

Article

Pitch Range and Voice Quality in Dimasa Focus Intonation

Shakuntala Mahanta ^{1,*} , Amalesh Gope ²  and Priti Raychoudhury ^{1,*}

¹ Department of Humanities and Social Sciences, Indian Institute of Technology Guwahati, Guwahati 781039, India

² Department of English, Tezpur University, Tezpur 784028, India; amalesh@tezu.ac.in

* Correspondence: smahanta@iitg.ac.in (S.M.); pritionline89@gmail.com (P.R.)

Abstract: This paper presents an analysis of Dimasa focus intonation. The acoustic analysis shows that narrow focus sentences undergo a jump in the pitch range irrespective of the underlying tonal value of the morpheme it attaches to. In addition to f0 expansion, the prosodic property of focus in Dimasa was found to have different (tense) phonation in morphologically marked narrow focus sentences when compared to the broad focus context. Thus, the tense phonation property of sentences bearing morphological focus is not only an acoustic property of a higher pitch range but may also be an acoustic cue of discourse-level intonation.

Keywords: Dimasa; pitch range; voice quality; morphological focus; discourse intonation



Citation: Mahanta, Shakuntala, Amalesh Gope, and Priti Raychoudhury. 2021. Pitch Range and Voice Quality in Dimasa Focus Intonation. *Languages* 6: 185. <https://doi.org/10.3390/languages6040185>

Academic Editors: Christina Esposito and Sameer Khan

Received: 7 May 2021

Accepted: 21 October 2021

Published: 8 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In this paper, we analyze sentences bearing morphological focus in Dimasa. We show that the prosodic cues for focus marking in a language such as Dimasa rest heavily on pitch range modification. Focus marking is morpho-syntactically indicated using a morphological focus marker which triggers a rise in the pitch range of the sentence. The characteristic pitch trends of the lexical tones are preserved along with expansion of the corresponding pitch range. We show in this paper that pitch range expansion is accompanied by tense phonation, and therefore we claim that the morphological focus marker leads to a voice quality change in Dimasa in discourse-level intonation. This study therefore addresses the role of voice quality in focus intonation, an area which has received very little attention in the literature. Previous studies (Kuang and Keating 2014; Kuang and Liberman 2016; Kuang et al. 2016; Kuang and Liberman 2018) have shown that voice quality plays a role in pitch range perception. This study shows from the perspective of production that pitch range differences at the level of intonation are accompanied by additional phonation differences.

1.1. Background: Phonation and Focus

Ladefoged (1971) observed that voicing varies from a relatively open glottis (breathy) through a stable glottal position (modal), to a relatively closed glottis (creaky). Along this continuum, there are two intermediate positions commonly referred to as tense and lax. Tense and lax phonation basically indicate two sides of the continuum, where lax is positioned on the breathier half (more open), and tense is positioned on the creakier half (more closed) (Esposito and Khan 2020). Based on segmental studies, Gordon and Ladefoged (2001) distinguished a falling spectrum from a flat spectrum depending upon whether the glottal constriction was narrower as in creaky phonation, or less so, as in breathy phonation. Unlike breathy and creaky, the spectral properties of tense and lax phonation remained unclear until more recent studies. The difference between the first and second harmonics has been reported to distinguish the lax vs. tense phonation contrasts where lax phonation is realized with relatively higher values (Kuang 2011; Kuang and Keating 2014). Spectral tilt measurements successfully distinguished phonological phonation types in many languages

of the world: for example, Zapotec (Esposito 2004), Hmong varieties (Garellek et al. 2013; Keating et al. 2010), Gujarati (Khan 2012), Chong (DiCanio 2012), Yi languages (Kuang and Keating 2014), and Mazatec (Garellek and Keating 2011; Keating et al. 2010). In general, various spectral measurements associate higher and positive spectral values (in dB) with breathy vowels, an intermediate value with modal vowels, and less and often negative values with creaky vowels (Blankenship 2002; DiCanio 2009; Wayland and Jongman 2002; Khan 2012). Since tense and lax phonation (somehow) represent the continuum between the breathy and modal and creaky and modal ranges, it is expected that lax phonation will be associated with higher (and positive) spectral values when compared to its tense phonation counterparts. Previous studies have shown that when ‘tense voice’ (Kuang and Liberman 2016; Kuang and Liberman 2018) was used as stimuli, it was perceived as ‘higher’ in pitch. Listeners generally hear higher pitches when the spectrum includes more high-frequency energy (i.e., tenser phonation). The results presented from our experiments further support the hypothesis that voice quality cues are strong indicators of higher-pitch range even at the level of discourse-level intonation. Another important contribution of this paper is, evidently, the use of natural speech to understand the role of phonation in intonational focus. While Kuang and Liberman (2016) and Kuang and Liberman (2018) used synthetic stimuli in their experiments, our results are based on natural speech.

As already mentioned, this study attempts to bridge the gap between research on phonation and the prosodic property of focus. Studies on focus marking in tone languages have shown that discourse-level intonation and lexical tones can use different acoustic properties of f0 movement and f0 range to convey discourse information structure (Chen and Gussenhoven 2008; Xu 1997). Xu (1999) showed that in Mandarin, non-final focus leads to pitch range expansion in the words under focus. Xu et al. (2012) reported focused raising of f0, and an increase in intensity and duration. Pitch expansion of the focus constituent is also attested in Hausa (Leben et al. 1989; Hartmann 2008). Thus, there are multiple instances where pitch modification is employed by a tone language for discourse information structure. Pitch range modification (called register in their work) in Hausa and Jita has previously been discussed in Inkelas and Leben (1990) and in Downing (1996). Das (2016) showed similar pitch range modification in Boro, although no phonation studies were carried out. As the overview of the literature shows, studies on f0 and focus are abundantly available, but in this paper, we address the often overlooked property of phonation when there is pitch range expansion under focus. The study of Esposito (2010) is among the rare instances which looks at the interaction between focus and phonation. She showed that in Santa Ana del Valle Zapotec, the otherwise contrastive phonation in the language is minimized in focus positions. The production of focus is also accompanied by high F0, but the phonation properties are attenuated.

The approach preferred here is to locate the notion of ‘focus’ and ‘prominence’ independent of the different branches of linguistics, and to also understand the influence of phonetic (and, to a lesser extent, phonological) properties on focus structure (determined by syntax/semantics). We consider the sub-categories of ‘broad’ and ‘narrow’ focus in this paper¹. The specific instances of the narrow focus sentences that we consider here are utterances which signal new information by means of a morphological marker. Hence, we call them morphologically marked focus or morphological focus (abbreviated as MF) throughout this paper. The marker’s presence indicates narrow focus, and it is also identified as an MF marker or morphological focus marker.

1.2. The Dimasa Language and Its Tonal Properties

Dimasa belongs to the Boro Garo subgroup of the Tibeto Burman family of languages. There are 16 consonants: p, b, t, d, k, g, ʒ, s, h, r, m, n, ŋ, l, w, j, and six vowels: a, i, u, e, o, ə.² Most of the languages in this subgroup are tonal languages, except for Garo. Although Dimasa monosyllables have a few instances of glottal codas, we did not find a consistent correspondence between these glottal stops and the lexical tones. Phonation is also not contrastive in Dimasa. Dimasa attests two lexical tones, namely, high and low

tones (Jacquesson 2008). Sarmah (2009) found three tones—high, mid, and low. The mid tone varies from the high tone in that the high tone has a rise and forms a contour in its latter part, whereas the mid tone remains a level tone. The distinction between all three tones is essentially realized in the latter part of the syllable in monosyllabic words, and in the second syllable in disyllabic words (Sarmah 2009). The language assigns tone only to the prosodic word, and contrastive tone is attested in the second syllable. Some examples of Dimasa citation tones where three lexical tones can be found are presented below in Table 1a. In Table 1b, we show the words used for illustration of tones in Figures 1 and 2.

Table 1. (a) Dimasa tones; (b) Dimasa tones illustrated in the figures.

(a) Tone	Word	Gloss	Word	Gloss	Word	Gloss
High	tí	‘speak’	lái	‘page’	maitái	‘year’
Mid	tī	‘die’	laī	‘easy’	maitāi	‘crop’
Low	tì	‘blood’	laì	‘wish	maitài	‘source’
(b) Tone	Word	Gloss	Word	Gloss		
High	bái	‘announce’,	dáij	‘moon’		
Low	bài	‘dance’	dàij	‘chop’		

Figure 1 below plots the tonal contrast for the high–low pair /bái/ ‘announce’ and /bài/ ‘dance’ for the f₀ direction. Figure 2 plots the pitch tracks of the high–low pair /dáj/ ‘moon’ and /dàj/ ‘chop’. A three-way lexical tone contrast was not found for these words. The reason for using these words for illustration was the stable tonal contrast produced by all the speakers, as against the variable production of contrastive tone of the words in Table 1.

