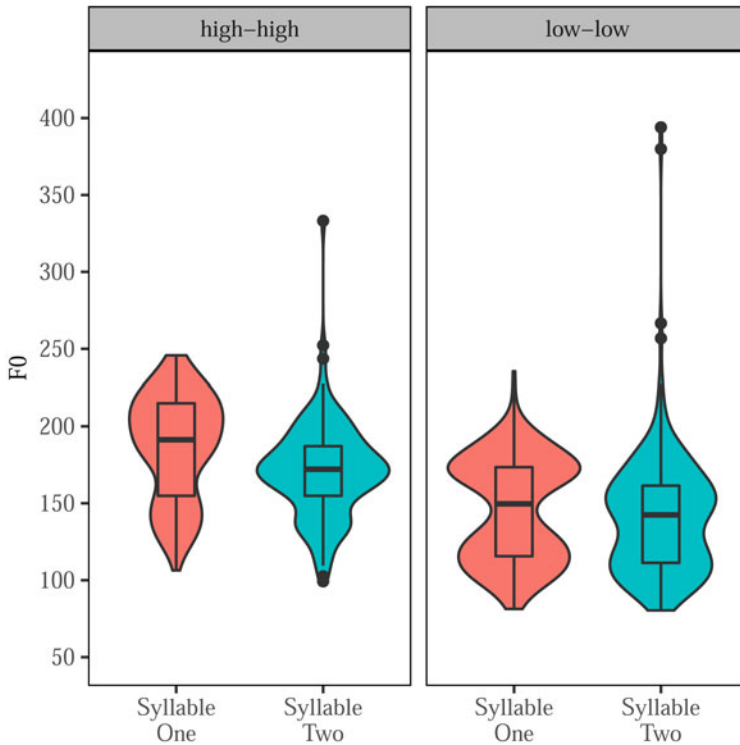
Figure 7. *Violin plot of the mean F0 for high–high (left) and low–low (right) by syllable position*

6.4 Linear discriminant analysis: comparing duration, intensity and F0

The LDA for two-syllable words produced correct predictions 74.62% of the time (stressed: 74.34%; unstressed 74.9%). The LDA revealed that the strongest predictor of stress was duration ($R^2 = 0.19$), whereas both intensity ($R^2 = 0.06$) and F0 ($R^2 = 0.04$)

Table VI. *Summary of fixed effects for the linear model of F0 in three-syllable words*

	Estimate	Std. Error	df	t-value	Pr(>|t|)
(Intercept)	186.628	20.089	3	9.290	0.001
Syllable (Syllable Two)	−10.159	2.345	371	−4.332	< 0.001
Syllable (Syllable Three)	−11.810	2.546	371	−4.640	< 0.001
Tone (Low)	−39.413	4.420	20	−8.918	< 0.001
Vowel (ã)	−12.124	8.485	98	−1.429	0.156
Vowel (e)	0.983	3.658	15	0.269	0.792
Vowel (i)	−4.845	5.101	60	−0.950	0.346
Vowel (ĩ)	11.781	3.367	327	3.396	< 0.001
Vowel (u)	−6.429	6.094	62	−1.055	0.296

Table VII. *LDA results for the prediction of stressed and unstressed syllables in two-syllable words with variables duration, intensity, and F0. Mean values for each variable in the stressed and unstressed syllables, R-squared and p-values are presented*

	Stressed n = 799	Unstressed n = 785	R-squared	p-values
Duration	0.51	−0.32	0.19	< 0.001
Intensity	0.94	0.58	0.06	< 0.001
F0	0.30	−0.12	0.04	< 0.001

Table VIII. *LDA results for the prediction of stressed and unstressed syllables in three-syllable words with variables duration, intensity and F0. Mean values for each variable in the stressed and unstressed syllables, R-squared and p-values are presented*

	Stressed n = 904	Unstressed n = 668	R-squared	p-values
Duration	−0.07	−0.95	0.23	< 0.001
Intensity	0.76	0.46	0.03	< 0.001
F0	0.06	−0.12	0.01	< 0.001

were weaker predictors. Table VII presents the results of the LDA. The model correctly predicted stressed syllables 660 times (139 miscategorisations as unstressed), and unstressed syllables 586 times (199 miscategorisations as stressed).

