Chapter 5: Focus Marking Prosody in Tonal Languages: Evidence from Chokri and Sylheti

5.1. Introduction

In most languages of the world, intonation serves the function of conveying different post-lexical meanings. Apart from marking sentence types and dividing stretches of speech into smaller units, intonation also encodes focus or prominence in languages (Gussenhoven, 2004; Ladd, 1996; Cruttenden, 1997). Studies on focus prosody in many intonation-only languages show that f0 plays a primary role, particularly in the form of prominence-lending pitch accents (Hartmann, 2008). However, prosodic focus through pitch accent is scarce in tone languages. The realization of focus prosody in these languages significantly differs from that of non-tonal languages and thus requires special attention. Cross-linguistically, focus strategies frequently attested in tone languages over the words include modification of pitch (on target words or/and pre-focal and post-focal domains), prosodic phrasing, and manipulation of duration and/or intensity (Xu, 1999; Kügler and Genzel, 2011; Pan, 2007; Inkelas and Leben, 1990; Mahanta, Das, and Gope, 2016). Using the same pitch property for lexical and post-lexical meaning, like focus in some tone languages, offers exciting insights into tone-intonation interaction. This Chapter aims to explore how Syheti and Chokri encode three types of in-situ focus viz-informational, contrastive, and corrective in object positions through prosodic means.

5.1.1. General Background

5.1.1.1. Focus Marking Strategies

Some portions of an utterance can be focused on or highlighted to signal newness or contrastivity, which requires the hearer's special attention. Focus is one of the primary elements of information structure in any language. Jackendoff (1972) defines focus as ‘the information in the sentence that is assumed by the speaker not to be shared by him and the hearer.’ In terms of information packaging, focus plays a central role as the determinant of surface word order as well as prosodic structure (Chafe, 1976; Prince, 1986; Vallduví, 1991, 1995; Wilbur, 2012). Depending on the size or scope, focus is classified as Broad Focus (where the whole sentence is in focus) and Narrow Focus (where only a phrase or word is focused). Moreover, narrow focus can be of many types, such as informational, contrastive, and corrective focus. Informational or new information focus signals the non-given elements of an utterance, usually found in constituents answering wh-questions. Contrastive focus presents juxtaposition between a pair of exclusive alternatives. Corrective focus occurs on elements that reject an alternative that has been offered in a previous utterance (Gussenhoven, 2008; Buring, 2011).

A language may employ various phonological, lexical, morphological, and grammatical means or a combination thereof for marking focus. In most Intonation only languages like English (Gussenhoven, 2007; Xu and Xu, 2005) and German (Baumann, Grice, and Steindamm, 2006), the prosodic means for encoding focus is placing a nuclear pitch accent on the primary stressed syllable of the focused element. Focus is prosodically marked through *culminativity*, i.e., making the constituent more prominent by accenting it or *demarcation*- adding a juncture at the beginning or the end of the constituent in focus (Downing and Pompino-Marschall, 2013; Zubizarreta, 2010). Another intonational strategy for focus includes prosodic phrasing as seen in languages like Chichewa (Kanerva, 1990), Korean (Jun and Lee, 1998; Lee and Xu, 2010), and Japanese (Pierrehumbert and Beckman, 1988). Some languages like Bangla (Hayes and Lahiri, 1991) and European Portuguese (Frota, 2002) also use a combination of both nuclear pitch accent and phrasing for focus marking.

5.1.1.2. Focus Prosody on Tone Languages

Many works on intonation in tonal languages indicate that languages with lexical tones generally avoid complex intonation systems (Gussenhoven, 2004; Yip, 2002; Cruttendan, 1997; Ladd, 1996). As tonal languages use pitch variation for lexical contrast, the intonational function of the pitch in these languages is relatively limited. Examining whether they mark focus through any prosodic means is useful to understand the extent to which intonation is relevant in these languages. Analysis of prosodic focus marking in a tone language also offers valuable insights into the nature of tone-intonation interactions. Some of the prosodic cues of focus that are seen to be employed in tonal languages include pitch register modification, phrasing, lengthening, alteration of intensity, or post-focal compression (Xu, 1999; Downing and Rialland, 2017).

Based on the type of focus prosody, tonal languages can be broadly categorized into three groups (Kugler and Genzel, 2011). The tonal languages of one group are known for using pitch register modification. This involves changes in f0 scaling on the target constituent and in the post-focal and pre-focal domains. Mandarin, for instance, uses register expansion for focus marking, leading to higher scaling of H and lower scaling of L tone on the constituent in focus (Xu, 1999). Mandarin also uses post-focal compression (PFC) for encoding focus. PFC is seen to be present in Bodo, a two-tone language of North-East India, as well, indicating in-situ focus (Mahanta et al., 2016; Das, 2017). The expansion of pitch on constituents in focus has also been reported in Hausa. In this Chadic language, a local high raising occurs on both subject foci and ex-situ non-subject foci (Leben et al.,1989.) Focus on Bemba is realized through both pre-focus pitch raising and PFC (Kula and Hamann, 2017). Tibetan uses f0 rising, pitch range expansion, and lengthening to indicate focus (Zhang et al., 2012). Akan employs pitch register lowering and phrasing as a prosodic means for marking focus (Kügler, 2017). In Dimasa, another Tibeto Burman language from Northeast India, the morphological focus marker triggers the pitch range expansion of the sentence (Mahanta et al., 2021).

The second group comprises tone languages that use prosodic phrasing for marking focus. The focused element in these languages is distinctly marked as a single prosodic unit with features like final lengthening, pause, and pitch register reset at boundaries. In Nkho-takota Chichewa, for instance, focused elements constitute a phrase marked by penultimate lengthening and tone lowering on the final vowel (Downing et al., 2004). Focus is indicated by an H boundary tone in Kammu (Karlsson et al., 2007). A phonological phrase boundary (that blocks H tone spreading) follows the on-focus element in Shingazidja (Patin, 2017). On the other hand, prosodic marking of focus is not always at work in all the tone languages. A lack of any focus prosody characterizes many tonal languages like Bàsàá (Makasso et al., 2017), Hausa (Hartmann and Zimmermann, 2007), Northern Sotho (Zerbian, 2006), and Navajo (McDonough, 2002). These languages can be grouped together as the third category.

5.2. Experimental Methodology

In a separate experiment conducted to analyze prosodic focus in the languages, five native speakers each from Sylheti and Chokri were chosen. The Sylheti native speakers (three female and two male) were from the Dharmanagar district of north Tripura. The five native Chokri speakers (three female, and two male) were from the Thipuzu village, under the Phek district of Nagaland, India. All the speakers of both languages were between the age group of 20 and 33 years.

5.2.1. Materials and Procedure

For these production experiments, utterances containing objects in focus were elicited in the form of scripted sentences in Chokri and Sylheti. The target words specified for different lexical tones were considered to examine the effect of focus prosody on contrastive tones. Sentences with in-situ informational focus on objects were elicited as responses to a wh-question. An example of the sentence frame used is shown below:

Q: Subject QW Verb_J?

In-situ focus: Subject $[Y]_{\text{narrow foc}}\text{Verb}_J$

A yes/no question-answer template was used to prompt objects with corrective focus with the following sentence frame:

Q: Subject [X] Verb_J?

Corrective Focus: NO, Subject $[Y]_{\textbf{correfoc}}\text{Verb}_J$

Sentences with contrastive focus were elicited as embedded sentences, in which the object of the matrix clause was in contrast with that of the embedded clause:

Subject [X] Verb_i BUT Subject $[Y]_{\textbf{contfoc}}\text{Verb}_J$

Here, Verb_i and Verb_J are replaced with two different verbs. The target word replaces [Y] and is realized with any of the three kinds of focus under consideration. The corpus for Sylheti consisted of 44 sentences, 11 sentences in four focus conditions. Similarly, 32 sentences were considered for the analysis of Chokri focus elements (See Appendix I and II for the complete dataset used for the study)

A portable recorder (Tascam DR 100) connected with a unidirectional head-worn microphone (Shure SM10A) was used to record the sentences produced by the subjects. The recordings were digitized at a sampling frequency of 44.1 kHz and 32-bit resolution. Each subject repeated the dataset containing different sentence types five times with a considerable pause between each repetition. One of the native speakers supplied questions on in-situ, contrastive, and corrective focus. The best five repetitions are considered for the analysis. Ten tokens had to be discarded due to the presence of abrupt pauses. The total sentence tokens observed for these two separate production experiments conducted in two different languages is 1840 (32 sentences in Sylheti x 5 subjects x 5 repetitions) + (42 sentences in Chokri x 5 subjects x 5 repetitions) - 10= 1840.

5.2.2. Data Annotation and Acoustic Measure

The individual sound files of each sentence are segmented at the word and syllable levels using PRAAT (version 6.0.43) (Boersma and Weenick, 2012). The acoustic components considered in these experiments include mean f0, duration, and intensity. The values for all three parameters are extracted using the *ProsodyPro* (Xu, 2013) Praat script. The resulting values are compared between the sentences carrying in-situ focus, contrastive focus, and corrective focus with their broad focus counterparts.

For the normalized pitch value, the pitch was measured at ten consecutive points starting from the onset until the offset of the TBU (word in Sylheti, syllable in Chokri). The averaged f0 of the time-normalized values of each speaker's production of sentences in different focus conditions were drawn to visualize and detect the uniform patterns, if any. Once a general pattern was observed, the average pitch value of the time-normalized values of full sentences under different focus as well as target words and pre- and post-focal words was drawn. All the data were plotted using *matplotlib* on *Python.*

5.2.3. Statistical Analysis

Statistical analysis and modeling were conducted using the R and RStudioversion 4.0.5 (R Core Team, 2013). The mixed-effects regression was computed using the *lme4* and *lmerTest* libraries (Bates et al., 2014, Sonderegger, 2023). Linear mixed model fit is by REML, and the t-tests use Satterthwaite's method ['lmerModLmerTest']. To answer the first research question, we did two different models: generalized linear modeling (GLM) and linear mixed modeling. We used *glm(response ~ focus)* for generalized modeling and *lmer (response ~ focus + (1|speaker))*, where the speakers are used as random effects. Since this model accounts for random effects by including a random intercept for "speaker," it provides a better approach for handling nested data structures or repeated measurements within the same speaker. It allows us to account for individual differences among speakers and is more robust in the presence of unexplained variance related to speaker-specific variation.

On the other hand, the GLM model does not account for random effects or individual differences among speakers and does not fully capture the variance in the data. Though we did not see a significant difference in the two modeling outputs, we can conclude that the speaker variation and repetitions do not affect the conclusion. Since mixed modeling is more robust, we used it for all the response variables (f0, intensity, and duration), as well as for different sentence positions, subjects, objects, and verbs.

For Chokri, One-way repeated measure ANOVA (RM ANOVA) was conducted to examine the statistical significance of differences observed in terms of f0 and intensity at different locatons of the utterances. The *anova_test()* function is utilized in the RM ANOVA.

5.3. Prosodic focus in Sylheti: Results and Analysis

5.3.1. Focus and Fundamental Frequency

The results of the production experiment show that Sylheti's in-situ informational focus is prosodically marked through pitch register lowering. The target words specified with both (underlying) L and H lexical tones surfaced with a lower f0 than their broad focus counterparts. The lexical tone is maintained as the change causes reduced f0 scaling within the tonal space specified for the H or L tone. This recurs consistently in all the sequences of lexically H or L tone words examined in this study. Apart from the pitch lowering of the target word, the f0 of the pre-focal words is also observed to be lowered in all the combinations examined in this study. On the other hand, the pitch in the post-focal domain remains neutral.

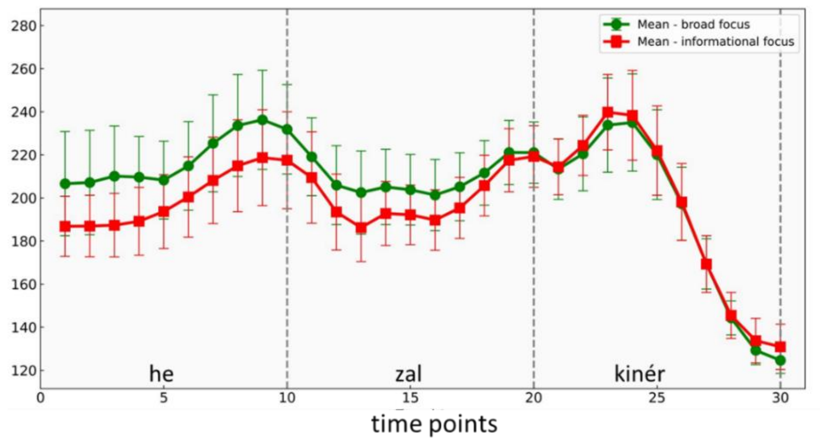

Figure 5.1. Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the sentence [he zal kinér] '*he is buying net*'. The error bars depict the standard error of data post-aggregation.