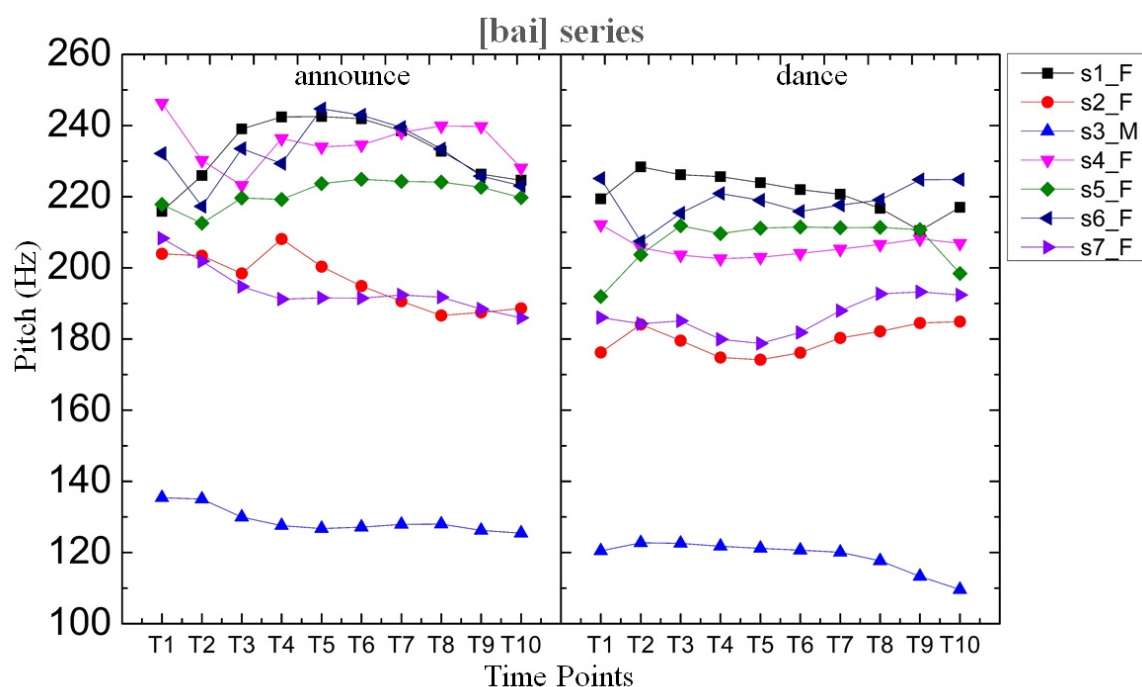


Figure 1. Time-normalized pitch tracks for [bai] series: /bái/ ‘announce’, and /bài/ ‘dance’. N = 28, 7 speakers x 2 iterations x 2 lexical items.

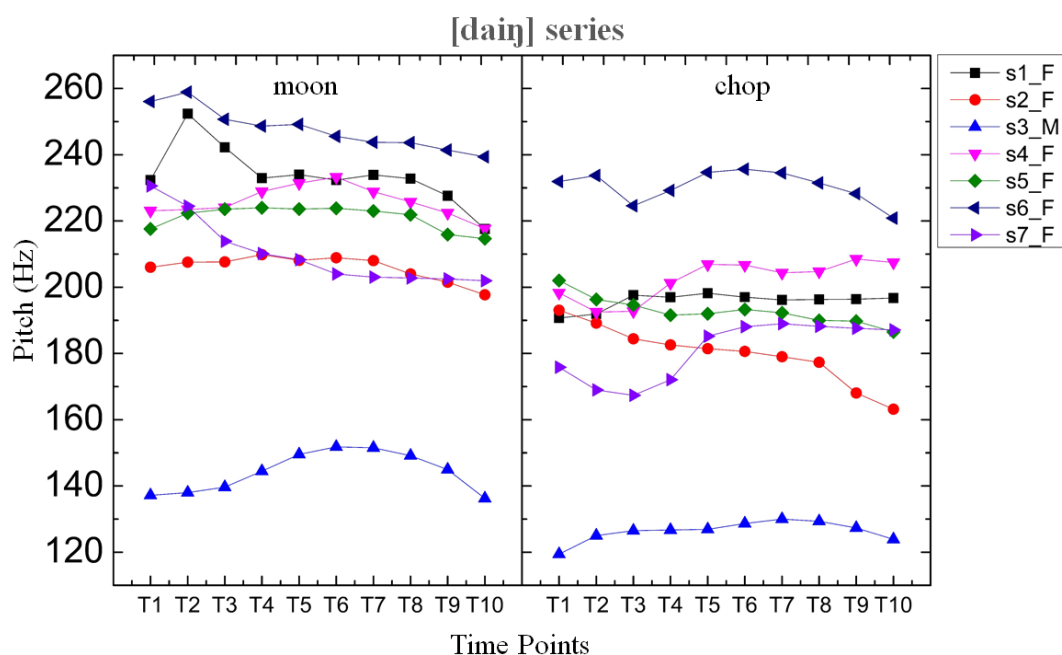


Figure 2. Time-normalized pitch tracks for [daiŋ] series: /dáiŋ/‘moon’, and /dàiŋ/‘chop’. N = 28, 7 speakers x 2 iterations x 2 lexical items.

We concur with [Jacquesson \(2008\)](#) that only two tones play a role in synchronic Dimasa phonology. The mid tone does not surface contrastively in most sentential patterns, and it seems to be variable, sometimes appearing as a level high tone and, at other times, indistinguishable from the high and low³ tones. In disyllables, Dimasa distinguishes tones based on the pitch variations in the second syllable. The lexical tone is received by the penultimate syllable of the same word when followed by a boundary tone⁴ in the final syllable. The high tone in Dimasa may surface with a high level pitch contour as in the word for ‘write’/réb/ and with a high rising tone as in the word for ‘blow’/jú/. These two variants of the lexical high tone have been argued to be allophonic in Dimasa by [Sarmah \(2009\)](#). He attributes this variance to the inherent property of voiceless consonants which raise the f₀ for the words with an underlying high tone. Sarmah shows that the voiceless onsets /ʃ/ (/s/ in the inventory we have shown) and /t/ embody their higher inherent frequency on the following vowels and notices a lower frequency in words with the liquid onset /r/.

Tonal alignment in Dimasa is always to the right of the prosodic word, as with its cognate language Boro. [Das and Mahanta \(2019\)](#) showed that, in Boro, the lexical tone shifts to a toneless suffix when the suffix is attached to a monosyllabic stem. However, suffixes with a prespecified underlying tone are not amenable to tonal change since the right edge of the prosodic word is already associated with a lexical tone. As shown in [Table 2](#) and the following figures in [Figure 3](#), the underlying tone of the stem tone aligns with the suffix.

Table 2. Tonal alignment in Dimasa-derived disyllables.

Word	Gloss	Perfective	Word	Gloss	Accusative	Word	Gloss	Reflexive
kái	‘run’	kai-bá	wái	‘fire’	wai-ké	tsì	‘say’	ti-là
kài	‘carry’	kai-bà	Bù ⁵	‘he/she’	bu-kè	réb	‘write’	reb-lá

This pattern is demonstrated by the time-normalized pitch contours in [Figure 3](#). The lexical tone of the word for fire, wái, is high which surfaces as a rise and then a fall.