The LDA for three-syllable words produced 73.60% correct predictions (stressed: 76.88%; unstressed 69.16%). The LDA revealed that the strongest predictor of stress was duration ($R^2 = 0.23$), whereas both intensity ($R^2 = 0.03$) and F0 ($R^2 = 0.01$) were weaker predictors. Table VIII presents the results of the of the LDA. The model correctly predicted stressed syllables 803 times (101 miscategorisations as unstressed), and unstressed syllables 487 times (181 miscategorisations as stressed).

The data revealed that the strongest correlate of stress was duration ($R^2 = 0.19$; 0.23). This was far stronger than intensity or F0, although all three variables do play a significant role in correlates of stress (all p < 0.001). In both cases, intensity was the second-strongest predictor, although only marginally (e.g. Intensity: 0.03 *vs.* F0: 0.01). Therefore, we can determine the following hierarchy for correlates of stress in this language: duration > intensity > F0.

It should be noted, however, that our LDA results for F0 are not directly comparable with the F0 results reported in §6.3. This is because the F0 results in §6.3 included only words with level tone sequences such as high–high or low–low. The LDA results in Table VII, however, include words with all possible tone patterns,

including, for example, high–low words with initial stress and low–high words with final stress.

7. Discussion

In the preceding sections, we saw that there was a statistically significant effect of stress on all three stress correlates in Tetsǫ́t'ıné, which is broadly consistent with the evidence presented earlier in §3 that stress plays an active role in morphophonemic alternations in this language. The effect of intensity was always significant and in the direction expected, as seen in Figures 5 and 6, where stressed syllables had greater intensity than unstressed syllables. The effect on F0, on the other hand, was not statistically significant in two-syllable words. In three-syllable words, the effect actually went in the opposite direction than expected, where we observed an overall gradual decline in F0 across all three syllables. Since stress is typically associated with higher F0 cross-linguistically, (DeLacy 2002, 2007; Gordon and Roettger 2017),[5] it is likely that other factors may have been responsible for the lowering of F0 in our results – most probably, the effect of phrase-final boundary tones. Thus, our results could be interpreted as showing a lack of effect of stress on F0 in this language. We also found a significant effect of stress on duration. Recall that our stimuli were designed such that all two-syllable words exhibited a (light–heavy) weight pattern, whereas three-syllable words exhibited a (light–light)(heavy) weight pattern, due to the fact that we chose stems containing long (full) vowels. However, this does not constitute a confound in our analysis, because the weight pattern was held constant throughout all the stimuli, and our analysis compared the same position (first, second or third syllable) when it was either stressed or unstressed. Thus, we found that for each position within the word, the same vowel (whether phonologically long or short) was longer when stressed than when unstressed.

We now return to the more general question with which we began this paper: How should we expect stress to be realised phonetically in a language which has both contrastive vowel length and four contrastive tones? That is, which of the stress correlates – duration, F0, or intensity – should we expect to be the primary stress correlate in this language? It is, of course, somewhat problematic to try to answer this question directly, because these three correlates are measured in three distinct units of measurement – milliseconds, decibels and Hertz – which are not directly comparable with each other. However, the discriminant analysis presented in §6.4 suggests that whereas all three stress correlates play a role in distinguishing stressed from unstressed syllables, duration is by far the most accurate predictor of stress in this language, with over ten times greater weight than the other two stress correlates. In other words, it appears that in Tetsǫ́t'ıné, duration is the primary correlate of stress.