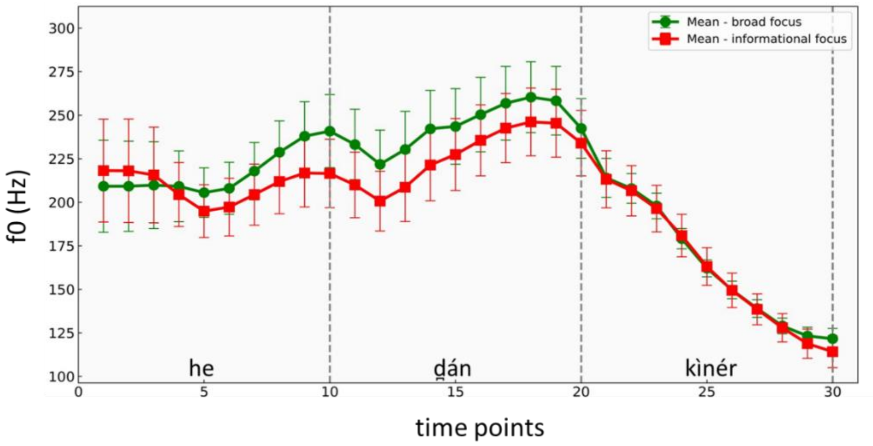

Figure 5.2. Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the sentence [he ḍán kinér] '*he is buying paddy'*. The error bars depict the standard error of data post-aggregation.

Figures 5.1 – 5.2 show the representation of f0 contours of broad focus (green color) and in-situ focus (red color). The target words that carry the in-situ focus, viz the objects [zàl] (in Figure 1) and [ḍán] (in Figure 2), are realized with lowered pitch. In both cases, the pre-focal elements, the subject word [he], 'he,' also undergo pitch lowering.

Contrastive focus in Sylheti is marked by lowering the pitch register on the target word with underlying L or H lexical tones. While the lexical tonal specifications of target words are retained, the effect of contrastive focus is manifested as changes in f0 scaling on the target words, which is lower than their broad focus counterparts. Interestingly, the words preceding the target words have a higher pitch than the corresponding broad-focus sentences. Pitch register compression is observed on the post-focal domains in all the sentences representing different tonal sequences (see Figure 5.3).

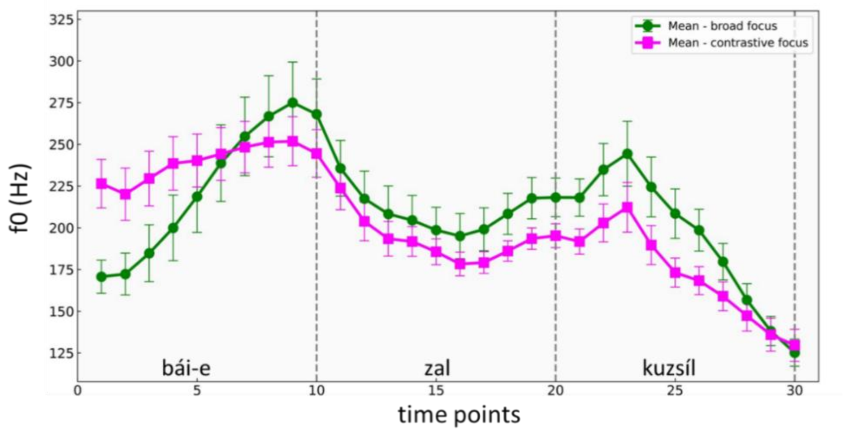

Figure 5.3. Time-normalized f0 contours representing broad focus (in green) and contrastive focus (in magenta); the f0 contours are averaged across all the tokens produced by all the speakers for the sentence [bái-e zal kuzsíl] '*(my) brother was looking for (a) net'*. The error bars depict the standard error of data post-aggregation.

Corrective focus, too, is realized through pitch register lowering on the target words. However, pitch levels of both pre-focus and post-focal domains are realized with a lowered f0 compared to their broad-focus counterparts. This holds true for all the combinations of tonal sequences considered in this study (see **Figure 5.4**)

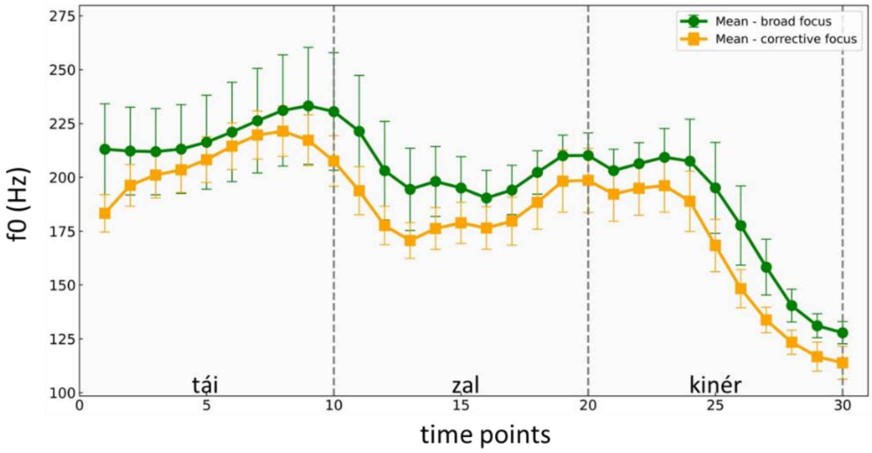

Figure 5.4. Time-normalized f0 contours representing broad focus (in black) and corrective focus (in blue); the f0 contours are averaged across all the tokens produced by all the speakers for the sentence [tái zàl kìnér] *she is buying (a) net.* The error bars depict the standard error of data post-aggregation.

A linear mixed model *lmer(response ~ focus + (1|speaker))* was used where the speakers are used as random effects. This model accounts for random effects by including a random intercept for "speaker" to compare the variable f0 of broad focus sentences w.r.t. the contrastive focus, in-situ, and corrective focus sentences at different positions; subject (pre-focus), object (target), and verb (post-focus).

We used linear mixed modeling with the interaction term formulated with *lmer(object ~*

Table 5.1: Linear mixed model fit: Response (f0) ~ focus type + (1 | speaker)

Comparison		β Estimate	Std. Error	df	t value	Pr(>|t|)
a) f0 at subject	(Intercept)	222.0984	28.1203	3.0042	7.898	0.00421
	Broad vs. corrective focus	-2.6236	0.4775	3289.0031	-5.494	<0.001*
	(Intercept)	223.635	35.178	3.002	6.357	0.00786*
	Broad vs. in-situ focus	-10.219	1.34	3015.0	-7.625	<0.001*
	(Intercept)	249.13	10.12	4.56	24.63	<0.001*
	Broad vs. contrastive focus	-14.52	4.14	326.14	-3.507	0.0005*
b) f0 at object	(Intercept)	224.238	26.218	3.003	8.553	0.00334*

	Broad vs. corrective focus	-6.227	0.402	3291.003	-15.488	<0.001*
	(Intercept)	225.108	33.242	3.001	6.772	0.00657*
	Broad vs. in-situ focus	-8.178	1.058	3001.0	-7.726	<0.001*
	(Intercept)	299.478	7.492	5.376	39.974	<0.001*
	Broad vs. contrastive focus	5.555	3.06	375.274	1.815	0.003*
c) f0 at verb	(Intercept)	174.1817	15.6118	3.0161	11.157	0.00151*
	Broad vs. corrective focus	-4.884	0.5078	3335.0114	-9.617	<0.001*
	(Intercept)	174.449	19.411	3.01	8.987	0.00287*
	Broad vs. in-situ focus	-2.287	1.624	3029.001	-1.408	**0.15917**
	(Intercept)	386.482	13.905	3.61	27.794	<0.001*
	Broad vs. contrastive focus	-10.342	4.676	277.077	-2.212	0.0278*

The linear mixed-effects model was employed to examine the impact of focus type on the response variable (f0) at different positions (subject, object, and verb) while accounting for individual variability with a random effect for the speaker.

The model output, as shown in Table 5.1(a-c), revealed a main effect off0 at the subject, verb, and object positions for different foci— board vs. corrective, board vs. contrastive, and board vs. in-situ informational focus. For the board vs. corrective focus, at the subject position, β, and p-values are found to be- 2.6236, and <0.001. Similarly, (β, p-value) for f0 at the object position is (-6.227, <0.001), and the verb position is (-4.8840, <0.001). Irrespective of the subject, verb, and object positions, f0 of the corrective focus is found to be significantly lower than the board focuses [see **Figure 5.5 (III)].**

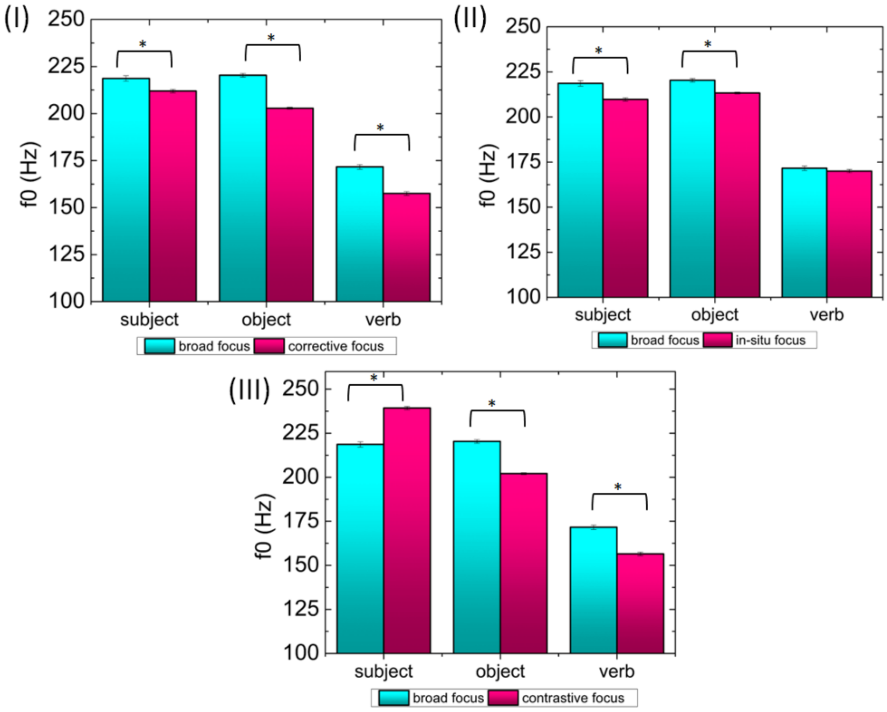

Figure 5.5. The average f0 (in Hertz) for the subject, object, and verb positions: (I) Broad vs. corrective focus, (II) Broad vs. In-situ focus, and (III) Broad vs. Contrastive focus. The asterisk (*) marks the significant difference with $p < 0.05$. The error bars represent the standard error.

Figure 5.5 (II) shows the comparison of the board focus and in-situ informational focus for subject, object, and verb positions. The β and p-values are(-10.219,<0.001) for the subject, (-8.976,<0.001) for the object, and (-2.287, 0.15917) for verb positions. This implies that f0 is significantly lower for in-situ focus from the broad focus on the subject and object positions, whereas no significant effect is found for the verb position.

Interestingly, f0 of the contrastive focus is found to be significantly different from the board focus irrespective of the subject, verb, and object positions, similar to the board-corrective focus case. However, a general trend is not observed. The f0 of the contrastive focus is significantly lower than the board focus by 10-20 Hz for the object and verb positions, whereas a 20 Hz increase in the f0 values is observed at the subject position for the contrastive focus [see **Figure 5.5(III)]**.

5.3.2. Focus and Duration

The linear mixed modeling was also incorporated with the interaction term formulated with *lmer(object ~ subject * verb + (1|speaker)* to compare the variable ‘duration’ of broad focus sentences w.r.t. the contrastive focus, in-situ, and corrective focus sentences at different positions; subject (pre-focus), object (target), and verb (post-focus).