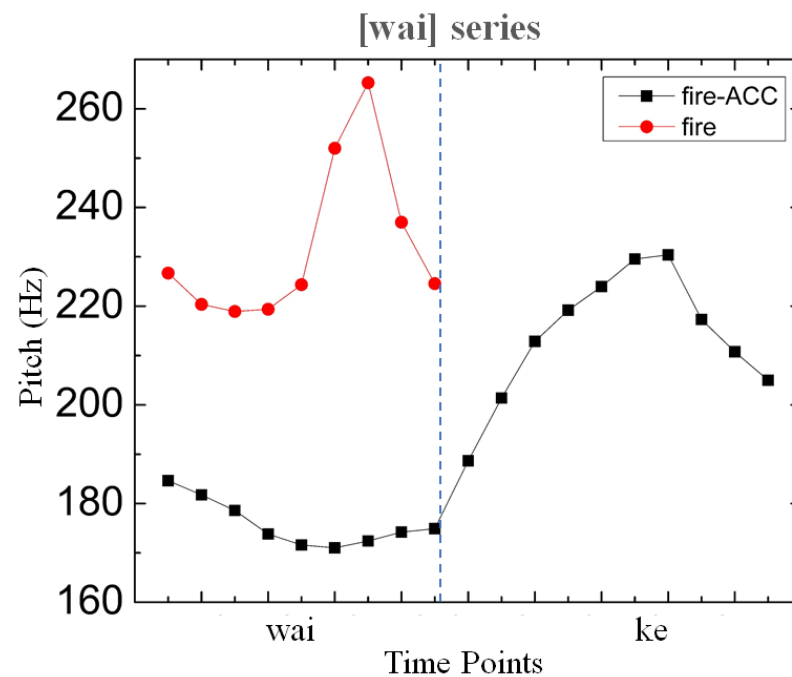


Figure 3. Averaged time-normalized pitch tracks of /wái/‘fire’ series with and without the ACC element /-ké/. N = 28, 7 speakers \times 2 iterations \times 2 lexical items.

In the derived disyllabic word /wai-ké/‘fire-ACC’, the high tone follows the right-alignment pattern as the underlying high tone of the stem /wái/‘fire’ shifts to the accusative marker /-ké/. Each morpheme is represented by ten separate time points in Figure 4. It is distinctly visible that the pitch rises significantly from the onset of the accusative suffix here.

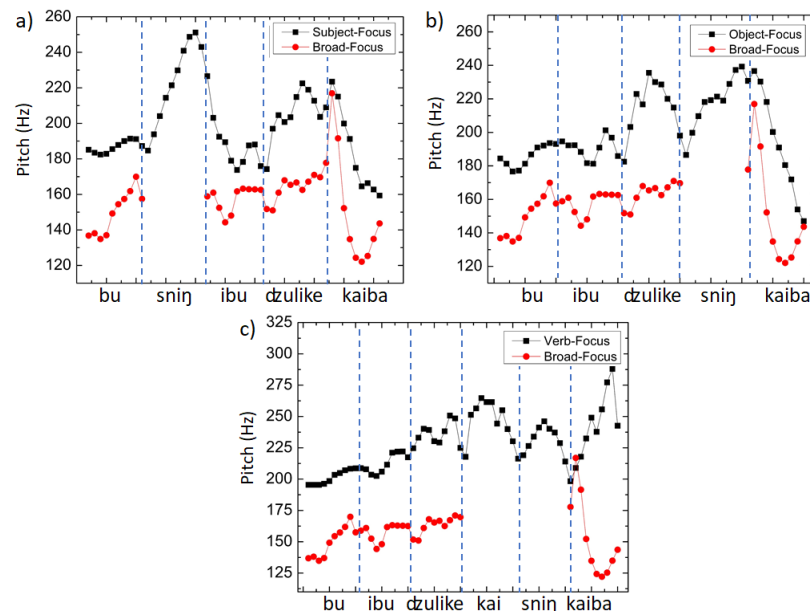


Figure 4. Morphological focus and broad focus: Time-normalized course of mean f0 for all tokens across all speakers. The diagram in panel (a) compares a broad focus sentence with the subject in MF sentences. The diagram in panel (b) compares the same broad focus sentence with the object in MF. The diagram in (c) compares the sentence in broad focus with a sentence with a verb in MF. The x-axis represents the average of the time-normalized f0 for 8 speakers in 2 iterations. Each word is plotted over 10 interval points. The y-axis represents pitch in Hz.

2. Recording Procedures, Materials, and Data Collection

This study was conducted on intonation patterns produced by Dimasa speakers residing in Haflong, in the northern half of Dima Hasao District, an administrative district of the Indian state of Assam. Nine native speakers of Dimasa (5 females, 4 males) of the age group 22–32 were recorded in a quiet room. Most speakers shared a fluency of English and Haflong-Hindi as well. For the experiment, native speakers of Dimasa were recorded in a quiet classroom of Haflong College and Dimasa Students' Union's office. Sentences were elicited in response to previously prepared questions, where the corresponding questions to elicit broad focus vs. narrow focus utterances were asked in Dimasa by one speaker to another. They practiced the task once before the recording started. Speech data were recorded with a Shure unidirectional head-worn microphone connected to a Tascam DR MKII 100 recorder (ensuring a constant microphone-to-mouth distance) via an xlr jack. Praat (Boersma and Weenik 2019) was used for segmentation and annotation of the data. For broad focus, declarative sentential forms in SOV order were elicited. In order to understand the influence of tones adjacent to the focus target, we tried to elicit as many tonal combinations as possible, i.e., HHH, HMH, LHL, LLL, and other possible combinations. For narrow focus, the morphological focus markers on subjects, objects, and verbs in three different forms of each sentence were adjoined. This experiment was part of a bigger experiment on Dimasa tones and intonation. Hence, various types of question and answer pairs (topicalized, contrastive, broad, narrow, and some other declarative utterances not related to the focus experiment) were mixed in the elicitation process. This ensured that there was enough variation in the elicitation process, precluding the need for separate sets of filler sentences. Dimasa has been reported to have three tones, and each word is assigned only one tone. In our work, we consider only two tones because of the variable behavior of the mid tone which may well be a neutral tone. Although previous acoustic work on Dimasa tones is available (Sarmah 2009), we also carried out an acoustic analysis on our own using a corpus of 20 monosyllabic lexical pairs.

The MF marker follows the argument while marking focus prominence on the subject or the object and precedes the tense or aspectual markers while the focus is on the verb. The suffix /-sníŋ/ is used to strongly mark prominence on the argument or the predicate that it attaches to and receives phonetic prominence. The Dimasa verb complex can include focus marking for a number of syntactic categories. The verb under focus is reduplicated in Dimasa. The first occurrence of the verb is uninflected and is followed by the MF marker /-sníŋ/, and it is not inflected for tense or number.

We considered MF sentences which were compared to broad focus sentences for this study. Our data for broad focus had 28 sentences in 3 iterations elicited from 8 speakers. The data for MF sentences consisted of 18 sentences from the same list which were elicited from 9 speakers with 3 iterations each. We segmented, labeled, and analyzed the data across all the speakers and iterations. The suffix /-sníŋ/ which follows the argument is the MF marker of the subject. See Appendix A for the sentences used in this paper.

This phenomenon is presented for visual inspection in Figure 4a, where the contrast between the sentences in broad focus and MF sentences is presented in their time-normalized f0 courses. The difference between the subject and the object in broad focus and MF sentences is visualized in the time-normalized course of f0 in Figure 4a,b. The difference between the verb in the broad focus and MF sentences is visualized in Figure 4c.

2.1. Acoustic Measurements: F0 and Duration

Both f0 and duration measurements were obtained from the nucleus of each syllable. For the normalized pitch value, each target word was measured for pitch at 10 consecutive points starting from the onset until the offset of the vowel/rime, i.e., start pitch to end pitch (0–100%), each point representing 10% of the total length of the pitch track. The average pitch value of the time-normalized values of each speaker's production of homophonous words was drawn in contours to see the pitch contrast.

Wherever required (especially for the tone experiment), the average pitch value of the time-normalized values of homophonous words was drawn in contours to see the pitch variation, as shown in Figure 1. The target words were measured by segmenting individual sound files into phonemes, and PRAAT text grid files were created for the acoustic measurement of each target word.

The duration of each vowel in the subject, object, and verb positions in both MF and broad focus positions was obtained by running the *prosody pro* script.

The time-normalized pitch range of the entire sentence was considered for measuring pitch contours of individual sentences. Pitch range differences between MF and broad focus sentences for each sentence type were compared. The normalized pitch range of MF sentences was compared with that of the corresponding broad focus sentences. The MF sentences with focus on the subject, object, and verb in all the sentences (expressed by the morphological focus marker /-sníŋ/) were compared with their corresponding sentences with no focus. Duration, minf0, maxf0, and meanf0 of the focus vowels/words from the sentences were further measured by running two different PRAAT scripts to extract all the values.