We are aware of only two detailed phonetic studies of stress in a language with both contrastive vowel length and tone. The first is Potisuk *et al.*'s (1996) study of

---

[5]A reviewer notes that much of the higher F0, which is often attributed to stress, may actually be due to phrasal high-pitch accents (i.e. intonation; Beckman 1986). However, perceptual experiments have shown that high F0 has an effect similar to increased intensity in its effect on rhythmic grouping biases (Crowhurst 2018: 97), suggesting that, at least to some degree, greater F0 is intrinsically associated with stress.

the acoustic correlates of stress in Thai. Their main finding was similar to ours, in that duration was the primary correlate of stress in Thai. However, the authors found that change in the shape of F0 contours was a secondary cue to stress, whereas there was no significant effect of stress on intensity (Potisuk *et al*. 1996: 210). Just as in our study, the fact that F0 plays a minor role in realising stress was expected (since Thai has five contrastive tones), whereas the primary role of duration was unexpected (since Thai also has contrastive vowel length). Potisuk *et al*. (1996: 211) speculate that the different behaviours of F0 and duration may be due to the different functional loads of the two phonetic properties: there are far more minimal pairs involving tone in Thai than there are involving vowel length. It is not clear to us whether this line of explanation could be extended to Tetsǫ́t'ıné. That is, it is not obvious to us that in Tetsǫ́t'ıné, tone has a higher functional load than does vowel length, except in the very abstract sense that vowels contrast two degrees of length, but there are four contrastive tones. Additional lexicostatistical work on Tetsǫ́t'ıné would be necessary to properly evaluate this hypothesis. The other such study was by Everett (1998) on Pirahã. Pirahã has both contrastive vowel length and tone. Everett finds that the primary correlates of stress in this language are amplitude and duration, whereas F0 and vowel formant frequencies seem to play little or no role. Thus, both studies found that duration played a role in realising stress, in tone languages with vowel length, though only one study also found a significant effect of amplitude.

How might one interpret these results? It is possible, contrary to the Functional Load Hypothesis in (2), that there is actually a universal preference for realising stress using duration rather than F0. Based on a survey of 110 studies of 75 languages, (Gordon and Roettger 2017: 8) found that duration was 'the most successful marker of stress, distinguishing stress in 90% […] of the languages studied'. However, in line with the Functional Load Hypothesis, they also found that F0 was used to cue stress in only two out of nine tone languages examined (Gordon and Roettger 2017: 9), and that six out of seven tone languages used intensity as a marker of stress (Gordon and Roettger 2017: 10). Thus, our Tetsǫ́t'ıné results might be seen as typologically normal for a tone language, except that Gordon and Roettger's sample size of tone languages was small, and the authors also do not specify how many of the tone languages examined also had contrastive vowel length. A study which specifically addressed the Functional Load Hypothesis in relation to contrastive vowel length (Lunden *et al*. 2017) found no support for the Functional Load Hypothesis: that is, languages with contrastive vowel length were just as likely to employ duration as a correlate of stress as languages without contrastive vowel length.

In the remainder of this section, we will situate our results within a broader theoretical context. We will consider the predictions of Dispersion Theory (Flemming 1995, 2001), which is conceptually related to the Functional Load Hypothesis, as well as the Iambic–Trochaic Law (Kager 1993; Hayes 1995), and we will discuss how the predictions of these theories compare with our results.

The predictions of the Functional Load Hypothesis in (2) for Tetsǫ́t'ıné are fairly straightforward. Although the Functional Load Hypothesis posits a universal bias in favour of using F0 to realise stress, in a language in which both tone and vowel length are contrastive, it predicts that stress should be realised mainly by intensity. As we

have seen, our results follow from the Functional Load Hypothesis only to the extent that F0 does not seem to be a correlate of stress in Tetsǫ́t'ıné. The fact that duration is the primary correlate of stress, whereas intensity plays a secondary role, is unexpected.

Dispersion Theory (Flemming 1995, 2001) seems to make very similar predictions to the Functional Load Hypothesis and runs into similar challenges in relation to our data. Dispersion Theory proposes a unified model of phonetics and phonology in which phonological inventories and their phonetic realisations, as well as their phonological behaviours, can be explained based on three general principles, as described in (23).

(23) *General principles of Dispersion Theory* (based on Flemming 2001: 25)
    a. Maximise the number of contrasts (in any given context).
    b. Maximise the distinctiveness of contrasts.
    c. Minimise effort.