Table 5.2: Linear mixed model fit: Response (Duration) ~ focus type + (1 | speaker)

Comparison		Estimate	Std. Error	df	t value	Pr (>|t|)
a) duration at subject	(Intercept)	239.774	7.238	8.708	33.126	<0.001*
	Broad vs. corrective focus	-10.320	2.941	296.247	-3.509	0.00051*
	(Intercept)	249.120	13.360	3.832	18.647	<0.001*
	Broad vs. in-situ focus	-32.781	9.337	296.973	-3.511	0.00051*
	(Intercept)	249.13	10.12	4.56	24.630	<0.001*
	Broad vs. contrastive focus	-14.52	4.14	326.14	-3.507	0.00051*
b) duration at object	(Intercept)	302.531	9.614	3.337	31.47	<0.001*
	Broad vs. corrective focus	-5.566	1.259	310.047	-4.42	<0.001*
	(Intercept)	298.061	18.316	3.055	16.273	0.00045*
	Broad vs. in-situ focus	8.916	3.331	297.006	2.677	0.00784*
	(Intercept)	299.478	7.492	5.376	39.974	<0.001*
	Broad vs. contrastive focus	5.555	3.060	375.274	1.815	**0.0703**
c) duration at verb	(Intercept)	379.739	15.626	3.585	24.30	<0.001*
	Broad vs. corrective focus	-8.976	2.972	282.044	-3.02	0.00276*
	(Intercept)	385.508	14.794	3.565	26.06	<0.001*
	Broad vs. in-situ focus	-6.191	8.969	296.953	-0.69	**0.491**
	(Intercept)	386.482	13.905	3.610	27.794	<0.001*
	Broad vs. contrastive focus	-10.342	4.676	277.077	-2.212	0.0278*

The linear mixed-effects model was applied to assess how the response variable (duration) at distinct linguistic positions (subject, object, and verb) get affected for different focus (in-situ, contrastive, and corrective focus w.r.t. broad focus), considering individual variability with a random effect for the speaker.

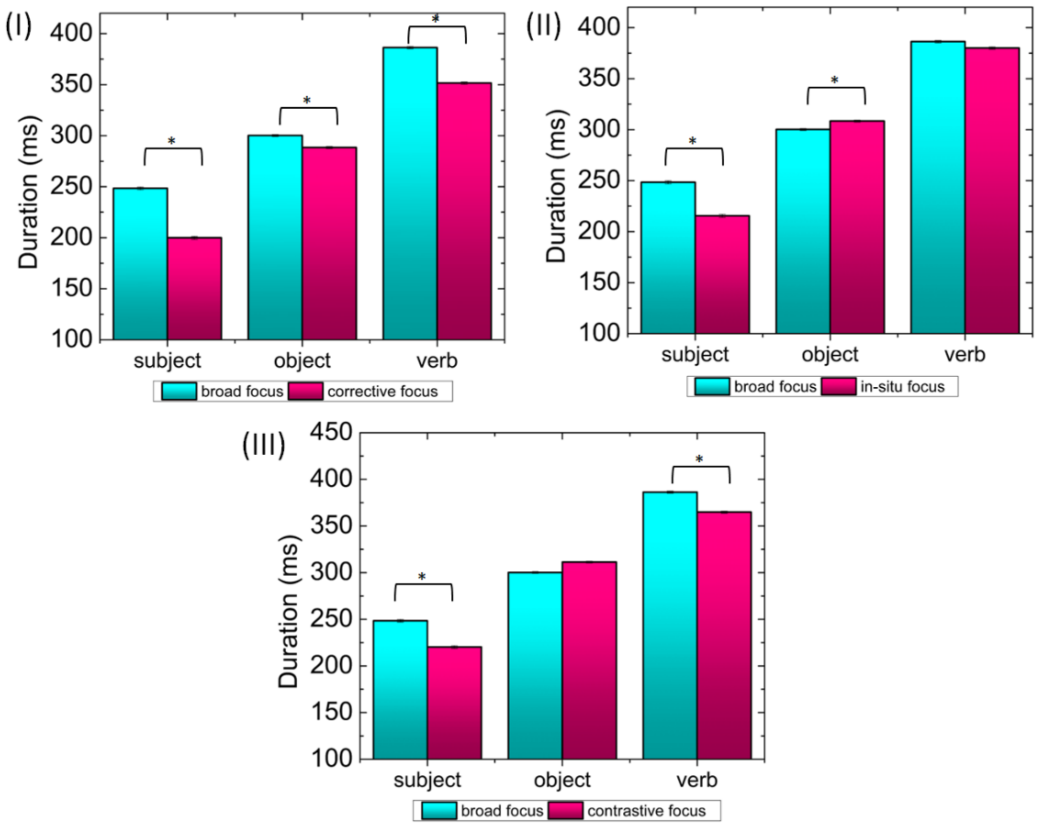

Figure 5.6: average duration (in milliseconds) for subject, object, and verb positions: (I) Broad vs. corrective focus, (II) Broad vs. In-situ focus, and (III) Broad vs. Contrastive focus. The asterisk (*) marks the significant difference with $p < 0.05$. The error bars represent the standard error.

The model results, detailed in Table 5.2(a-c), illuminated significant effects of duration at subject, verb, and object positions for various foci—broad vs. corrective, broad vs. contrastive, and broad vs. in-situ informational focus. For broad vs. corrective focus at the subject position, the estimate (β) and p-value were -10.320 and 0.000519, respectively. Similarly, at the object position, they were (-5.566, <0.001), and at the verb position, they were (-8.976, 0.00276). The average duration of in-situ focus for subject, object, and verb is found to be ~50 ms, ~10 ms, and ~35 ms lower than that of broad focus. Regardless of position, the duration for corrective focus was consistently significantly lower than for broad focus [see **Figure 5.6(I)].**

In **Figure 5.6(II)**, the comparison of broad focus and in-situ focus reveals estimates (β) and p-values of (-32.781, 0.000516) for the subject, (8.916,0.007842) for the object, and (-6.191, 0.491) for the verb positions. This implies that duration is significantly different for in-situ focus than for broad focus at the subject and object positions, while no significant effect is observed at the verb position. The average duration of in-situ focus at the subject position is found to be ~30 ms lower than that of broad focus. It is to be noted that the duration and f0 exhibit a similar trend in terms of statistical significance, i.e., the verb position does not show any difference when the broad focus is compared to the in-situ focus. In contrast, the duration for the contrastive focus at the subject and verb positions is significantly lower than for broad focus by 20-30 ms, whereas at the object position, no significant difference is observed [see **Figure 5.6(III)**].

5.3.3. Focus and Intensity

The same linear mixed modeling with the interaction term formulated with *lmer(object ~ subject * verb + (1|speaker)* was incorporated to compare the variable 'intensity' of broad focus sentences w.r.t. the contrastive focus, in-situ, and corrective focus sentences at different positions; subject (pre-focus), object (target), and verb (post-focus).

Table 5.3: Linear mixed model fit: Response (Intensity) ~ focus type + (1 | speaker)

Comparison		Estimate	Std. Error	df	t value	Pr(>|t|)
a) intensity at subject	(Intercept)	76.6378	1.4008	3.0888	54.710	<0.001*
	Broad vs. corrective focus	-0.7873	0.1075	296.0020	-7.321	<0.001*
	(Intercept)	76.5452	1.3945	3.0473	54.890	<0.001*
	Broad vs. in-situ focus	-0.7395	0.2551	296.9976	-2.899	0.00402*
	(Intercept)	76.606	3.436	329.000	22.296	<0.001*
	Broad vs. contrastive focus	8.344	2.336	329.000	3.572	0.000407*
b) intensity at object	(Intercept)	75.8957	0.8872	3.2687	85.54	<0.001*
	Broad vs. corrective focus	-0.4500	0.1103	310.0136	-4.08	<0.001*
	(Intercept)	75.8029	1.0070	3.0487	75.280	<0.001*
	Broad vs. in-situ focus	0.1586	0.1835	296.9994	0.864	**0.388**
	(Intercept)	75.964	3.181	378.000	23.880	<0.001*
	Broad vs. contrastive focus	4.973	2.053	378.000	2.422	0.0159*
c) intensity at verb	(Intercept)	66.9559	3.1080	3.0440	21.543	0.000199*
	Broad vs. corrective focus	-1.2062	0.1741	282.0026	-6.928	<0.001*
	(Intercept)	66.9307	2.9495	3.0181	22.692	0.00018*
	Broad vs. in-situ focus	-0.5651	0.3280	296.9999	-1.723	**0.0859**
	(Intercept)	67.073	7.206	280.000	9.308	<0.001*
	Broad vs. contrastive focus	22.461	5.307	280.000	4.233	<0.001*

The model results, detailed in Table 5.3(a-c), illuminated significant effects of intensity of the subject, verb, and object positions for various focus conditions; viz., broad vs.

corrective, broad vs. contrastive, and broad vs. in-situ. For broad vs. corrective focus at the subject position, the estimate (β) and p-value are -0.7873 and <0.001. For the object, (β and p-value) = (-0.4500, <0.001), and for the verb (β and p-value) = (-1.2062, <0.001). Regardless of position, the intensity for corrective focus is consistently significantly lower than for broad focus [see **Figure 5.7(I)**], akin to duration [see **Figure 5.6(I)**], and f0 [see **Figure 5.5(I)**]. Interestingly, for broad vs. in-situ focus, the intensity at the subject position is significantly different by β = -0.7395, and p-value= 0.00402. In contrast, the intensity at the object (0.1586, 0.388) and verb (-0.5651, 0.08591) positions are not significantly different when in-situ informational focus and broad focus are compared [see **Figure 5.7(II)**]. Though it is observed that the intensity of contrastive focus is consistently significantly different than that of broad focus for the subject with (β and p-value) = (8.344, 0.000407), the object with (β and p-value) = (4.973, 0.0159), and the verb with (β and p-value) = (22.461, <0.001) similar to the broad-corrective focus scenario, the trend is opposite. The intensity of contrastive focus is more than the broad focus by ~15 dB at the subject position, ~10 dB at the object position, and ~45 dB at the verb position [see Figure **5.7**(III)].

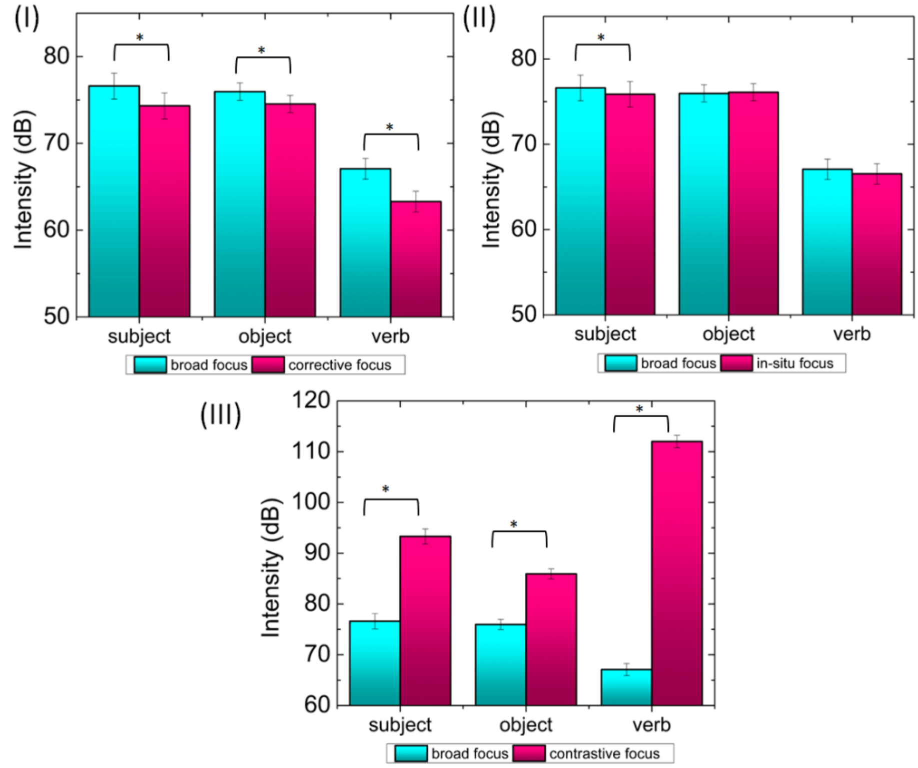

Figure 5.7: The average intensity (in decibels) for the subject, object, and verb positions: (I) Broad vs. corrective focus, (II) Broad vs. In-situ focus, and (III) Broad vs. Contrastive focus. The asterisk (*) marks the significant difference with $p < 0.05$. The error bars represent the standard error.