2.2. Phonation Analysis

To examine the effect of voice quality of the vowels in broad focus vis-a-vis MF (narrow focus) contexts, we used the same sentences which were used to explore the focus-related properties as discussed above (Section 2.1). The target vowels in the subject position of a given sentence include [a] as in /aŋ/‘I’, [i] as in /niŋ/‘you’, and [u] as in /bu/‘he’. Similarly, the vowels in the object position include [u], [e], and [i] as in [bumu] ‘name’, /bune/‘3P-ACC’, /ʒuli-ke/‘bag’, /ri/‘cloth’, and /sis-a-ke/‘dog’. We compared the acoustic properties of the vowels occurring in broad focus sentences with those of the vowels in MF contexts.

It must be noted that the quality of the vowel may also impact the spectral tilt measurements. For example, the low F1 of a vowel directly influences the amplitude of the first and second harmonics (Iseli et al. 2007). For a long time, there was a tendency to ignore using high vowels to determine the phonation quality in a given language. To address this issue, Shue et al. (2011) incorporated the advancements of algorithms to correct the formant frequency and bandwidths (Hanson 1995; Iseli et al. 2007) in VoiceSauce. VoiceSauce is a free program that computes a variety of acoustic measurements including f0, formant frequencies, and harmonic amplitudes with corrections for surrounding formant frequencies and bandwidths. In this study, we included both high and low vowels and generated the corrected spectral values in VoiceSauce for the purpose of analysis.

Furthermore, studies also reported that the phonation qualities might reflect differences in male and female speakers due to different physiological properties (Simpson 2012; Hanson and Chuang 1999). In our case, we observed consistent results across gender. Therefore, we did not separate the male and female speakers’ data in our analysis.

The acoustic components⁶ considered in this study include amplitude differences of various harmonics and formants, viz., H1*-H2*, H1*-A1*, H1*-A2*, and H1*-A3*. Among these four acoustic measurements, the difference between the amplitudes of the first and second harmonics (H1*-H2*) and the difference between the amplitudes of the first harmonic and third formant (H1*-A3*) have been reported to systematically distinguish different phonation types in many languages (Garellek and Keating 2011; Keating et al. 2010; Blankenship 2002; DiCanio 2009; Esposito 2010). In a cross-linguistic study, Keating et al. (2010) observed that only H1*-H2* is capable of distinguishing different phonation categories both across and within languages. Blankenship (2002), on the other hand, observed that H1*-A2* is better at distinguishing breathy voice from modal voice. The findings of these studies indicate that there is consistency in the way the phonation differences are reflected in different languages in terms of different acoustic components. In this study, we therefore decided to examine the above-mentioned acoustic measurements, viz., H1*-H2*, H1*-A1*, H1*-A2*, and H1*-A3*, to explore the phonation qualities (if any) of the vowels

bearing MF in different positions (viz., subject, object, and verb) of a sentence. The acoustic measurements were carried out automatically at every millisecond of the target vowel duration in VoiceSauce. The values were averaged across nine equal time points for each vowel to examine the acoustic changes (if any) over these time points.

3. Results

3.1. Pitch Range Results

An ANOVA test (conducted in IBM SPSS 20.0), where the duration of the vowel was the dependent variable and focus type (broad and MF) was the independent variable, was conducted. The corresponding vowels without focus for 220 subject, object, and verb tokens showed significant differences (subject [$(1, 38) = 13.06; p < 0.001$], object [$F(1, 38) = 12.01; p < 0.001$], and verb [$F(1, 38) = 10.3; p < 0.001$]). Duration was not found to be significantly different between MF and broad focus sentences.

Additionally, a visible difference in the pitch contour was found between MF and broad focus sentences. Irrespective of the difference in tonal sequences in the sentences, the pitch contour for the MF sentences always has an overall higher pitch when compared to that of the corresponding broad focus contour. The pitch contour variation between the broad focus and MF sentences shows up as a cue for identification of sentence patterns in Dimasa. The following figures provide an idea of the pitch range variations found among broad focus and MF sentences. The pitch tracks show that the contours for each type vary with respect to broad focus and narrow focus. Prosodic cues for focus marking in a tone language such as Dimasa rest heavily on pitch range modification in a way which allows lexical tones to preserve their characteristic f_0 trends. No difference among the focus and post-focus parts could be found. The results show the pitch range variation for broad and MF sentences for three representative sentences with three different tonal sequences, namely, HHH, LLHL, and HLLL. The diagrams below in Figure 4 demonstrate that the pitch range for MF is higher than that for the broad focus sentences.

Examination of the pitch contours of the sentences containing the morphological focus marker /-snij/ 'only' also revealed that they bear another local effect. The pitch contours of such sentences reveal that the MF markers occurring in non-final positions surface with a higher F_0 peak than the preceding tone. There is noticeably a downslope of the post-subject and post-object constituents, although the pitch range variations between MF and broad focus is still maintained. It can be seen in the pitch tracks in Figure 4 that it is not only the time-normalized F_0 contour for the MF marker /snij/ 'only' which rises to a distinct higher target but the entire pitch contour also undergoes pitch range expansion. A comparison between the vowels bearing MF and the corresponding vowels without focus for the subject, object, and verb showed significant differences (subject [$(1, 38) = 13.06; p < 0.001$], object [$F(1, 38) = 12.01; p < 0.001$], and verb [$F(1, 38) = 10.3; p < 0.001$]). The duration was not found to be significantly different between MF and broad focus sentences.

3.2. Voice Quality Results

Voice quality variations in the vowels occurring in the subject position: All the acoustic components⁷, viz., $H1^*-H2^*$, $H1^*-A1^*$, $H1^*-A2^*$, and $H1^*-A3^*$, showed a reduction in spectral values in the vowels occurring in the focus conditions. This was observed across all the vowels (viz., [a, u, i]) considered in this study. A one-way ANOVA was conducted to compare the spectral values of the vowels occurring in the subject focus condition with their counterparts in the broad focus conditions. In this test, focus conditions (viz., MF vs. broad focus) were the dependent variables, and the various acoustic components (viz., $H1^*-H2^*$, $H1^*-A1^*$, $H1^*-A2^*$, and $H1^*-A3^*$) were deemed to be the independent variables. A comparison of the mean values of the vowels occurring in the MF contexts indicated a significant interaction between the acoustic correlates and the focus types. Interestingly, all four acoustic correlates, viz., $H1^*-H2^*$ [$F(1, 80) = 33.52; p < 0.001$], $H1^*-A1^*$ [$F(1, 80) = 31.52; p < 0.0001$], $H1^*-A2^*$ [$F(1, 80) = 41.32; p < 0.0001$], and $H1^*-A3^*$ [$F(1, 80) = 19.55; p < 0.0001$], confirm that the (reduced) spectral values associated with the low vowel [a] (/aɪ/ 'I') in

subject focus conditions are significantly low when compared to the vowels in broad focus conditions. However, only two of those four acoustic components, viz., $H1^*-H2^*$ ($[F(1, 60) = 37.32; p < 0.001]$ and $[F(1, 60) = 17.52; p < 0.001]$), and $H1^*-A3^*$ ($[F(1, 60) = 30.52; p < 0.001]$ and $[F(1, 60) = 26.92; p < 0.0001]$), indicate a statistically significant reduction in terms of spectral values in the two high vowels [u and i] when they occur in the subject focus conditions, respectively, even though a consistent reduction in the spectrum is also observed for the other two acoustic components, viz., $H1^*-A1^*$ and $H1^*-A2^*$, for both these vowels. The results are shown below in Figure 5.