The majority of work that we are aware of within Dispersion Theory has focused on explaining the structure of vowel inventories (Schwartz *et al.* 1997; Padgett and Tabain 2005; Trudgill 2009; Becker-Kristal 2010; Hall 2011). Although the study by Padgett and Tabain (2005) examines the effect of stress on the Russian vowel space, we are not aware of any studies within the framework of Dispersion Theory which directly address the question of how stress itself should be realised, given other facts about the phonological inventory of a language. Nevertheless, we believe that if the principles in (23) are applied to the central question of this paper, a fairly clear prediction emerges. If we assume that vowel length and tonal contrasts are to be maintained in accordance with principle (23a), and that, by the same principle, stress must also be realised, then a conflict arises in the implementation of principle (23b). All other things being equal, to the extent to which stress is realised by increased F0, it makes stress more distinctive but tonal contrasts less distinctive (and conversely); likewise, to the extent to which stress is realised by increased duration, it makes stress more distinctive but vowel length contrasts less distinctive (and conversely). There is only one way to avoid this conflict, and this would be to employ a third stress correlate – intensity – which is not independently contrastive in the language. Thus, it seems clear to us that in a language such as Tetsǫ́t'ıné, in which both vowel length and tone are contrastive, Dispersion Theory predicts that intensity ought to be the primary correlate of stress.

Our results are only partially consistent with these predictions. As shown in Figures 7 and 8, the effect of stress on F0 was either non-significant or went in the opposite direction than expected, suggesting that stress itself may not be responsible for the effect on F0 in these data. This would be consistent with Dispersion Theory, in that we would not expect a large effect of stress on F0 in a language with four contrastive tones. On the other hand, the LDA in Tables VII and VIII suggests that duration is a more reliable predictor of stress in this language than intensity. From a Dispersion Theory perspective, this is surprising for a language which has contrastive vowel length. The only way in which Dispersion Theory could be reconciled with our results would be to invoke principle (23c), minimise effort. That is, it could be assumed that the realisation of stress as increased intensity inherently involves greater effort (perhaps due to the greater subglottal pressure required) than the realisation of
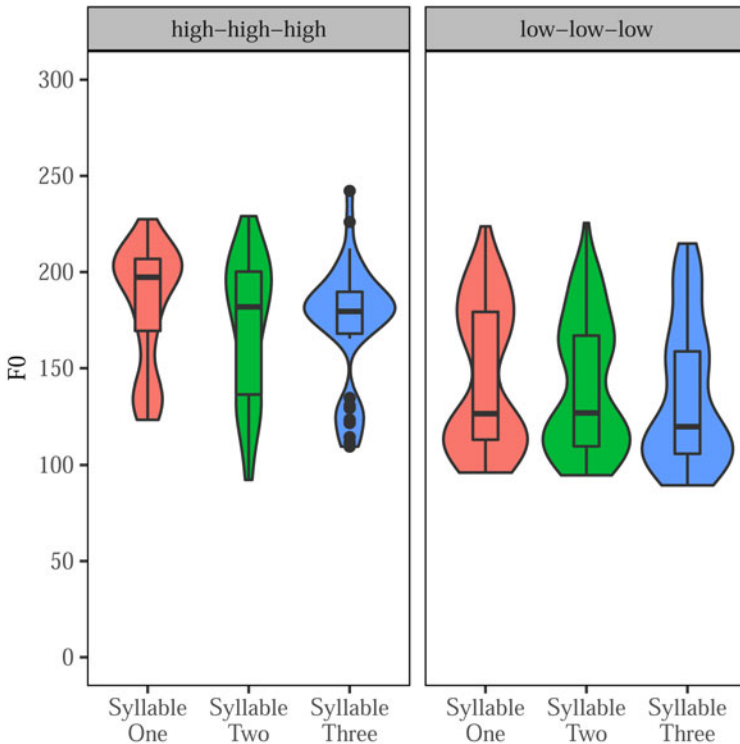
Figure 8. *Violin plot of the mean F0 for high–high–high (left) and low–low–low (right) by syllable position*

stress as F0 or duration. However, we are not aware of any means, at present, by which articulatory effort can be compared across different stress correlates. There is no way to determine, for example, that an increase in intensity of 3 dB requires greater effort than an increase in duration of 50 ms. Alternatively, a reviewer suggests another possible explanation based on perceptibility: while there are many languages in the world with durational contrasts, and many languages with tonal contrasts, there are no languages with a contrast based purely on intensity. It may thus be that intensity is inherently less perceptible than other phonetic correlates of stress.[6] However, this explanation also raises a similar issue of how to compare perceptibility across different stress correlates. Thus, while our results may still be compatible with Dispersion Theory on a conceptual level, we do not believe it is possible to formally model our results in this framework until such time as a method is devised to quantify articulatory effort and perceptibility across different stress correlates, as suggested above.