This analysis reveals a distinct comparison of focus type on acoustic features (variables, such as f0, intensity, and duration) across sentenced positions. The corrective focus consistently exhibits lower values for f0, duration, and intensity compared to broad focus. In-situ focus shows a significant difference for f0 and duration but not for intensity. In contrast, the contrastive focus, while consistently differing from broad focus, exhibits contrasting trends across the three features, demonstrating the focus induced variations of three different acoustic parameters.

5.4. Prosodic focus in Chokri: Results and Analysis

5.4.1. Focus and f0

Informational focus

Visual inspection of f0 contours of the utterances suggested a global lowering of pitch register in utterances carrying target words with informational focus. When compared to their broad focus counterparts, the target words, along with the pre-focal and post-focal elements, showed a lower f0 scaling. This is shown in **Figure 5.8.**

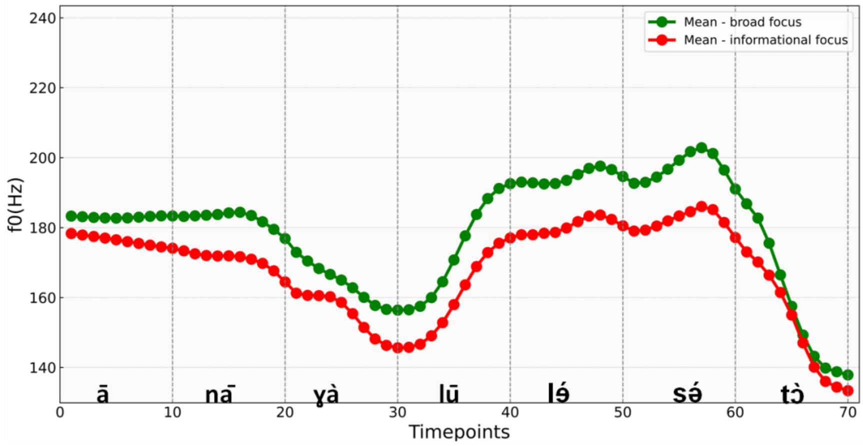

Fig 5.8: time normalized average f0 of the sentence [ā-nā ɣàlū lə́ sə́ tɔ̀] *'my aunt will cut the potatoes'* produced in broad focus (green) and informational focus on object (red) of all tokens by five speakers

The figure above shows a comparison of time normalized f0 values of the sentence [ā-nā ɣàlū lə́ sə́ tɔ̀] produced in broad focus condition (in green) and when the object [ɣàlū] 'potato' receives informational focus (in red). The overall contour of the one with focus is consistently lower from the onset of the utterance till the end. This global lowering is further evident when the pitch of the target word and the post-focal and pre-focal elements of the focused sentence are compared to that of their broad focus counterparts.

Figure 5.9 shows the f0 lowering induced by the informational focus on the target words. This kind of lowering was observed in all the sentences examined for this study.

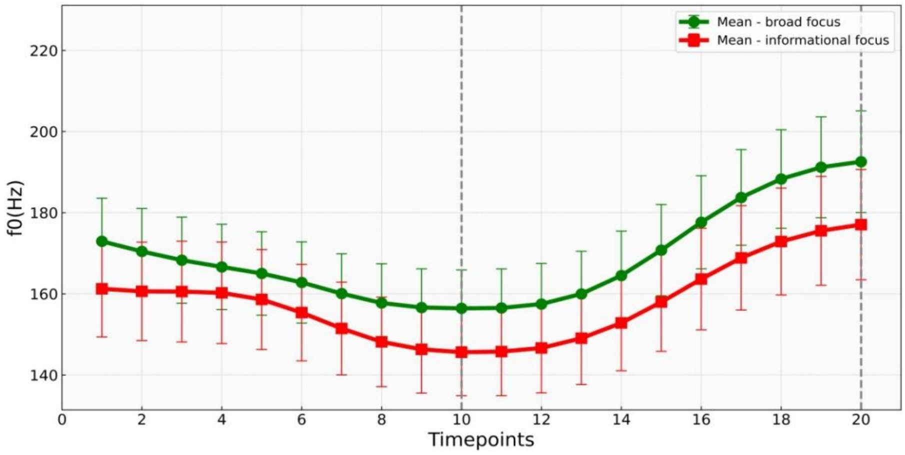

Figure 5.9: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the target word [ɣàlū] *'potato'* spoken in broad focus (green) and informational focus (red). The error bars depict the standard error of data post-aggregation.

In order to examine whether the observed differences of f0 scaling between the sentences with and without focused elements, a one-way *repeated measure ANOVA* (RM ANOVA) was performed. The AnovaRM function from the statsmodels.stats.anova in Python. The one-way repeated measure ANOVA was carried out by converting the 'speaker' as the categorical variable. The syntax *AnovaRM (data=data, depvar='dependent variable', subject='speaker', within=['type'], aggregate_func='mean').fit()* is used, where the dependent variable is f0. The $p < 0.05$ is considered as statistically significant. The results of the one-way repeated measures ANOVA revealed a significant effect of focus type on the dependent variable, $F(7, 28) = 15.2723$, $p < .0001$, indicating that the differences of f0 between sentences with the object in informational focus and their broad focus counterparts are statistically significant (**Figure 5.10**).

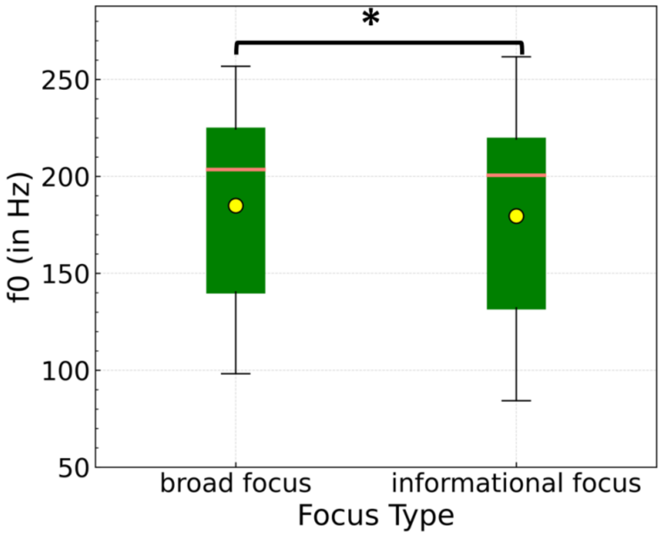

Figure 5.10: f0 (in Hz) of words carrying in situ informational focus and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with $p < 0.05$.

We also examined the differences in the f0 of words occurring at pre-focal and post-focal positions and that of their broad focus counterparts. In both positions, a pitch reduction induced by focus can be seen (**Figure 5.11**).

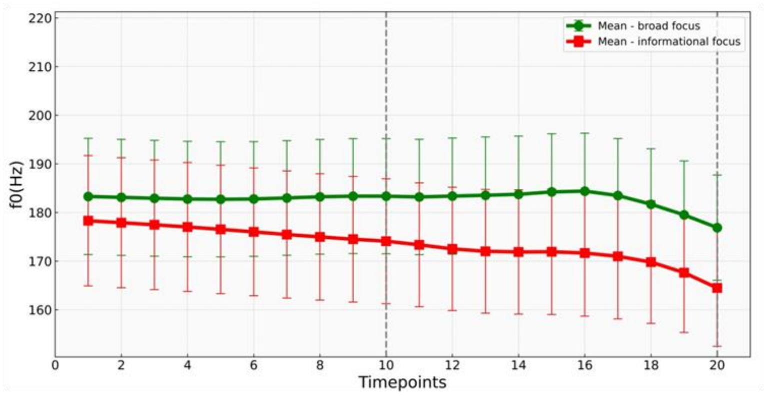

Figure 5.11: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the word [ā-nā] *'my-aunt'* at pre-focal position spoken in broad focus (green) and informational focus (red). The error bars depict the standard error of data post-aggregation.

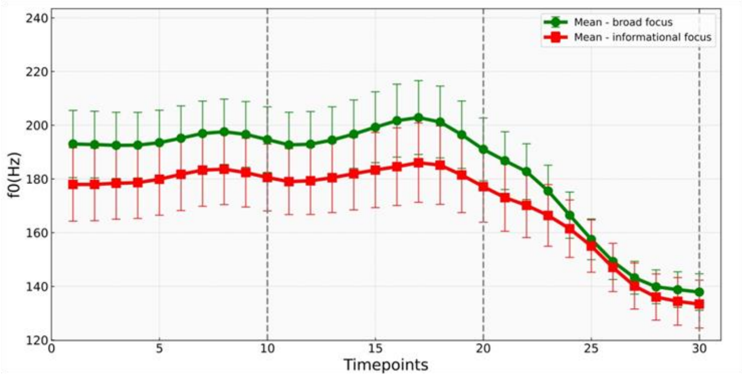

Figure 5.12: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the words [lə́ sə́ tɔ̀] *'will peel'* at post-focal position spoken in broad focus (green) and informational focus (red). The error bars depict the standard error of data post-aggregation.

The RM ANOVA revealed a significant effect of focus type on the f0 of the pre-focal element, $F(7, 28) = 2.4628$, $p = 0.0421$, as well as of the pre-focal element $F(7, 28) = 18.9744$, $p < .0001$ indicating that the differences in mean f0 between broad focus and informational focus sentences are statistically significant in both positions **(Figure 5.13).**

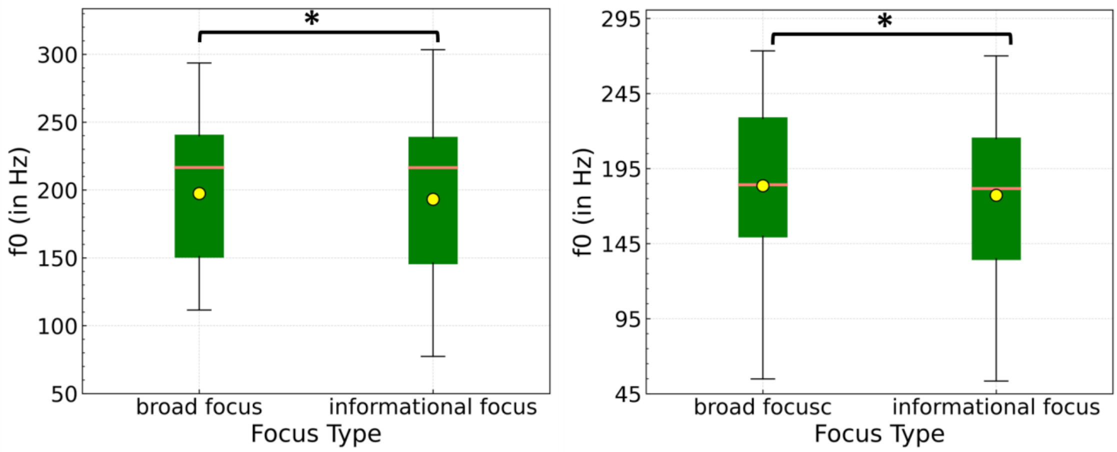

Figure 5.13: f0 (in Hz) of words in pre-focal (left panel) and post-focal (right panel) positions in informational focus sentences and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with $p < 0.05$.

Contrastive focus

Comparison of time normalized average f0 of sentences with objects in contrastive focus and that of sentences in broad focus showed a similar pattern as informational focus utterances. Sentences where objects receive in-situ contrastive focus undergo pitch lowering throughout the sentence (**Figure 5.14**).

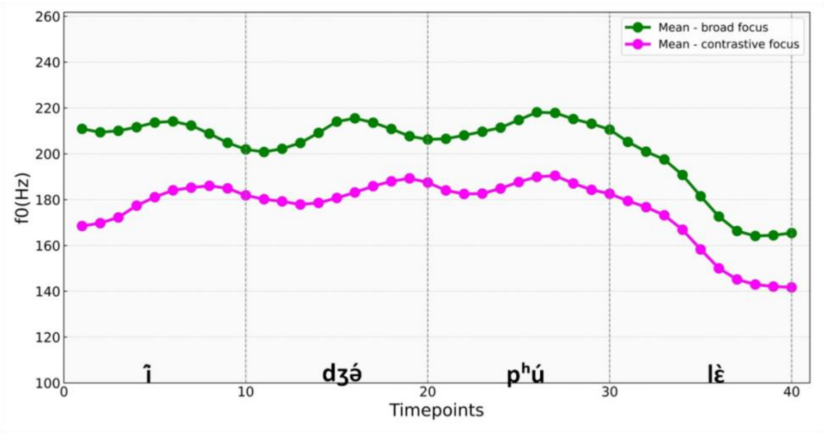

Figure 5.14: time normalized average f0 of the sentence [í dʒə́ pʰú lɛ̀] *'I am searching for some water'* produced in broad focus (green) and contrastive focus on object (purple) of all tokens by 5 speakers

A closer examination of f0 differences in the target words show a consistent lowering affect at all points of the unit. Results of RM ANOVA performed on mean f0 of all tokens of words in contrastive focus and broad focus further confirm statistical significance of f0 difference of target words in focus and their broad focus counterparts $F(5, 20) = 5.1738$, $p = 0.0033$,. **Figure 5.15-5.16** shows the significant difference indicated by RM ANOVA.