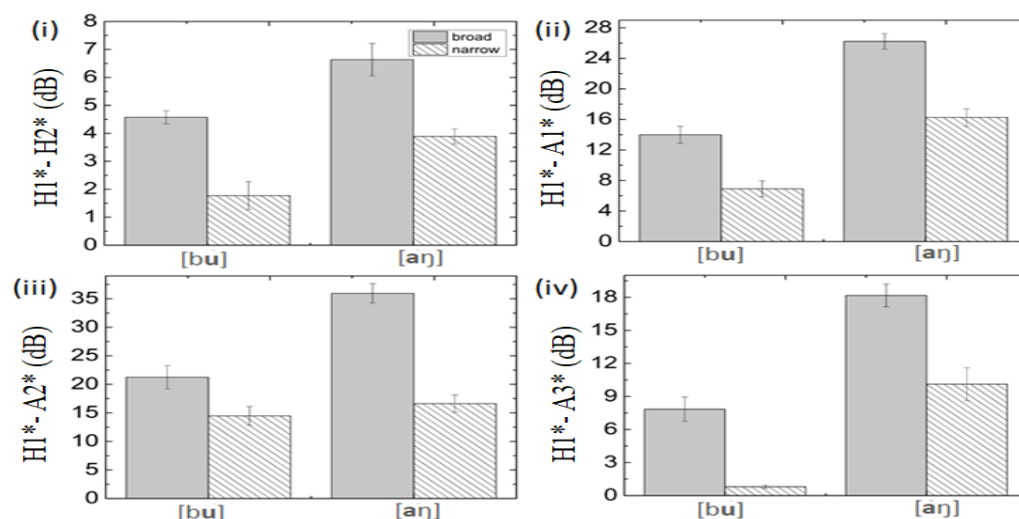


Figure 5. Mean spectral values calculated for all the acoustic components over total vowel duration for the vowels [u and a] occurring in the subject position of the words /bu/‘he’ and /aŋ/‘I.’ The solid line represents the vowels in broad focus contexts, and the striped line indicates the vowels in the MF contexts. The vowel in the MF condition is realized with low spectral values compared to the vowel in the broad focus context.

The overview of the distribution of spectral differences across nine time points for all the vowels occurring in different focus contexts indicates a steady low for the vowels occurring in the MF contexts. While an overall larger spectral difference could be observed for the vowel [a] across all acoustic components and all the time points, for the other two vowels, [u] and [i], such a difference is observed to be greater in the first six time points. This is shown in Figure 6.

When we survey the voice quality variations in the vowels occurring in the object position, we find that the object vowels representing the (object) focus contexts show a similar spectrum reduction compared to their counterparts occurring in the broad focus context. This trend is similar to the one we observed for the vowels occurring in the subject focus position. The comparison of the mean spectral values for all the acoustic components, viz., $H1^*-H2^*$ ($[F(1, 54) = 32.32; p < 0.0001]$ and $[F(1, 54) = 15.22; p < 0.001]$), $H1^*-A1^*$ ($[F(1, 54) = 38.24; p < 0.0001]$ and $[F(1, 54) = 28.22; p < 0.001]$), $H1^*-A2^*$ ($[F(1, 54) = 53.32; p < 0.0001]$ and $[F(1, 54) = 35.52; p < 0.001]$), and $H1^*-A3^*$ ($[F(1, 54) = 26.42; p < 0.0001]$ and $[F(1, 54) = 19.51; p < 0.001]$), confirmed a significant spectrum reduction for both the vowels occurring in the object focus context compared to their counterparts in broad focus conditions. The mean spectral values of /u/ and /e/ in the subject position are presented in Figure 7.

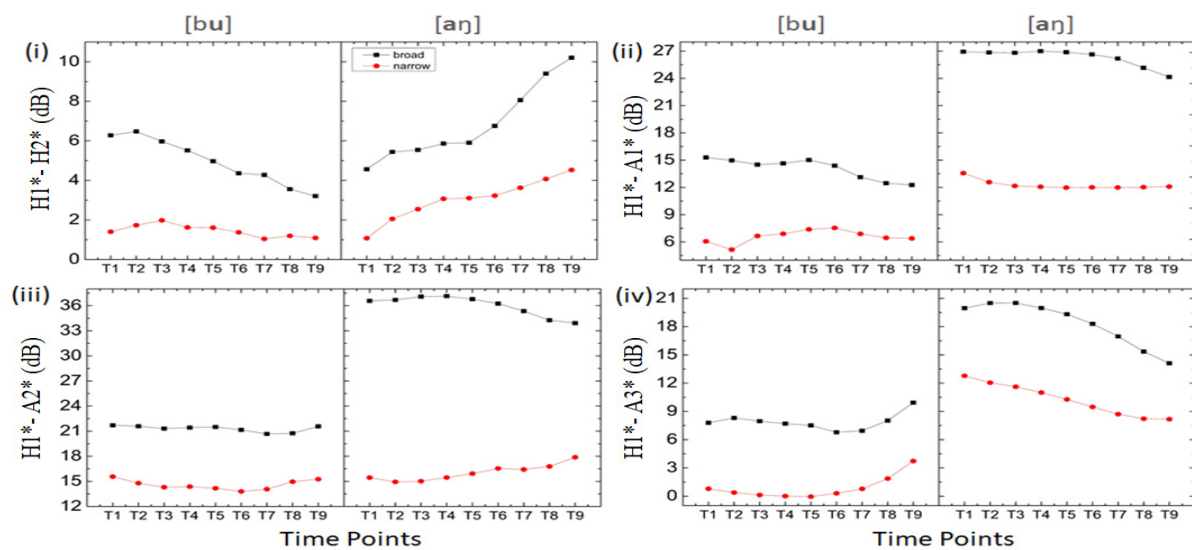


Figure 6. Time point spectral values calculated for all the acoustic components over total vowel duration for the vowels [u and a] occurring in the subject position of the words /bu/‘he’ and /aŋ/‘I.’ The blue line represents the vowels in broad focus contexts, and the red line indicates the vowels in the MF contexts. The vowel in the MF condition is realized with low spectral values compared to the vowel in the broad focus context.

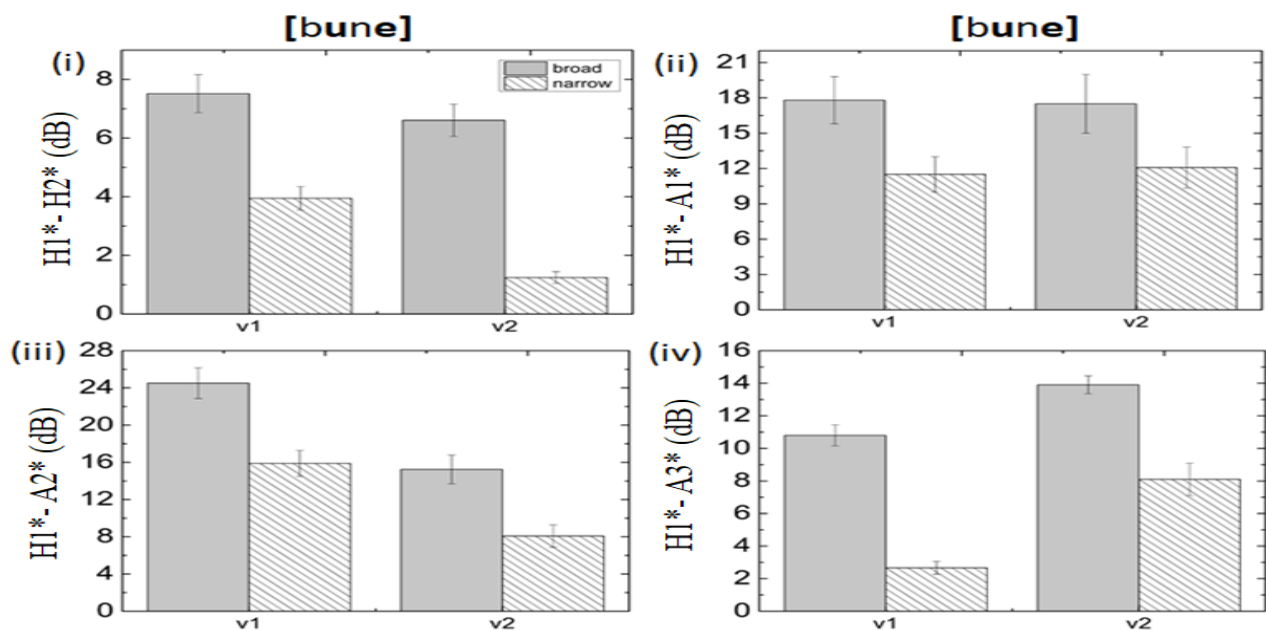


Figure 7. Mean spectral values calculated for all the acoustic components over total vowel duration of [u and e] occurring in the subject position of the word /bune/‘3P-ACC’. The solid line represents the vowels in broad focus contexts, and the striped line indicates the vowels in the MF focus contexts. v1 is the vowel [u] of the first syllable, and v2 [e] is the vowel of the second syllable. The vowel in the MF condition is realized with low spectral values compared to the vowel in the broad focus context.

As expected, a spectrum reduction was observed across all nine time points in both vowels occurring in the MF context. This is shown in Figure 8.