---

[6]A reviewer offers a related explanation, which is that intensity may be to some extent parasitic on duration at the perceptual level. That is, a longer sound will sound louder than a shorter one at the same intensity. Thus Gordon (2002: 60, 2006: 191) introduces the concept of 'total perceptual energy', which is loudness measured as a function of both intensity and duration. It may thus be that increasing intensity or duration are two different strategies to achieve the same perceptual goal.

Another theoretical proposal we wish to consider in relation to our results is the Iambic–Trochaic Law (Bolton 1894; Kager 1993; Hayes 1995; Mellander 2003). The definition of the Iambic–Trochaic Law given by Hayes (1995) is reproduced in (24).

(24) *Iambic–Trochaic Law* (Hayes 1995: 80)
    a. Elements contrasting in intensity naturally form groupings with initial prominence.
    b. Elements contrasting in duration naturally form groupings with final prominence.

As stated by Hayes (1995), the Iambic–Trochaic Law refers to foot parsing: syllables may be grouped together into units of initial prominence (trochaic feet) or final prominence (iambic feet), depending on how the syllable nuclei differ from each other in intensity or in duration. However, the converse of this principle has also been assumed: that the grouping of syllables into groups with initial prominence or final prominence predicts how that prominence will be realised. Thus, Kager's (1993) formulation of the Iambic–Trochaic Law is given in (25).

(25) *Iambic–Trochaic Law* (Kager 1993: 382)
    a. Trochaic systems have durationally even feet.
    b. Iambic systems have durationally uneven feet.

The logic in (25) could similarly be applied to intensity: under the Iambic–Trochaic Law, we expect that trochaic systems would have feet with uneven intensity, whereas iambic systems would have feet with even intensity. This generalisation may ultimately be grounded in perceptibility: if an iambic system were to employ intensity as its primary stress correlate, then a sequence of (soft–loud)(soft–loud) syllables would be reinterpreted as soft(loud–soft)(loud), given what appears to be a universal tendency for syllables varying in intensity to be grouped head-initially (Crowhurst 2016, 2018, 2020; Crowhurst and Teodocio Olivares 2014). In other words, listeners would re-analyse the system as trochaic. Viewed in this light, the Iambic–Trochaic Law predicts a correlation between metrical rhythm type and the primary stress correlate used: trochaic systems are predicted to employ intensity as a primary stress correlate, whereas iambic systems are predicted to employ duration.

As we saw in §3.4, there is ample distributional evidence that Tetsǫ́t'ıné is an iambic language. Iambic feet condition vowel length adjustments (§3.4.1), movement of tone (§3.4.2) and deletion of consonants to avoid stress lapse (§3.4.3). As we also saw in §6, duration also appears to be the most robust correlate of stress in this language. From the perspective of the Iambic–Trochaic Law, these facts are expected: Tetsǫ́t'ıné employs duration as its main stress correlate rather than intensity because it is an iambic language. Thus, on the whole, our results seem to be more consistent with the Iambic–Trochaic Law than Dispersion Theory or the Functional Load Hypothesis, to the extent that these models make predictions about the phonetic realisation of stress cross-linguistically.