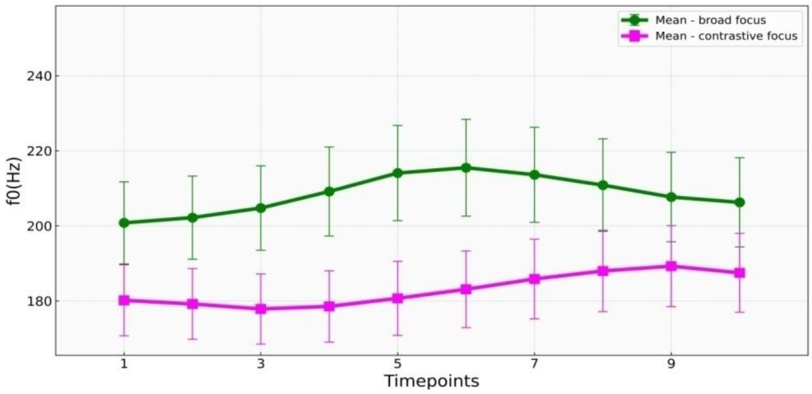

Figure 5.15: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the target word [dʒə́] *'water'* spoken in broad focus (green) and contrastive focus (purple). The error bars depict the standard error of data post-aggregation

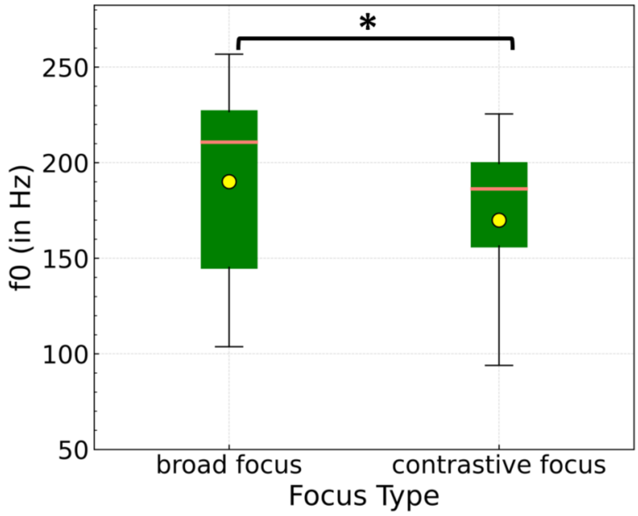

Figure 5.16: f0 (in Hz) of words carrying in situ contrastive focus and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with $p < 0.05$.

Visual inspection of f0 contours suggested that contrastive focus induced pitch lowering impacts both entities at both pre-focal and post focal positions. This was confirmed by results obtained from RM ANOVA conducted to test the statistical significance of the mean f0 difference between pre- and post-focal words in contrastive focus sentences and broad focus sentences.

The outcome of the ANOVA showed significant difference in f0 of pre-focal entities $F(5, 20) = 4.3313$, $p = 0.0078$, and of post focal entities $F(5, 20) = 8.9847$, $p = 0.0001$. This provides evidence to our analysis of global pitch lowering in contrastive focus sentences (see **Figures 5.17** to **5.19**).

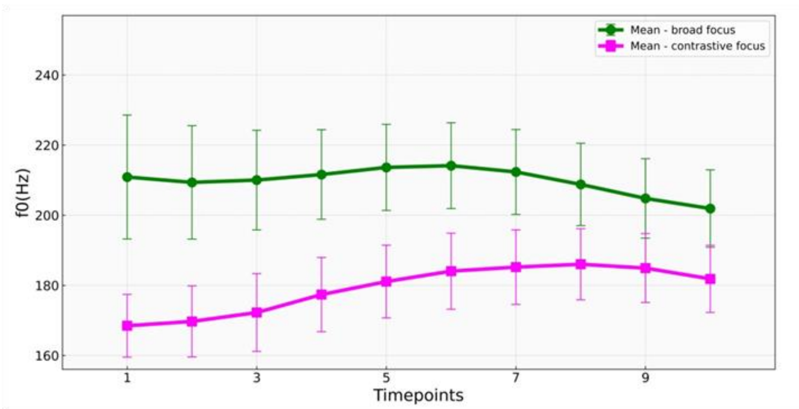

Figure 5.17: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the word [í] *'I'* at the pre-focal position spoken in broad focus (green) and contrastive focus (purple). The error bars depict the standard error of data post-aggregation.

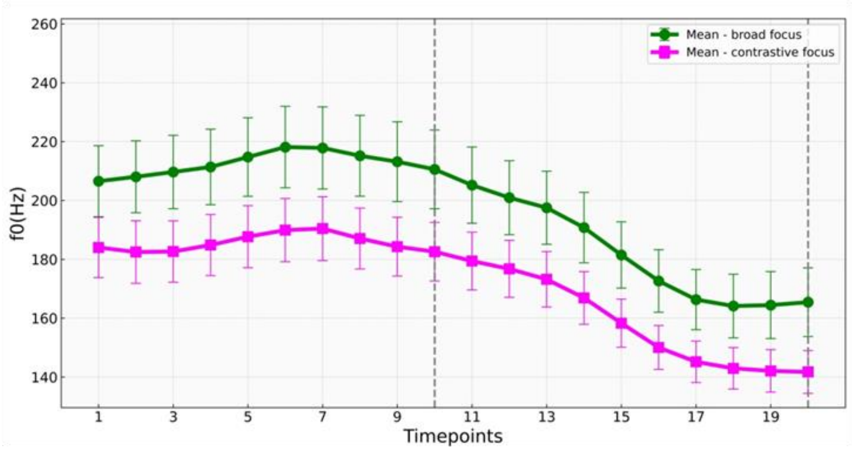

Figure 5.18: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the word [pʰú lɛ̀] *'searching'* at the post-focal position spoken in broad focus (green) and contrastive focus (purple). The error bars depict the standard error of data post-aggregation.

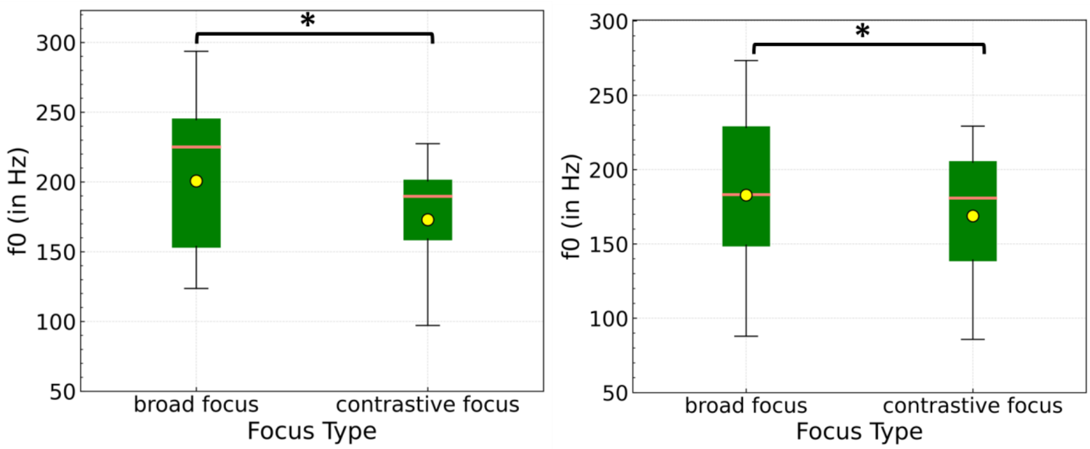

Figure 5.19: f0 (in Hz) of words in pre-focal (left panel) and post-focal (right panel) positions in contrastive focus sentences and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with $p < 0.05$.

Corrective focus:

In terms of f0 contour, corrective focus in Chokri exhibits similar trends as informational focus and contrastive focus. The target word, i.e., the object, along with its preceding and following words, are realized with a lower f0 compared to their broad focus counterparts. Fig shows this global lowering of sentences with corrective focus:

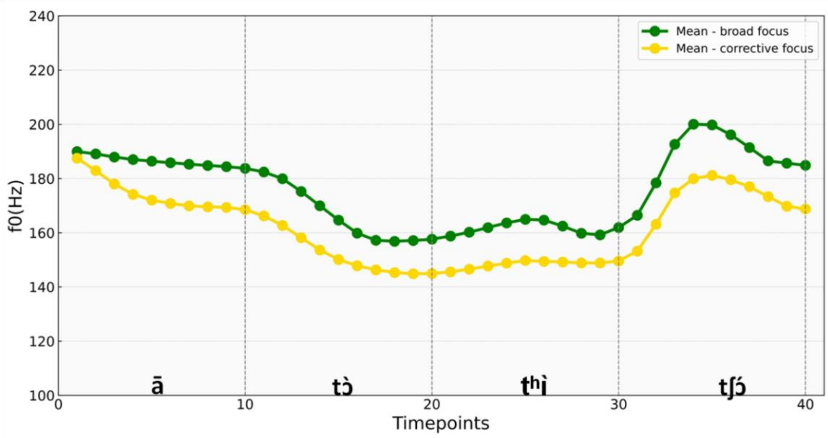

Figure 5.20: time normalized average f0 of the sentence [ātɔ̀ tʰì tʃɔ́] *'Ato cooked meat'* produced in broad focus (green) and corrective focus on object (yellow) of all tokens by five speakers

The f0 values of the object carrying in-situ corrective focus in comparison to the same word produced with the wide focus can be seen in Figure 5.20. Such lowering was seen in all tokens of the data considered for this study regardless of their underlying lexical tone. Results of RM ANOVA performed on mean f0 of all tokens of words in corrective and broad focus further confirm the statistical significance of the f0 difference of target words in focus and their broad focus counterparts $F(1, 4) = 37.9399$, $p = 0.0035$. **Figure 5.22** shows the significant difference indicated by RM ANOVA.

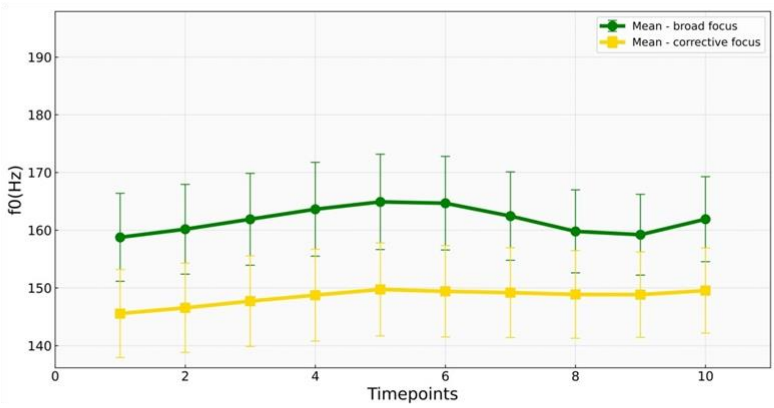

Figure 5.21: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the target word [tʰì] *'meat'* spoken in broad focus (green) and corrective focus (yellow). The error bars depict the standard error of data post-aggregation

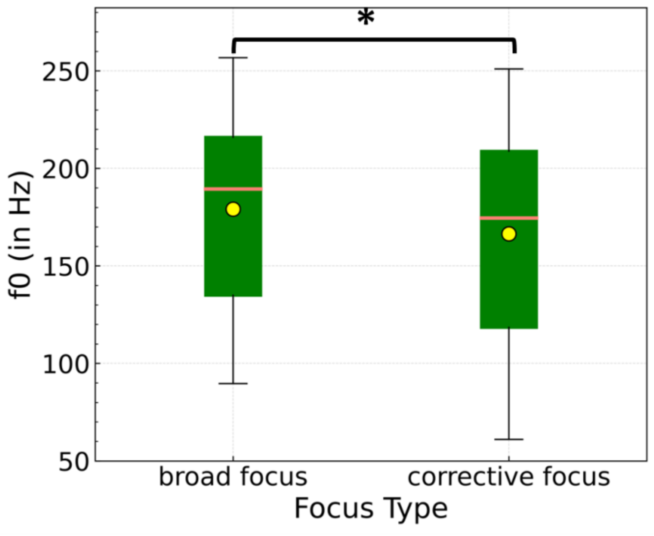

Figure 5.22: f0 (in Hz) of words carrying in situ corrective focus and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with p < 0.05.