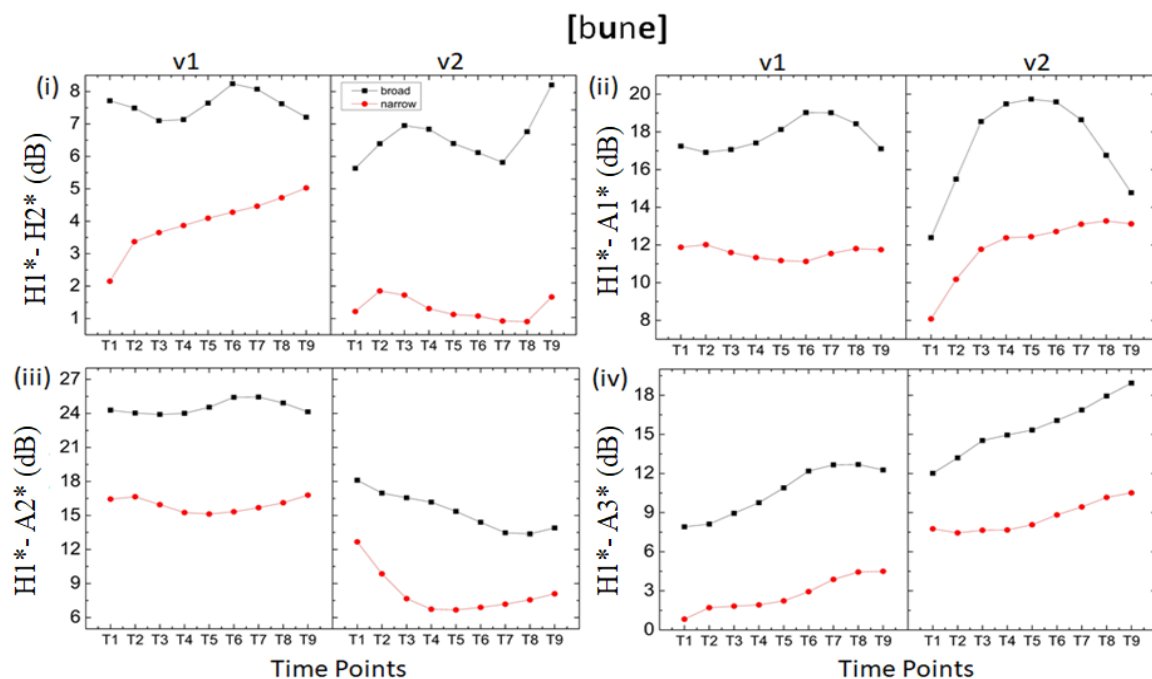


Figure 8. Time point spectral values calculated for all the acoustic components over total vowel duration for the vowels [u and e] occurring in the object position of the word /bune/ ‘3P-ACC’. The blue line represents the vowels in broad focus contexts, and the red line indicates the vowels in the MF contexts. v1 is the vowel [u] of the first syllable, and v2 [e] is the vowel of the second syllable. The vowel in the MF condition is realized with low spectral values compared to the vowel in the broad focus context.

4. Discussion

The spectral reduction in the vowels occurring in the morphologically marked narrow focus context confirms that native Dimasa speakers make use of voice quality properties to distinguish MF contexts from broad focus conditions (along with pitch range differences). Recall the discussion in Section 1.1, where the distribution of tense and lax in the breathiness–modal–creaky continuum was discussed. In this context, [Gordon and Ladefoged \(2001\)](#) observed that the involvement of increased constriction in the laryngeal area may lead to a resemblance of creaky phonation; however, tense phonation does not precisely fall under the rubric of creaky phonation. Tense phonation, thus, falls under the creaky–modal range. Lax phonation, on the other hand, is expected to carry a similar glottal configuration to breathy phonation. However, the glottal configuration of lax phonation is likely to be less extreme, since it is only *slightly breathy* ([Ladefoged and Maddieson 1996](#)). Lax phonation, therefore, falls under the breathy–modal range. Generally, the spectral balance and tilt measures are reported to be the highest for breathy phonation, intermediate for modal, and the lowest (and often negative) for creaky phonation. Concomitantly, relatively higher and positive spectral values are expected for both lax and tense phonation types. In such cases, spectral values are expected to be higher for lax phonation compared to tense phonation types. In this context, it must be noted that a higher f_0 is often realized in tense voice due to coarticulatory effects. In Dimasa, we observed an increased f_0 (Figure 9) in the vowels associated with MF conditions.

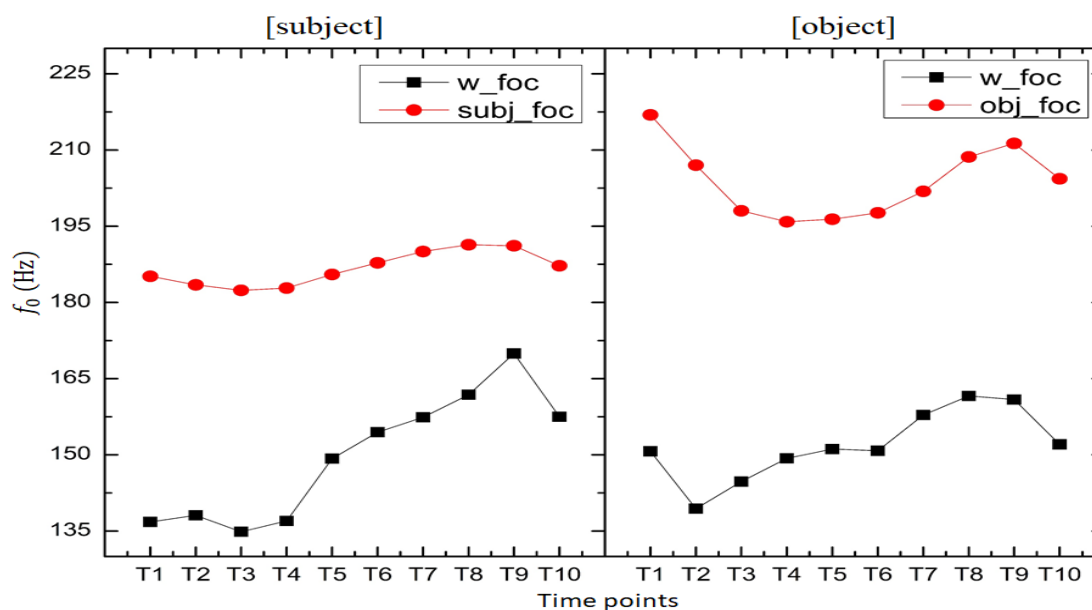


Figure 9. Time-normalized f_0 tracks of vowels representing MF and broad focus conditions. The left panel represents the vowels occurring in subject positions, while the right panel represents the vowels occurring in the object positions. The blue line represents the vowels in broad focus contexts, and the red line indicates the vowels in the MF contexts.

Naturally, an important question which arises at this point is whether the spectral differences observed through different acoustic measurements (reduced spectral values for the vowels associated with MF contexts) are a by-product of a higher f_0 borne out due to MF conditions. In this context, it is relevant to note that several Tibeto-Burman languages are known to have phonation-based register contrast (tense vs. lax) that is historically derived (Kuang 2011). Furthermore, Dimasa does not have breathiness contrasts in its consonant inventory. The language, at present⁸, attests a two-way tonal contrast in utterances (viz., high vs. low). While it is not possible to justify, at this point, that the phonation differences play an absolute role (and are the sole differentiating element) in identifying the differences in Dimasa intonation, we can argue that the phonation differences reported in this paper may facilitate the perception of focus-related information in this language (along with the f_0 differences). The significant reduction in spectral values associated with the vowels in MF contexts appears to be in the continuum of modal to creaky (hence, tense phonation), for all the vowels representing the MF condition occurring in different positions of a sentence, viz., subject, object, and verb. On the other hand, the vowels representing the broad focus context are associated with higher spectral values and thus fall under the modal-to-breathy continuum (lax phonation). Although perception studies could not be carried out⁹, we predict that tense phonation would be a good cue to perceive the MF contexts in Dimasa.

5. Conclusions

These results, that is, the way that tense phonation is associated with pitch range in the MF condition, bear important implications for any theory of focus prosody. Dimasa presents an instance where pitch raising and voice quality changes are recruited for prosodic purposes. This generates great interest for the typology of intonation since it employs neither phonological rephrasing (Ladd 1996; Downing 2006, 2007, 2008) nor sentence stress (Szendrői 2003; Truckenbrodt 1995; Samek-Lodovici 2005) for its focus prosody. In sum, the results for focus marking in Dimasa show how prosodic information and voice quality can play a vital role in information structure in addition to the support it provides to lexical tones. They also show how the acoustic aspects of a tone language can employ different aspects of acoustic possibilities for tone and intonation. Further, they also show the role of voice quality in the organization of information structure in tone languages.