8. Conclusion

It has been observed that, cross-linguistically, stress is typically realised by some combination of increased F0 and/or increased duration (Gordon and Roettger 2017). It has also been claimed that languages avoid using phonetic properties which are independently contrastive as correlates of stress and also have a dispreference for employing intensity as a stress correlate (Berinstein 1979; Hayes 1995). This set of assumptions is potentially problematic for a language such as Tetsǫ́t'ıné, in which vowel length and tone are both contrastive. It has been claimed that in such a language, having a phonologically active stress system is impossible, and that stress can be 'covert' only (Spahr 2016). However, in this paper, we have provided both phonological distributional evidence as well as acoustic phonetic evidence for iambic stress in Tetsǫ́t'ıné. Regarding the manner in which stress is realised, we have found that duration rather than intensity is the primary correlate of stress in this language, which would follow from our understanding of the Iambic–Trochaic Law (Hayes 1995). This is in spite of the fact that vowel length is independently contrastive in Tetsǫ́t'ıné. However, as Lunden *et al*. (2017) have observed, increased vowel duration under stress does not necessarily obscure contrastive vowel length distinctions. Phonetic data from Tetsǫ́t'ıné itself would seem to support this view: in the present study, word-final stressed vowels were approximately 13% longer than unstressed vowels (Figures 3 and 4). In contrast, a study of contrastive vowel length by Jaker (2019) found that phonemically full (long) stem vowels were approximately 50% longer than reduced (short) vowels, and that this contrastive length difference was also accompanied by marked differences in vowel quality. In other words, phonetic lengthening under stress in Tetsǫ́t'ıné is non-neutralising, because phonetic lengthening is of a much lesser magnitude than phonemic length differences, and the latter are also enhanced by other acoustic cues.

If the use of duration as the primary stress cue in Tetsǫ́t'ıné is indeed a result of the Iambic–Trochaic Law, as we have suggested, a question for future research is whether what we have observed in Tetsǫ́t'ıné may hold true of iambic languages more generally. The Iambic–Trochaic Law predicts that, in iambic languages, duration will be the primary correlate of stress, whereas intensity should play only a minor role. Indeed, this is our impression. For example, in Menominee, an iambic language, the vowel in the head syllable of a foot is phonetically lengthened (Milligan 2006). In this context, it is noteworthy that in a recent survey of the literature on perceptual studies relating to the Iambic–Trochaic Law (Crowhurst 2020), all of the languages surveyed were either trochaic (e.g. English, German and Spanish) or else had no clear rhythm type (e.g. French and Japanese); there are no studies of which we are aware of rhythmic grouping biases in speakers of iambic languages. There have been similarly few production studies examining the realisation of stress in iambic languages, even though it is our impression that this seems to be the most widespread rhythm type in the indigenous languages of North America. Therefore, we believe our results underscore the need for additional phonetic studies on stress in iambic languages to verify whether these languages do, as a group, differ fundamentally from trochaic languages at the phonetic level.

References

Ackroyd, Lynda (1982). Dogrib gramar. Ms., University of Toronto.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**. 1–48.

Becker-Kristal, Roy (2010). *Acoustic typology of vowel inventories and Dispersion Theory: insights from a large cross-linguistic corpus*. PhD dissertation, University of California, Los Angeles.

Beckman, Mary (1986). *Stress and non-stress accent*. Dordrecht: Foris.

Berinstein, Ava (1979). *A cross-linguistic study on the perception and production of stress*. Master's thesis, University of California, Los Angeles.

Bloomfield, Leonard (1933). *Language*. Holt, Rinehart & Winston.

Boersma, Paul & David Weenink (2020). Praat: doing phonetics by computer. Version 6.1.34, https://praat.org/.

Bolton, Thaddeus (1894). Rhythm. *American Journal of Psychology* **6**. 145–238.

Cardinal, Emerence, Alessandro Jaker & Dora Cardinal (2021). *Tetsǫ́t'ıné dictionary / Tetsǫ́t'ıné Yatıé K'ı́zı́ T'asíe Hudzí ʔErehtł'ís*. Fairbanks, AK: ANLC Publications.

Cook, Eung-Do (1983). Chipewyan vowels. *IJAL* **49**. 413–427.

Cook, Eung-Do (2004). *A grammar of Dëne Sųłıné (Chipewyan)*. Number 17 in Algonquian and Iroquoian Linguistics Memoirs. Winnipeg: University of Manitoba.

Crowhurst, Megan (2016). Iambic-trochaic law effects among native speakers of Spanish and English. *Laboratory Phonology* **7**. 1–41.