RM ANOVA also confirms a significant difference in f0 between words surrounding words with corrective focus and their broad focus counterparts, F(1, 4) = 51.5936 p = 0.0020 (**Figure 5.23-5.25**).

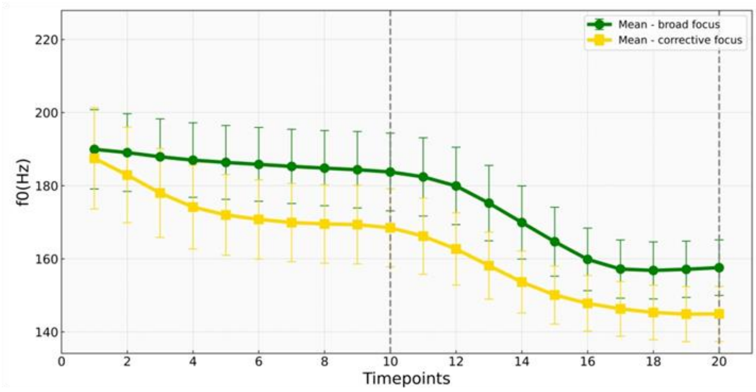

Figure 5.23: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the word [ātɔ̀] *'Ato'* at pre-focal position spoken in broad focus (green) and corrective (purple). The error bars depict the standard error of data post-aggregation.

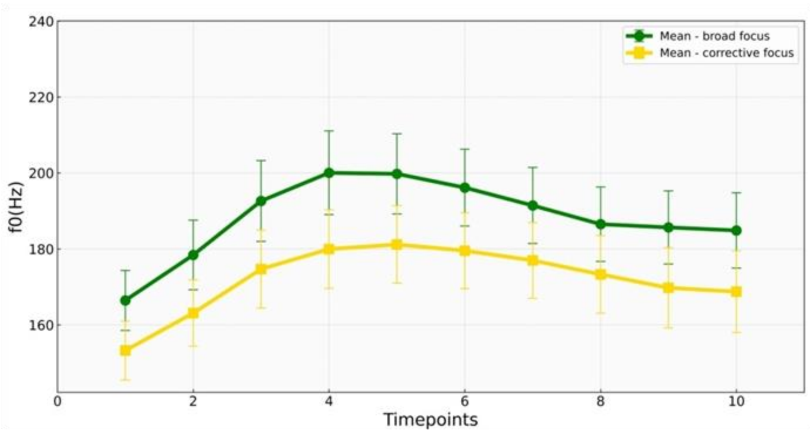

Figure 5.24: Time-normalized f0 contours averaged across all the tokens produced by all the speakers for the word [tʃɔ̃]] *'cook'* at pre-focal position spoken in broad focus (green) and corrective (yellow). The error bars depict the standard error of data post-aggregation.

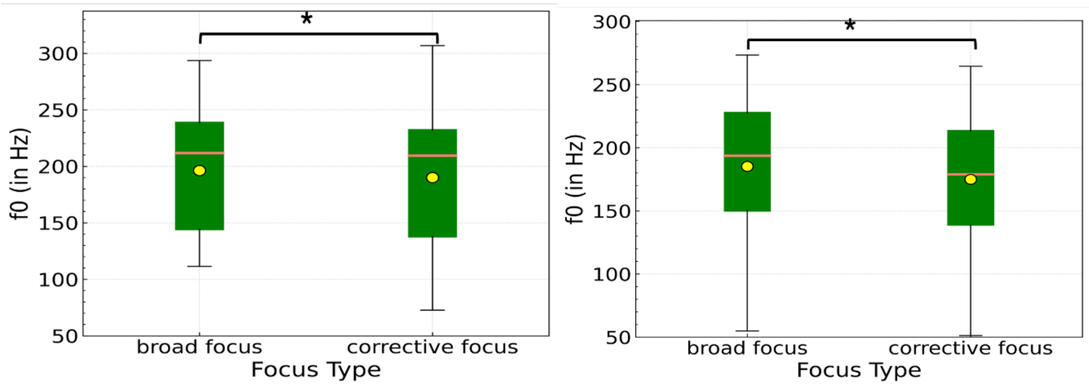

Figure 5.25: f0 (in Hz) of words in pre-focal (left panel) and post-focal (right panel) positions in corrective focus sentences and their broad focus counterparts averaged across all tokens by all five speakers. The error bars depict the standard error of data post-aggregation. The asterisk (*) marks the significant difference with $p < 0.05$.

5.4.2. Focus and Duration

The mean duration of target words, i.e., the objects with in-situ informational, contrastive, and corrective focus and focus adjacent words, were compared to their broad focus counterparts to examine whether duration is a prosodic marker of focus in the language. However, no consistent and uniform pattern of duration variation was observed while analyzing average duration differences between the narrow-focuse and wide-focus utterances in each speaker. In all three focus conditions, viz, informational (**Figure 5.26**), contrastive **(Figure 5.27**), and corrective focus (**Figure 5.28**), there were considerable variations across speakers in all three positions- in focus, pre-focus, and post-focus conditions. This shows that duration is not a prosodic parameter employed for focus in the language.

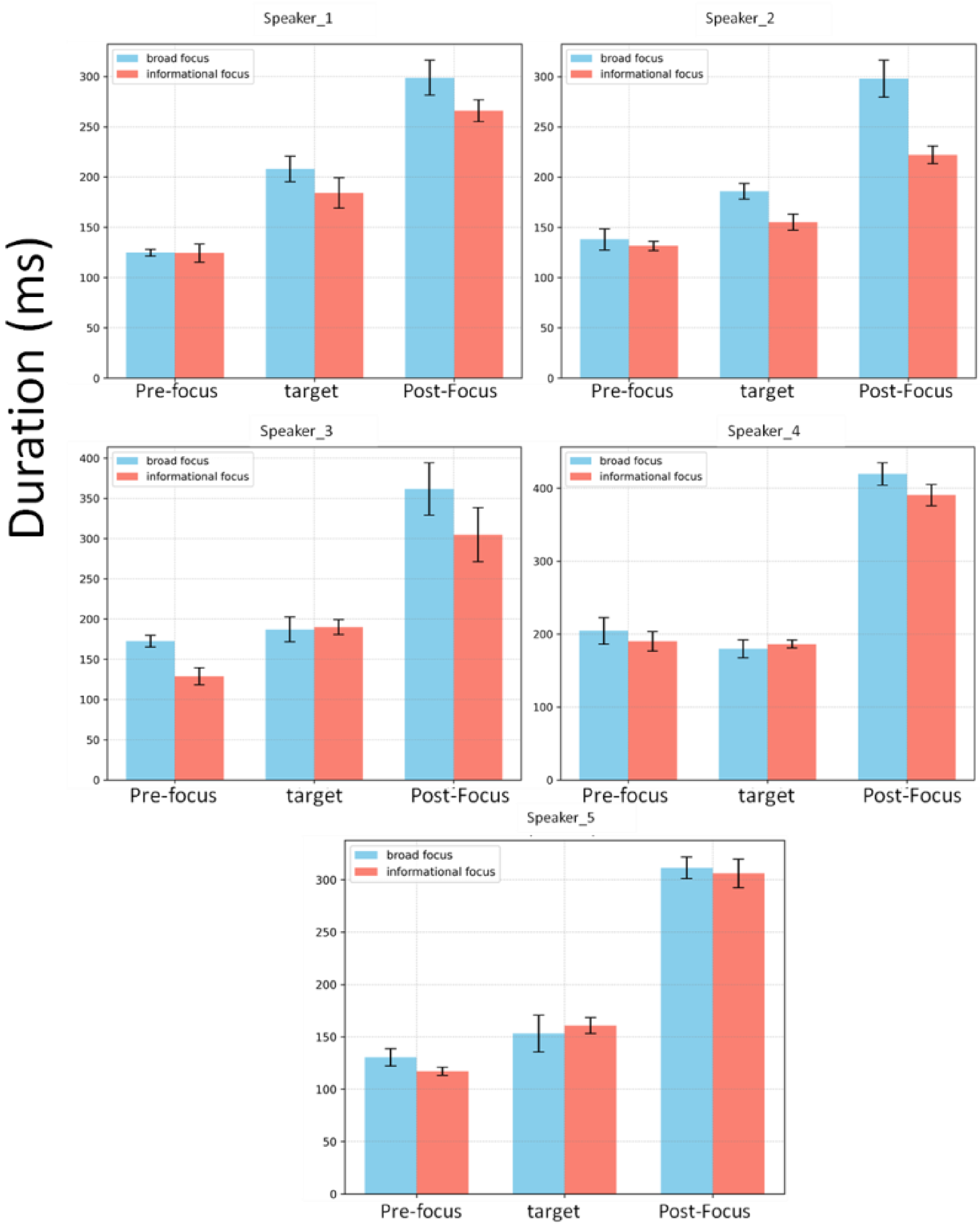

Figure 5.26: speakerwise comparison of average duration of the target word, pre-focal and post-focal elements of the sentence [ātɔ̀ tʰì tʃɔ́] *'Ato cooked meat'* in broad focus (in blue) and informational (in salmon) conditions. The error bars indicate standard error.

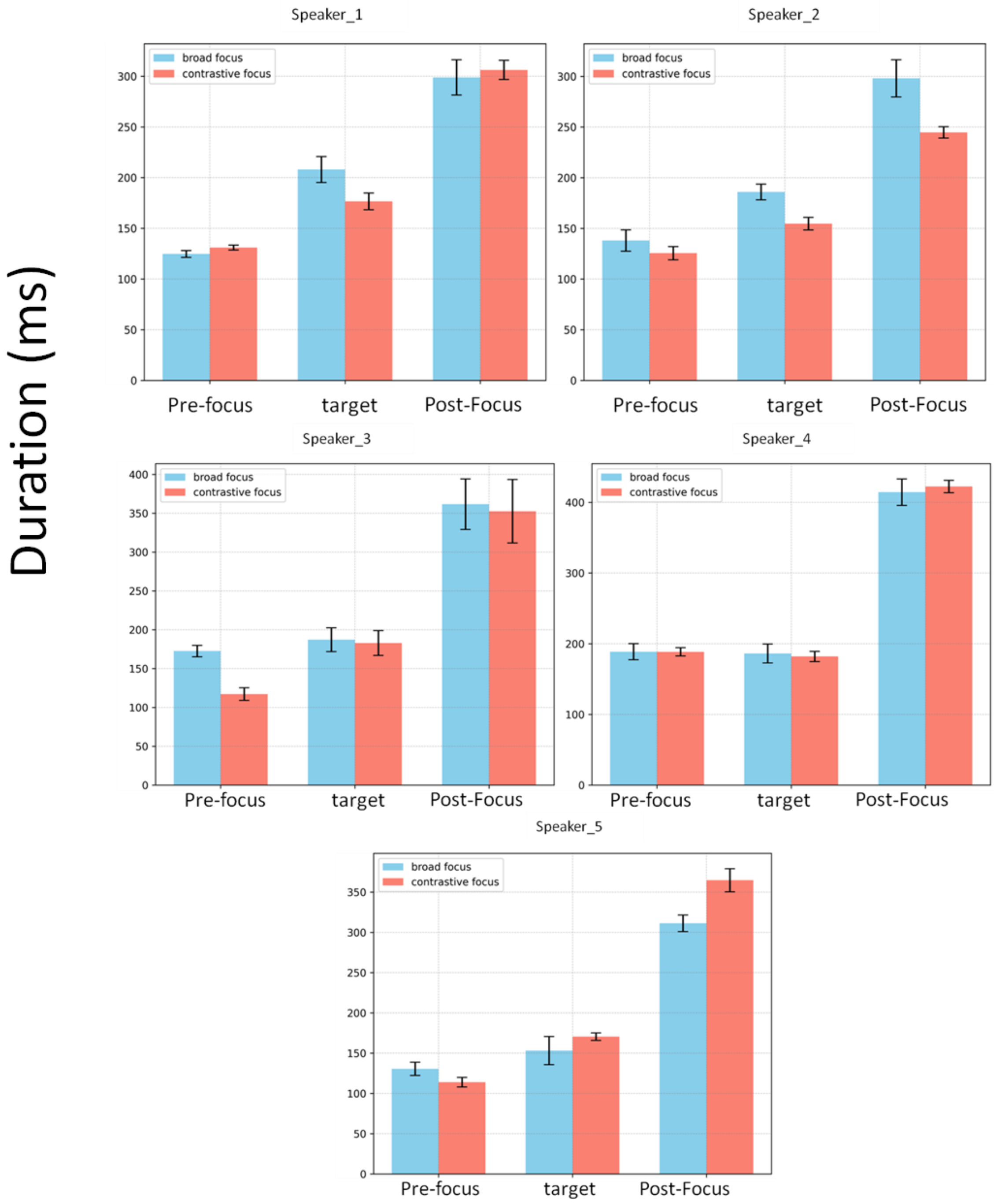

Figure 5.27: speakerwise comparison of average duration of target word, pre-focal and post-focal elements of the sentence [ātɔ̀ tʰì tʃɔ́] *'Ato cooked meat'* in broad focus (in blue) and contrastive focus (in salmon) conditions. The error bars indicate standard error.