Previous studies of tone languages have shown that f_0 is not the sole property of tonally contrastive lexical items. There are languages where voice quality plays a major role in addition to f_0 or even instead of f_0 , indicating tonal or register contrast (Bradley 1982; Brunelle 2005; Brunelle 2009; DiCanio 2009; Esposito 2012; Kuang 2013; Laver 1980; Ladefoged 1971). In this paper, we report findings from Dimasa, a language where there is no contrastive phonation, but we show that at the discourse structure level, higher pitch correlates with phonation differences only in specific focus types. Therefore, in addition to f_0 , the prosodic property of morphologically marked narrow focus sentences attests a different phonation (in the range of modal to tense) than that from the broad focus context. Thus, unlike other tone languages where voice quality plays a role at the segmental level, in Dimasa, the phonation property of modal to tense along with the higher pitch range of the sentence bearing morphological focus can be considered to be the acoustic cues, apart from the morphological information of the presence of the morphological focus marker. Finally, this paper also indicates that voice quality may serve as a perceptual cue for prosodic information at the discourse structure level which may be explored in further research on Tibeto-Burman languages.

Author Contributions: Conceptualization, S.M.; methodology, S.M. and A.G.; software, S.M. and A.G.; validation, S.M.; formal analysis, S.M.; investigation, S.M. and P.R.; resources, S.M. and P.R.; data curation, P.R.; writing—original draft preparation, S.M.; writing—review and editing, S.M. (whole) A.G. and P.R. (parts); visualization, S.M.; supervision, S.M.; project administration, S.M.; funding acquisition, S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by DeITy (Department of Information Technology), awarded to Shakuntala Mahanta IITG/SM/P/03.

Institutional Review Board Statement: Ethical review and approval were waived for this study, as it was not related to bio-medical and health research.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data is available with the corresponding author and can be made available for non-profit, academic and research purposes.

Acknowledgments: We would like to express our sincere thanks to the Dimasa Students' Union in Haflong and Guwahati for helping us to conduct this study, and gratitude to Shri Amit Daulagupu and Vimi Daulagupu for their kindness and hospitality during our stay in Haflong.

Conflicts of Interest: The authors declare no conflict of interest.

List of Abbreviations

LOC	Locative marker
ACC	Accusative marker
GEN	Genitive marker
INST	Instrumental marker
PRES	Present tense marker
PST	Past tense marker
FUT	Future tense marker
PERF	Perfective marker
GNO	Gnomic aspect
REFLEX	Reflexive marker
ASRT	Assertive
SING	Singular
PI	Plural
COP	Copulative
VN	Verbal noun

Appendix A

In the sentences from 1 to 18, /-sníŋ/ was added to the arguments to elicit narrow focus. The rest are for broad focus alone.

1. nĩŋ nĩnĩ bumú tsi-lá pu-duì LLHLL
2PSing your name say-REFLX can-PRES
you can tell (me) your name
2. nĩŋ nĩnĩ bumú reb-lá pu-duì LLHHL
2PSing your name write-REFLX can-PRES
you can write down your name
3. áŋ wai-ké ʃu-ku-má naŋ-dú HHLH
1PSing fire-ACC blow-ASP-FUT need-PRES
I need to blow the fire off
4. sisà aŋ-ké wai-bà LHL
dog 1PSing-DAT bite-PERF
the dog bit me
5. ʒĩŋ bu-ké do-ká LLH
1P-PL 3P-ACC hit-PST
we hit him
6. áŋ ri-ké brai-bá HHH
1P SING cloth-ACC buy-PERF
I bought a cloth
7. áŋ bu-nè kĩm ri-ká HLLL
1P SING 3P SING-ACC flower buy-PST
I gave the flower to him
8. bú sisa-ké ʒig-bá LLH
3P SING dog-ACC kick-PERF
he kicked the dog
9. bú ibù ʒuli-ké kai-bà LLHL
3P SING this bag-ACC carry-PERF
he carried this bag
10. bú aŋ-ké bai-bà nu-ká LHLL
3PSing 1P-DAT dance-PST see-PST
he saw me dance
11. ʒĩŋ buʃu-há bai-duì LHL
1P Pl festival-LOC dance-PRES
we dance at Bushu
12. ansaraò jao-bí LH
kids make noise-Assrtv
the kids are making noise
13. ʒĩŋ bu-ké do-ká LLH
1P Pl 3P-ACC beat-PST
we beat him up

14. áŋ rí brai-bá HHH
1P cloth buy-PERF
I bought a cloth
15. ané dī naŋ-duì HLL
1P Sing-ACC water need-PRES
I want water
16. áŋ oraha-ni-sáng kai-lang-bá HHH
1PSING there-GEN-from run-GNO-Pst
I ran away from there
17. bù tsi-ká LH
3PSing die-PST
he died
18. bù sainlai-bà hemsínín di-bí LLHH
3P SING talk-PST very sweet-Assertv
he can talk very sweetly
19. áŋ duhá ʒi-má HLH
1P SING now eat-FUT
I will eat now
20. nìŋ ibù mogon-ké lai-ʒán reb pu-duì LLHHLL
2Psing this meat-ACC banana leaf-INST wrap can-PRES
you wrap this meat in banana leaf
21. áŋ aní jao-ké ʒu-má HHHL
1P SING 1P-GEN hand-ACC-FUT
I will wash my hand
22. áŋ makám ʒi-má naŋ-dú HLHH
1P SING rice eat-FUT need-PRES
I need to eat rice
23. bunǝĩ naolai-hà graoʒámá bai-bá HLHH
3P-Pl village-LOC words-declaration announce-PST
they made an announcement in the village
24. nìŋ ibù mogon-ké basà basà klaihi dáŋ LLHLLLH
2P SING this meat-ACC small small do-VN chop
chop this meat into little pieces
25. ibù kím bufù rao-bí LLLH
this flower thorn has Assrtv
this flower has thorns
26. áŋ aní jao-ké ʒu-má HHHL
1P 1P-GEN hand-ACC wash-FUT
I will wash my hand
27. ʒìŋ dò ɖa-duì LLL
1P Pl six COP-PRES
Only we are six in number

28. ané	di	nan-du	HLL
1P-ACC	water	need-PRES	
I want water			

Notes

- ¹ Broad focus and narrow focus have been understood and elicited for the purpose of this paper as responses to the two types of questions below:(1a) What happened?(1b) What did he eat?The two questions are understood to give rise to two different focus conditions (where capitalization indicates that the word receives a nuclear accent):(2a) [F He ate a PIZZA]—Broad Focus(2b) He ate [F a PIZZA]—Narrow FocusIn 2b, the new information receives the focus, whereas in 2a, the entire sentence is under focus, and hence broad focus.
- ² Both Jacquesson (2008) and Sarmah (2009) reported 16 consonants. Jacquesson (2008) reported five vowels, while Sarmah (2009) reported six vowels. Sarmah reported the sixth vowel (the so-called sixth vowel in Boro Garo languages, Joseph and Burling 2006) as ə. We found it to be akin to the schwa.
- ³ This distribution may very well be the result of it being the neutral tone, and more investigation is required to decide on the status of the mid tone.
- ⁴ The boundary tone in declarative utterances is an L% or an HL% in Dimasa, resulting in boundary lowering.
- ⁵ The third-person singular is not marked for gender in Dimasa. It is transcribed as /bo/ in Longmailai (2012).
- ⁶ We examined a few other measurements as well, viz., energy, CPP, and H2*–H4*. However, none of these measurements display any consistency. Hence, we do not report these measurements in this paper.
- ⁷ One of the reviewers suggested the use of normalized spectral values. It must be noted that spectral values are automatically generated in VoiceSauce (Shue et al. 2011), and the corrected values have been used in this paper for further analysis.
- ⁸ We recorded and examined the tonal properties in Dimasa as part of the current project and found a two-way tonal contrast in the variety that we examined.
- ⁹ The current pandemic situation did not allow us to reach out to an adequate number of respondents for responses.