Crowhurst, Megan (2018). The influence of varying vowel phonation and duration on rhythmic grouping biases among Spanish and English speakers. *JPh* **66**. 82–99.

Crowhurst, Megan (2020). The iambic/trochaic law: nature or nurture? *Language and Linguistics Compass* **14**. 1–16.

Crowhurst, Megan & Amador Teodocio Olivares (2014). Beyond the iambic-trochaic law: the joint influence of duration and intensity on the perception of rhythmic speech. *Phonology* **31**. 51–94.

DeLacy, Paul (2002). The interaction of tone and stress in Optimality Theory. *Phonology* **19**. 1–32.

DeLacy, Paul (2007). The interaction of tone, sonority, and prosodic structure. In Paul DeLacy (ed.) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press, 281–307.

Displayr (2021). flipMultivariates: multivariate models. R package, version 1.0.8.

Everett, Keren Madora (1998). The acoustic correlates of stress in Pirahã. *Journal of Amazonian Languages* **1**. 104–162.

Flemming, Edward (1995). *Auditory representations in phonology*. PhD dissertation, University of California, Los Angeles.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* **18**. 7–44.

Flemming, Edward & Hyesun Cho (2017). The phonetic specification of contour tones: evidence from the Mandarin rising tone. *Phonology* **34**. 1–40.

Gordon, Matthew (2002). A phonetically driven account of syllable weight. *Lg* **78**. 51–80.

Gordon, Matthew (2006). *Phonological typology*. Oxford: Oxford University Press.

Gordon, Matthew & Timo Roettger (2017). Acoustic correlates of word stress: a cross-linguistic survey. *Linguistics Vanguard* **3**.

Haas, Mary (1968). Notes on a Chipewyan dialect. *IJAL* **34**. 165–175.

Hall, Daniel Currie (2011). Phonological contrast and its phonetic enhancement: dispersedness without dispersion. *Phonology* **28**. 1–54.

Hargus, Sharon (2005a). Prosody in two Athabaskan languages of northern British Columbia. In Rice & Hargus (2005a), 393–423.

Hargus, Sharon (2005b). Stress in polysyllabic morphemes: Sekani and Deg Xinag. In Suzanne Gessner (ed.) *Proceedings of the 2005 Athabaskan Languages Conference*. Fairbanks, AK: ANLC Publications, 39–66.

Hargus, Sharon & Siri Tuttle (2004). Explaining variability in affix order: the Athabaskan areal and third person prefixes. *ANLC Working Papers* **4**. 70–98.

Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Hoijer, H., L. Bloomfield, M. R. Haas, A. M. Halpern, F. K. Li, S. S. Newman, M. Swadesh, G. L. Trager, C. F. Voegelin & B. L. Whorf (eds.) (1946). *Linguistic structures of Native America*. Number 6 in Viking Fund Publications in Anthropology. New York: Viking Fund.

Hoijer, Harry (1946). Chiricahua Apache. In Hoijer *et al.* (1946), 55–84.

House, Arthur (1961). On vowel duration in English. *JASA* **33**. 1174–1178.

Jaker, Alessandro (2019). The full ∼ reduced vowel contrast in Tetsǫ́t'ıné: evidence for an 8 vowel system. *ANLC Working Papers* **15**.

Jaker, Alessandro (2020). On the historical source of a ∼ u alternations in Dëne Sųłíné optative paradigms. *Glossa* **5**. 1–33.

Jaker, Alessandro (to appear). Tetsǫ́t'ıné prefix vowel length: evidence for systematic underspecification. *NLLT*.

Jaker, Alessandro & Emerence Cardinal (2020). *Tetsǫ́t'ıné verb grammar / Tetsǫ́t'ıné Yatıé k'ízı̨́ t'at'ú henádhër yatıé ʔełtth'ı ʔełaníílye*. Fairbanks, AK: ANLC Publications.

Jaker, Alessandro & Paul Kiparsky (2020). Conjugation tone mapping and level ordering in Tetsǫ́t'ıné. *Phonology* **37**. 617–655.

Kager, René (1993). Alternatives to the iambic-trochaic law. *NLLT* **11**. 381–432.