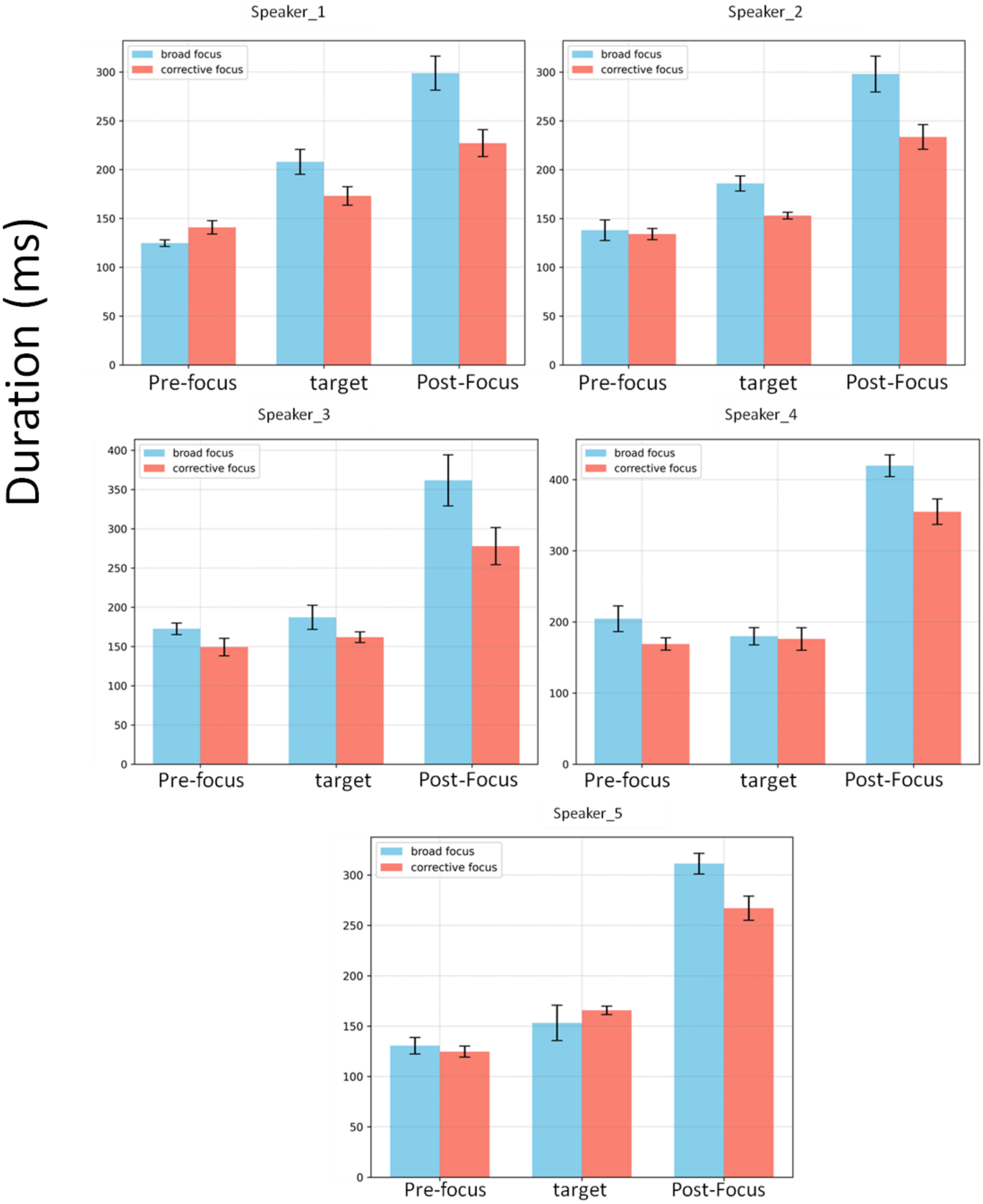

Figure 5.28: speakerwise comparison of average duration of the target word, pre-focal and post-focal elements of the sentence [ātɔ̀ tʰì tʃɔ́] *'Ato cooked meat'* in broad focus (in blue) and corrective (in salmon) conditions. The error bars indicate standard error.

5.4.3. Focus and Intensity

Analysis of intensity differences of elements from focused utterances and the corresponding broad focus utterances showed that sentences with informational and corrective focus on the object are subject to intensity reduction. Reduced intensity was observed not only on the target words but also in the words in pre-focal and post-focal positions. A one RM ANOVA conducted to check the significance of variation observed in focused and non-focused utterances confirmed that focus induces intensity lowering in target ($F(1, 4) = 26.5959$, $p = 0.0067$), pre-focal ($F(1, 4) = 11.6764$, $p = 0.0269$) and post-focal words $F(1, 4) = 7.7451$, $p = 0.0497$) in sentences under corrective narrow focus (**Figure 5.29**). However, no significant difference in intensity was found between words of contrastive focus sentences, informational focus sentences, and those of their broad focus counterparts.

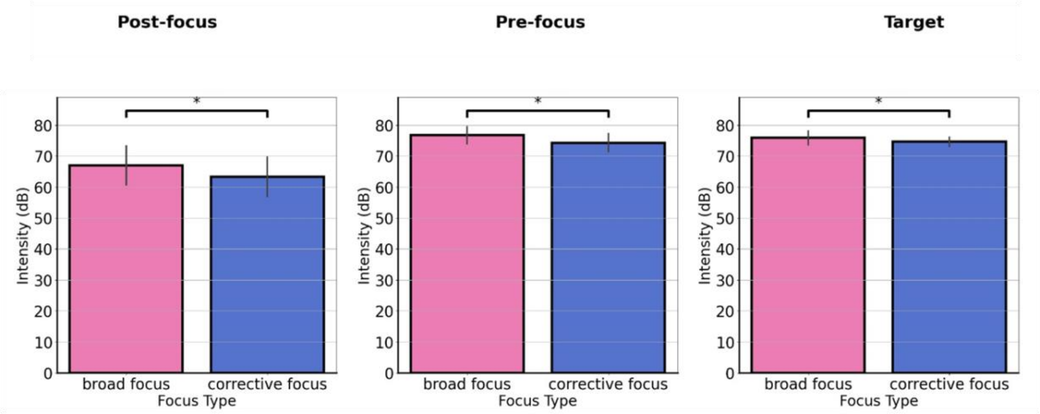

Figure 5.29: Comparison of the average intensity of pre-focal and post-focal elements and target word of the broad focus and corrective conditions. The error bars indicate standard error.

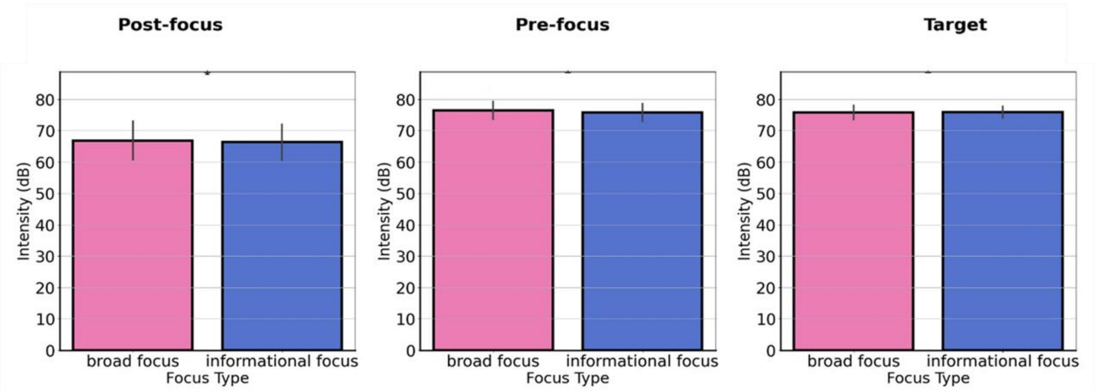

Figure 5.30: Comparison of the average intensity of pre-focal and post-focal elements and target word of the broad focus and informational focus conditions. The error bars indicate standard error.

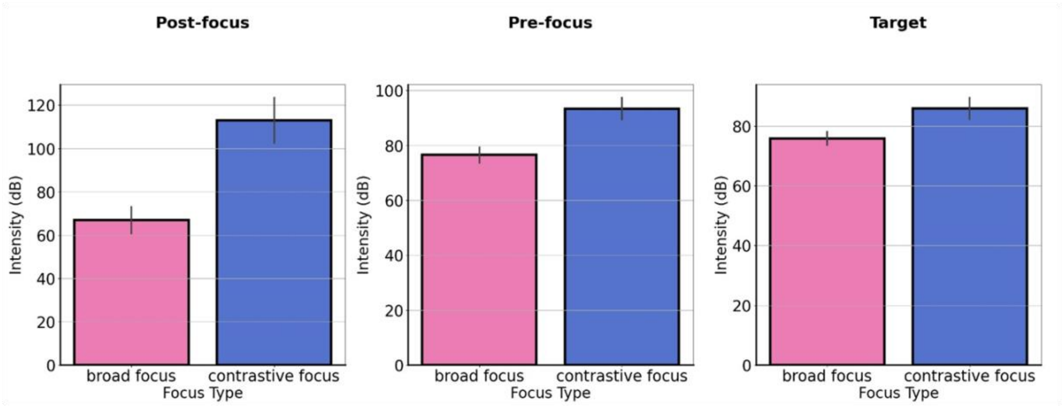

Fig 5.31: Comparison of the average intensity of pre-focal and post-focal elements and target word of the broad focus and contrastive focus conditions. The error bars indicate standard error.

5.5. Discussion and Conclusion:

Cross-linguistic research has shown that while many non-tonal languages use accenting as a strategy for focus marking, it is rare in languages with lexical and/or grammatical tones (Güldemann et al., 2015; Downing and Rialland, 2017). Several tone languages avoid the use of focus prosody to preserve lexical contrast (Makasso, Hamloui and Lee, 2017 for Bàsàá; Hartmann, Zimmermann, 2007 for Hausa; Rialland and Aborobongui, 2017 for Embosi); whereas others make use of strategies that include phrasing, lengthening, pitch register modification on target or/and post-focal and pre-focal domains (Xu, 1999 for Mandarin; Kügler, 2017 for Akan; Mahanta, Das, and Gope, 2016 for Bodo, Kula and Hamann, 2017 for Bemba, Patin, 2017 for Shingazidja). As Sylheti and Chokri are tonal languages with two and five tones, respectively, this study investigated whether they make use of focus prosody or not. Results from the experiments conducted for this paper show that both languages do employ prosodic means to mark prominence. A shared general prosodic mechanism for encoding in-situ focus is present in both languages despite having varied tonal inventories. Pitch register modification is a primary means for prosodic marking of focus in both Chokri and Sylheti. This allows encoding prominence in utterances while preserving the underlying lexical tones of the constituent words. A reduced f0 on target words, in particular, is the most significant finding of this analysis.