References

- Blankenship, Barbara. 2002. The timing of nonmodal phonation in vowels. *Journal of Phonetics* 30: 163–91. [CrossRef]
- Boersma, Paul, and David Weenik. 2019. Praat: Doing Phonetics by Computer (Version 5.3.04_win32) [Computer Programme]. Available online: <http://www.praat.org/> (accessed on 12 January 2012).
- Bradley, David. 1982. Register in Burmese. In *Tonation*. Edited by Bradley David. Canberra: Pacific Linguistics, the Australian National University, pp. 117–32.
- Brunelle, Marc. 2005. Register and tone in Eastern Cham: Evidence from a word game. *Mon-Khmer Studies* 35: 121–31.
- Brunelle, Marc. 2009. Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics* 37: 79–96. [CrossRef]
- Chen, Yiya, and Carlos Gussenhoven. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36: 724–46. [CrossRef]
- Das, Kalyan. 2016. Tone and Intonation in Boro. Unpublished Ph.D. dissertation, IIT Guwahati, Guwahati, India.
- Das, Kalyan, and Shakuntala Mahanta. 2019. Intonational phonology of Boro. *Glossa: A Journal of General Linguistics* 4: 103. [CrossRef]
- DiCanio, Christian. 2009. The Phonetics of Register in Takhian Thong Chong. *Journal of the International Phonetic Association* 39: 162–88. [CrossRef]
- DiCanio, Christian. 2012. Coarticulation between tone and glottal consonants in Itunyoso Trique. *Journal of Phonetics* 40: 1–15. [CrossRef]
- Downing, Laura J. 1996. *The Tonal Phonology of Jita*. Munich: LINCOM Europa.
- Downing, Laura J. 2006. The prosody and syntax of focus in Chitumbuka. *ZAS Papers in Linguistics* 43: 55–79. [CrossRef]
- Downing, Laura J. 2007. Focus prosody divorced from stress and intonation in Chichewa, Chitumbuka and Durban Zulu. Paper presented at the ICPHS 2007 Satellite Meeting, Intonational Phonology: Understudied or Fieldwork Languages, Saarbrücken, Germany, August 5.
- Downing, Laura J. 2008. Focus and prominence in Chichewa, Chitumbuka and Durban Zulu. *ZAS Papers in Linguistics* 49: 47–65. [CrossRef]
- Esposito, Cristina M. 2004. Santa Ana del Valle Zapotec phonation. *UCLA Working Paper on Phonetics* 103: 71–105.
- Esposito, Christina M. 2010. Variation in contrastive phonation in Santa Ana del Valle Zapotec. *Journal of the International Phonetic Association* 40: 181–98. [CrossRef]

- Esposito, Cristina M. 2012. An acoustic and electroglottographic study of White Hmong tone and phonation. *Journal of Phonetics* 40: 466–76. [CrossRef]
- Esposito, Christina, and Sameer ud Dowla Khan. 2020. The cross-linguistic patterns of phonation types. *Lang Linguist Compass* 14: e12392. [CrossRef]
- Garellek, Marc, and Patricia Keating. 2011. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *Journal of the International Phonetic Association* 41: 185–205. [CrossRef]
- Garellek, Marc, Patricia Keating, Christina M. Esposito, and Jody Kreiman. 2013. Voice quality and tone identification in White Hmong. *The Journal of the Acoustical Society of America* 133: 1078–89. [CrossRef] [PubMed]
- Gordon, Matthew, and Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29: 383–406. [CrossRef]
- Hanson, Helen. 1995. Glottal Characteristics of Female Speakers. Ph.D. dissertation, Harvard University, Cambridge, MA, USA.
- Hanson, Helen, and Erika Chuang. 1999. Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America* 106: 1064–77. [CrossRef]
- Hartmann, Katharina. 2008. Focus and Emphasis in Tone and Intonational Languages. Edited by Anita Steube. In *The Discourse Potential of Underspecified Structures*. Berlin and New York: De Gruyter, pp. 389–412. [CrossRef]
- Inkelas, Sharon, and William Leben. 1990. Where phonology and phonetics intersect: The case of Hausa intonation. In *Papers in Laboratory Phonology: Between the Grammar and Physics of Speech*. Edited by Beckman Mary and John Kingston. Cambridge: Cambridge University Press, pp. 17–34.
- Iseli, Markus, Yen-Liang Shue, and Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America* 121: 2283–95. [CrossRef]
- Jacquesson, Francois. 2008. A Dimasa Grammar. Available online: http://brahmaputra.ceh.vjf.cnrs.fr/bdd/IMG/pdf/Dimasa_Grammar-2.pdf (accessed on 5 May 2021).
- Joseph, Umbavu Varghese, and Robbins Burling. 2006. *The Comparative Phonology of the Boro Garo Languages*. Manasagangotri: Central Institute of Indian Languages.
- Keating, Patricia, Christina Esposito, Mark Garellek, Sameer ud Dowla Khan, and Jiangjing Kuang. 2010. Phonation contrasts across languages. *UCLA Working Papers in Phonetics* 108: 188–202.
- Khan, Sameer ud Dowla. 2012. The Phonetics of Contrastive Phonation in Gujrati. *Journal of Phonetics* 40: 780–95. [CrossRef]
- Kuang, Jianjing. 2011. Phonation contrast in two register contrast languages and its influence on vowel quality and tone. Paper presented at the 17th International Congress of Phonetics Sciences, Hong Kong, China, August 17–21; pp. 1146–49.
- Kuang, Jianjing. 2013. The tonal space of contrastive five level tones. *Phonetica* 70: 1–23. [CrossRef] [PubMed]
- Kuang, Jianjing, and Patricia Keating. 2014. Vocal fold vibratory patterns in tense versus lax phonation contrasts. *Journal of the Acoustical Society of America* 136: 2784–97. [CrossRef] [PubMed]
- Kuang, Jianjing, and Mark Liberman. 2016. Pitch-range perception: The dynamic interaction between voice quality and fundamental frequency. Paper presented at the Interspeech, San Francisco, CA, USA, September 8–12; pp. 1350–54.
- Kuang, Jianjing, and Mark Liberman. 2018. Integrating voice quality cues in the pitch perception of speech and non-speech utterances. *Frontiers in Psychology* 9: 21–47. [CrossRef]
- Kuang, Jianjing, Yixuan Guo, and Mark Liberman. 2016. Voice quality as a pitch-range indicator. *Proceedings of Speech Prosody 2016*: 1061–65.
- Ladd, Robert. 1996. *Intonational Phonology*, 1st ed. Cambridge: Cambridge University Press.
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press, pp. 1–123.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell, pp. 1–425.
- Laver, John. 1980. *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.
- Leben, William, Sharon Inkelas, and Mark Cobler. 1989. Phrases and Phrase Tones in Hausa. In *Current Approaches to African Linguistics*. Dordrecht: Foris, pp. 45–61.
- Longmailai, Manali. 2012. Personal Pronouns in Dimasa. In *Northeast Indian Linguistics Society 4, CUP*. Edited by Mark W. Post, Gwendolyn Hyslop and Stephen Morey. New Delhi: Foundation Books, pp. 105–18.
- Samek-Lodovici, Vieri. 2005. Prosody-syntax interaction in the expression of focus. *Natural Language and Linguistic Theory* 23: 687–755. [CrossRef]
- Sarmah, Priyankoo. 2009. Tone Systems of Dimasa and Rabha: A Phonetic and Phonological Study. Ph.D. dissertation, University of Florida, Gainesville, FL, USA.
- Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2011. VoiceSauce: A program for voice analysis. Paper presented at the International Congress of Phonetic Science XVII, Hong Kong, China, August 17–21; pp. 1846–49.
- Simpson, Adrian. 2012. The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics* 40: 477–90. [CrossRef]
- Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *Linguistic Review* 20: 37–78. [CrossRef]
- Truckenbrodt, Hubert. 1995. Phonological Phrases: Their Relation to Syntax, Focus, and Prominence. Ph.D. dissertation, MIT, Cambridge, MA, USA.
- Wayland, Ratre, and Allard Jongman. 2002. Registrogenesis in Khmer: A phonetic account. *Mon-Khmer Studies* 32: 101–15.
- Xu, Yi. 1997. Contextual tonal variations in Mandarin. *Journal of Phonetics* 25: 61–83. [CrossRef]

-
- Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27: 55–105. [[CrossRef](#)]
- Xu, Yi, Szu-wei Chen, and Bei Wang. 2012. Prosodic focus with and without post-focus compression: A typological divide within the same language family? *The Linguistic Review* 29: 131–47. [[CrossRef](#)]