Krauss, Michael (1964). Proto-Athapaskan-Eyak and the problem of Na-Dene: phonology. *IJAL* **30**. 118–131.

Krauss, Michael (1983). Vowels in Saskatchewan Chipewyan: Mary Jane Kasyon. Ms., Alaska Native Language Center. Available at http://www.uaf.edu/anla/item.xml?id=CA961K1983.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* **82**. 1–26.

Leer, Jeff (2005). How stress shapes the stem–suffix complex in Athabaskan. In Rice & Hargus (2005a), 277–318.

Li, Fang-Kuei (1933). A list of Chipewyan stems. *IJAL* **7**. 122–151.

Li, Fang-Kuei (1946). Chipewyan. In Hoijer *et al.* (1946), 398–423.

Lovick, Olga (2020). *A grammar of Upper Tanana, volume I: phonology, lexical classes, morphology*. Lincoln, NE: University of Nebraska Press.

Lunden, Anya, Jessica Campbell, Mark Hutchens & Nick Kalivoda (2017). Vowel-length contrasts and phonetic cues to stress: an investigation of their relation. *Phonology* **34**. 565–580.

McDonough, Joyce & Valerie Wood (2008). The stop contrasts of the Athabaskan languages. *JPh* **36**. 427–449.

Mellander, Evan (2003). (HL)-creating processes in a theory of foot structure. *The Linguistic Review* **20**. 243–280.

Milligan, Marianne (2006). *Menominee prosodic structure*. PhD dissertation, University of Wisconsin, Madison.

Padgett, Jaye & Marija Tabain (2005). Adaptive Dispersion Theory and phonological vowel reduction in Russian. *Phonetica* **62**. 14–54.

Potisuk, Siripong, Jack Gandour & Mary Harper (1996). Acoustic correlates of stress in Thai. *Phonetica* **53**. 200–220.

Prince, Alan (1990). Quantitative consequences of rhythmic organization. *CLS* **26**. 355–398.

Prince, Alan (1991). Quantitative consequences of rhythmic organization. Ms., Brandeis University.

Prince, Alan & Paul Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar*. Malden, MA & Oxford: Blackwell.

R Core Team (2021). R: a language and environment for statistical computing. Published online at https://www.R-project.org/.

Rice, Keren (1989). *A grammar of Slave*. Berlin: Mouton de Gruyter.

Rice, Keren (1990). Prosodic constituency in Hare (Algonquian): evidence for the foot. *Lingua* **82**. 201–244.

Rice, Keren (2000). *Morpheme order and semantic scope: word formation in the Athapaskan verb*. Cambridge: Cambridge University Press.

Rice, Keren & Sharon Hargus (eds.) (2005a). *Athabaskan prosody*. Amsterdam: John Benjamins.

Rice, Keren & Sharon Hargus (2005b). Introduction. In Rice & Hargus (2005a), 1–45.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* **2**. 110–114.

Schwartz, Jen-Luc, Louis-Jean Boë, Nathalie Vallée & Christain Abry (1997). The dispersion-focalization theory of vowel systems. *JPh* **25**. 255–286.

Spahr, Christopher (2016). *Contrastive representations in non-segmental phonology*. PhD dissertation, University of Toronto.

Toivonen, Ida, Lev Blumenfeld, Andrea Gormley, Leah Hoiting, John Logan, Nalini Ramlakhan & Adam Stone (2015). Vowel height and duration. *WCCFL* **32**. 64–71.

Trudgill, Peter (2009). Greek dialect vowel systems, vowel dispersion theory, and sociolinguistic typology. *Journal of Greek Linguistics* **9**. 165–182.

Tuttle, Siri (1998). *Metrical and tonal structures in Tanana Athabaskan*. PhD dissertation, University of Washington.

Wickham, Hadley (2016). *ggplot2: elegant graphics for data analysis*. 2nd edition. Cham, Switzerland: Springer.

Wood, Simon N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society* **73**. 3–36.

Zec, Draga (1999). Footed tones and tonal feet: rhythmic constituency in a pitch-accent language. *Phonology* **16**. 225–264.