In discourses on prosodic focus, a generally accepted assumption is that a focused element is "maximally prominent" in a sentence. It is believed that this maximal prominence is manifested through phonetic features like longer duration, greater intensity, and higher pitch (Büring, 2010; Samek-Lodovici, 2005; Truckenbrodt, 1995). Though results from studies focus on well-documented languages like English (Pierrehumbert, 1980; Gussenhoven, 1983), German (Alter et al., 2001), Mandarin Chinese (Xu, 1999), Japanese (Pierrehumbert and Beckman, 1988; Poser, 1984) and Korean (Jun, 2002) do conform to this assumption, recent works on many other languages indicate that positive correlation between focused elements and prosodic features are not universal. In languages like Chichewa and Chitumbuka (Downing, 2008), the prosodically prominent element in the sentence does not occur on the focused word. The assignment of prosodic prominence in these languages is conditioned by the sentential position. This notion of universality in focus prosody is also reflected in Gussenhoven's (2002, 2004) theory on the 'three biological codes of intonation.' According to this theory, variations in the biological configuration of vocal folds and the physical processes involved in pitch production underlie many grammaticalized phonological intonation patterns. One of these three biological codes is the effort code, which implies that greater effort put into production results in wider pitch excursions and is grammaticalized as focus. The association of nuclear pitch accent with focused words as seen in intonation-only languages like English (Gussenhoven, 2007; Xu and Xu, 2005) and German (Baumann et al., 2006), deaccentuation or f0 compression in post focal elements reported in Mandarin (Xu, 1999), English (Xu and Xu, 2005), German (Fery and Kügler, 2008) and Hindi (Kügler, 2020) are often interpreted as a realization of the effort code. The effort code assumes that parts of utterances are made more prominent by increased effort in production by the speaker. This entails that the pitch on the constituent will be raised and produced with increased duration and higher intensity when it receives focus. However, our analysis of in-situ focus in Sylheti and Chokri shows that prosodic focus marking does not always result in wider pitch expansion, causing a rise in the pitch register.

Contrary to the general expectation of focus attracting a higher pitch, prosodic focus in Sylheti and Chokri involves lowering the pitch register on target words. In all three kinds of narrow focus on objects in the two languages, the pitch of the target words undergoes lowering. Such rare instances of focus-induced reduction are also reported in Akan, where words in corrective focus are realized with a reduction in f0 scaling (Kügler and Genzel, 2011). A similar observation was also made in Bemba (Kula and Hamann, 2017), in which subjects and objects undergo dislocation to be focused and are realized with a compressed pitch. Bodo, another Tibeto Burman tonal language that is arealy proximate to Sylheti and Chokri, also makes use of pitch lowering to encode in-situ contrastive and corrective focus (Das, 2017). Kugler and Genzel (2011) have interpreted this focus-induced tonal lowering as a deviation from 'the effort code.'

Our results indicate that the prosodic focus marking in tone languages fundamentally involves a deviation from the non-focused counterparts in terms of one or more prosodic parameters. This deviation can be projected in either direction and it is not limited to the domain of focus (target word) only. This further exemplifies Kügler and Genzel's (2011) observation that the prosodic marking of focus does not involve deviation from the neutral register only in a particular direction. In addition to that, our study also shows that to create a distinction between different types of focus, deviation may be extended to pre- or post-focal domains in either direction-raising or lowering. To highlight the words that are more focused, Sylheti employs deviation in terms of three parameters: pitch register, duration, and intensity. In addition, to encode different categories of focus, the language makes use of the pre and post-focal domains. For the informational focus on the object, pitch lowering affects the pre-focal (subject) elements. In the case of contrastive focus, it is marked by the pitch of pre-focal raising and post-focal compression. The corrective focus, on the other hand, affects the entire utterance as it induces a global pitch lowering. This strategy particularly handles the need to balance lexical and intonational functions of the same prosodic feature, i.e., pitch. Extending pitch deviation to pre- and/or post-focal domains to achieve distinction between three focus types ensures that there is not as much pitch modification on the target word, which could obliterate the lexical tone. Chokri, on the other hand, primarily uses global pitch register lowering for all three types of focus, suggesting that focus prosody in the language is not employed to distinguish between focus types.

In terms of the other two parameters, viz., duration and intensity too, the differences between the broad focus and the other focused utterances that are statistically significant are all reductions instead of enhanced efforts. In Sylheti, the duration of the target words and/or words surrounding the word in focus in sentences carrying focused constituents is considerably reduced for informational (target and pre-focal words, corrective (target as well as both pre and post-focal words), and contrastive focus (pre-focus and post-focus elements). The focus induced lengthening is also not seen in the languages. Changes in intensity are also a strategy to indicate focus, as seen in the overall intensity reduction for corrective focus in Chokri; and the contrastive and corrective focus in Sylheti. The lowering of the pitch and reduced duration and intensity of the target word for encoding focus thus suggests a deviation from the expected prosodic effects of focus, such as increased pitch, duration, and intensity.

Lowering of acoustic parameters for in-situ focus has been observed in three tonal languages of North-East India, Bodo (Das, 2017), along with Chokri and Sylheti. This observation points towards the possibility of a shared prosodic strategy for focus marking in the tonal languages of the area. However, more studies on other languages from the region need to be undertaken to gain a better understanding of this phenomenon.

References:

Alter, K., Mleinek, I., Umbach, C., & Rohe, T. (2001). Kontrastprosodie in Sprachproduktion und -perzeption. In A. Steube & C. Umbach (Eds.), *Linguistische Arbeitsberichte 77* (pp. 59-81).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumann, S., Grice, M., & Steindamm, S. (2006). Prosodic marking of focus domains - Categorical or gradient? In *Proceedings of Speech Prosody 2006* (Paper 065). https://doi.org/10.21437/SpeechProsody.2006-73

Boersma, P., & Weenik, D. (2012). *Praat: Doing phonetics by computer* (Version 5.3.04_win64) [Computer program].

Büring, D. (2011). Focus and intonation. In G. Russell & D. Graff Fara (Eds.), *Routledge Companion to Philosophy of Language* (pp. 197-209). Routledge.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25-55). Academic Press.

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge University Press.

Das, K. (2017). *Tone and intonation in Bodo* (PhD Dissertation). Indian Institute of Technology, Guwahati.

Downing, L. J., Mtenje, A., & Pompino-Marschall, B. (2004). Prosody and information structure in Chichewa. *ZAS Papers in Linguistics*, *37*, 167–186. https://doi.org/10.21248/zaspil.37.2004.248

Downing, L. (2008). Focus and prominence in Chichewa, Chitumbuka, and Durban Zulu. *ZAS Papers in Linguistics*, *49*, 47–65.

Downing, L. J., & Pompino-Marschall, B. (2013). The focus prosody of Chichewa and the Stress-Focus constraint: A response to Samek-Lodovici (2005). *Natural Language & Linguistic Theory*, *31*, 647–681.

Downing, L., & Rialland, A. (Eds.). (2017). *Intonation in African tone languages*. De Gruyter Mouton. https://doi.org/10.1515/9783110503524

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new, and focused constituents in German. *Journal of Phonetics*, *36*(4), 680–703. https://doi.org/10.1016/j.wocn.2008.05.001

Frota, S. (2002). Tonal association and target alignment in European Portuguese nuclear falls. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 387-418). Mouton de Gruyter.

Güldemann, T., Zerbian, S., & Zimmermann, M. (2015). Variation in information structure with special reference to Africa. *Annual Review of Linguistics*, *1*(1), 155–178. https://doi.org/10.1146/annurev-linguist-030514-125134

Gussenhoven, C. (1983). Focus, mode, and the nucleus. *Journal of Linguistics*, *19*(2), 377–417. https://doi.org/10.1017/S0022226700007799

Gussenhoven, C. (2002). *Intonation and interpretation: Phonetics and phonology*. https://doi.org/10.21437/speechprosody.2002-7

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge University Press.

Gussenhoven, C. (2008). Types of focus in English. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Studies in Linguistics and Philosophy* (Vol. 82, pp. 83–102). Springer. https://doi.org/10.1007/978-1-4020-4796-1_5

Hayes, B., & Lahiri, A. (1991). Bengali intonational phonology. *Natural Language and Linguistic Theory*, *9*, 47–96. https://doi.org/10.1007/BF00133326

Inkelas, S., & Leben, W. (1990). Where phonology and phonetics intersect: The case of Hausa intonation. In *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (pp. 17–34).

Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. MIT Press.

Jun, S. A., & Lee, H. J. (1998). Phonetic and phonological markers of contrastive focus in Korean. In *Proceedings of the 5th International Conference on Spoken Language Processing* (Vol. 4, pp. 1295–1298). Sydney, Australia.

Jun, S. A. (2002). Syntax over focus. In J. H. L. Hansen & B. Pellom (Eds.), *Proceedings of International Conference on Spoken Language Processing* (pp. 2281–2284). Denver, Colorado.

Kanerva, J. (1990). *Focus and phrasing in Chichewa phonology*. Garland.

Karlsson, A. M., House, D., Svantesson, J. O., & Tayanin, D. (2007). Prosodic phrasing in tonal and non-tonal dialects of Kammu. *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1309–1312). https://lucris.lub.lu.se/ws/files/6181030/701750.pdf

Kügler, F., & Genzel, S. (2012). On the prosodic expression of pragmatic prominence: The case of pitch register lowering in Akan. *Language and Speech, 55*(3), 331–359.

Kügler, F. (2017). Tone and intonation in Akan. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 89–117). De Gruyter Mouton.

Kügler, F. (2020). Post-focal compression as a prosodic cue for focus perception in Hindi. *Journal of South Asian Linguistics, 10*(2), 38–59.

Kula, N., & Hamann, S. (2017). Intonation in Bemba. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 321–364). De Gruyter Mouton. https://doi.org/10.1515/9783110503524-010

Ladd, D. R. (1996). *Intonational phonology*. Cambridge University Press.

Leben, W. R., Inkelas, S., & Cobler, M. (1989). Phrases and phrase tones in Hausa. In P. Newman & R. Botne (Eds.), *Current approaches to African linguistics* (pp. 105–124). Foris.

Lee, A., & Xu, Y. (2012). Revisiting focus prosody in Japanese. In *Proceedings of Speech Prosody 2012* (pp. 274–277). Shanghai, China.

Mahanta, S., Das, K., & Gope, A. (2016). On the phonetics and phonology of focus marking in Boro. *Proceedings of the Annual Meetings on Phonology, 3*. https://doi.org/10.3765/amp.v3i0.3668

Mahanta, S., Gope, A., & Raychoudhury, P. (2021). Pitch range and voice quality in Dimasa focus intonation. *Languages, 6*(4), 185. https://doi.org/10.3390/languages6040185

Makasso, E., Hamlaoui, F., & Lee, S. (2017). Aspects of the intonational phonology of Bàsàá. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 167–194). De Gruyter Mouton.

McDonough, J. (2003). The prosody of interrogative and focus constructions in Navajo. In *Linguistik aktuell* (pp. 191–206). https://doi.org/10.1075/la.62.15mcd

Pan, H. (2007). Focus and Taiwanese unchecked tones. In C. Lee, M. Gordon, & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (pp. 195–213). Dordrecht: Springer.

Patin, C. (2017). Tone and intonation in Shingazida. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 147–166). De Gruyter Mouton.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* [PhD thesis]. MIT.

Pierrehumbert, J. B., & Beckman, M. E. (1988). *Japanese tone structure*. MIT Press.

Poser, W. J. (1984). *The phonetics and phonology of tone and intonation in Japanese* (PhD thesis). MIT.

Prince, E. F. (1986). On the syntactic marking of presupposed open propositions. Retrieved from http://www.linguist.univ-paris-diderot.fr/~amsili/papers/opcls.pdf

Rialland, A., & Aborobongui, M. E. (2017). How intonation interacts with tones in Embosi (Bantu C25), a two-tone language without downdrift. In L. J. Downing & A. Rialland (Eds.), *Intonation in African tone languages* (pp. 195–221). De Gruyter Mouton.

Samek-Lodovici, V. (2005). Prosody-syntax interaction in the expression of focus. *Natural Language & Linguistic Theory*, *23*, 687–755. https://doi.org/10.1007/s11049-005-2598-1

Sonderegger, M. (2023). *Regression modeling for linguistic data*. MIT Press.

Vallduví, E. (1992). *The informational component*. Garland.

Wilbur, R. (2012). Information structure. In C. Ceil, C. V. B. J. & K. M. C. (Eds.), *Sign Language: An International Handbook* (pp. 462-489). Mouton de Gruyter.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics, 27*(1), 55–105. https://doi.org/10.1006/jpho.1998.0085

Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. In *Proceedings of the TRASP Conference* (pp. 7-10).

Xu, Y., & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics, 33*(2), 159–197. https://doi.org/10.1016/j.wocn.2004.11.001

Yip, M. (2002). *Tone*. Cambridge University Press.

Zerbian, S. (2006). Expression of information structure in Northern Sotho. *ZAS Papers in Linguistics, 45*. Berlin: ZAS.

Zubizarreta, M. L. (2010). The syntax and prosody of focus: The Bantu-Italian connection. *IBERIA: An International Journal of Theoretical Linguistics, 2*(1). Retrieved from https://revistascientificas.us.es/index.php/iberia/article/view/89