BILINGUAL PUBLISHING GROUP

Forum for Linguistic Studies

https://journals.bilpubgroup.com/index.php/fls

ARTICLE

A Mathematical and Statistical Approach to Explore the Downtrend Properties in Chokri

Tulika Gogoi, Sekholu Tetseo, Amalesh Gope *

Department of Linguistics and Language Technology, Tezpur University, Tezpur 784028, Assam, India

ABSTRACT

This paper explores f0 downtrend processes in Chokri, a Tibeto-Burman language spoken in Nagaland, India. Chokri manifests five lexical tones, comprising four level tones- extra high (EH), high (H), mid (M), low (L), and one contour tone-mid-rising (MR). This study encompasses three distinct downtrend phenomena: temporal declination, final lowering, and automatic downstep. Three complementary methodologies- (a) visual, (b) statistical, and (c) mathematical modeling, are incorporated to examine time-dependent declination across all H, M, and L tone sequences. Visual and statistical analyses show the downtrend nature of these tonal sequences, revealing a linear downtrend in all H and L tone sequences, with M tone displaying a non-linear trajectory. However, while visual observation and statistical modeling do not precisely delineate the appropriate non-linear model, mathematical modeling confirms exponential decay as the fitting non-linear model for M tone sequences. The mathematical modeling further validates that only M tone sequences result in final lowering. Furthermore, an intervening L tone affects the f0 scaling of H tone sequences, confirming an automatic downstep in Chokri. Quantitative analysis revealed the pronounced effect of the intervening L tone on the preceding and following syllables containing an H tone and confirmed the declination process in Chokri.

Keywords: Chokri; Downtrends; Downstep; Declination; Statistical Modeling; Mathematical Modeling

*CORRESPONDING AUTHOR:

Amalesh Gope, Department of Linguistics and Language Technology, Tezpur University, Tezpur 784028, Assam, India; Email: amaleshtezu@gmail.com

ARTICLE INFO

Received: 1 August 2024 | Revised: 22 August 2024 | Accepted: 6 September 2024 | Published Online: 14 November 2024

DOI: https://doi.org/10.30564/fls.v6i5.6981

CITATION

Gogoi, T., Tetseo, S., Gope, A., 2024. A Mathematical and Statistical Approach to Explore the Downtrend Properties in Chokri. Forum for Linguistic Studies. 6(5): 664–677. DOI: https://doi.org/10.30564/fls.v6i5.6981

1. Introduction

Communication of post-lexical meanings of utterances through intonation is a universal phenomenon. Several crucial linguistic functions, such as the encoding of speech acts, information structure, and demarcation of prosodic units, are often carried out by intonation features. Paralinguistic meaning, which includes discoursal and attitudinal information, is also cued by intonation [1, 2]. The use of fundamental frequency or pitch for intonational meaning is well-researched in several non-tonal languages. Although fewer, intonational pitch variation in languages with lexical pitch contrast has also been explored in some tonal languages. In the extensive field of intonation research, a central theme captivating scholars is the exploitation of downtrend patterns, a phenomenon marked by the gradual descent of f0 from the initiation to the end of utterances. The presence of a downtrend in a language indicates that the pitch of a tone occurring later in the utterance will be perceptibly lower than the same tone at the beginning. Although debatable, such progressive pitch descent is often attributed to phonetic effects [3]. This intricate exploration unfolds into three distinctive dimensions: (i) declination, characterized by a time-dependent lowering of f0; (ii) downstep, the phenomenon of f0 lowering in high (H) tones influenced by the intervention of a low (L) tone; and (iii) final lowering [1].

In this context, zooming into the research on intonation-only languages, the predominant focus has historically gravitated toward the study of declination [4–7]. However, the exploration into downtrend patterns in numerous African tone languages has unearthed the existence of downsteps. This intriguing process involves alternating occurrences of H and L tones, where the subsequent H in an HLH sequence exhibits a lower f0 owing to the influence of the intervening L tone [3]. Downtrend analysis is an invaluable phenomenon, revealing the intricate interplay between the phonetics and phonology of pitch realization within a given language. Beyond its role in unraveling general intonational pitch patterns at the utterance level, downtrend analysis is pivotal in comprehending its dynamic interaction with other intonational properties, such as focus and phrase boundaries. In Mandarin, for instance, the declination rate is higher after narrow focus elements compared to plain utterances [8]. On the flip side, downstep encounters resistance through on-focus f0 raising and post-focus compression [9, 10]. In Tswana, downstep is blocked by phonological phrase boundaries, as evidenced by Ishihara [11], revealing its attenuation in the presence of a strong phrase boundary and a focused H tone following the L tone.

Declination is often deemed a universal effect in declarative utterances across languages and refers to the global downward trend of the f0 contour stretched over phrases or utterances [2, 4, 12–14]. Described as the "*gradual modification (throughout a phrase or utterance) of the phonetic backdrop against which the phonologically specified f0 targets are scaled*" [13, 15], declination has been extensively explored for its physiological underpinnings. Various researchers attribute its occurrence to falling subglottal pressure and laryngeal configurations [16–18]. Maeda's (1976) '*Tracheal Pull Hypothesis*' holds that the sternum is lowered when the lung volume decreases [4]. The larynx is physically connected to the sternum. It gradually goes down during the production of an utterance, resulting in a declining f0. Ohala and Ewan [19] reasoned that the laryngeal combinations that are required to produce a higher f0 are more complex than those needed to produce a fall. This means that speakers ought to favor making smaller rises instead of declines. Additionally, listeners would anticipate a f0 drop in utterances because of the f0 downward trending tendency. It would also become "a variable that speakers use with communicative intent" due to their usage of it as a perceptual cue for domain borders [20]. Xu [9] suggests that declination, rather than being a fundamental principle of intonation, is the total of various linguistic factors and local physiological constraints.

While earlier works on declination were conducted in non-tonal languages, viz., English [21] and Dutch [7], analyzing the actual features of declination is rather challenging. Declining pitch slopes in these languages often had to be estimated from sparsely situated f0 peaks or valleys. Additionally, as syllables were not specified for tone, observable pitch contours were influenced by numerous factors [8]. In this context, tone languages offer a more accurate understanding of declination with their specified tones for syllables. Although the declination effect can be seen for all tones, in tone languages, they are clearer through analysis of sentences containing words with the same lexical tones, like all H or all L sequences. Cantonese, for instance, is known to show declination in identical sequences of all six tones [22]. Declination is generally considered a universal effect, primarily present in declarative utterances. However, many researchers observed the absence of declination in some tonal languages, which has been one of the exciting findings from the works on intonational downtrends. For example, declination and final lowering are not found in Yoruba for sentences with all H and all M tones [15]. Mambila, on the other hand, shows little tendency for downtrends in Tone 1, the highest tone [3].

Another noteworthy downtrend in African tone languages involves alternate H and L tone sequences. H tones following an L tone manifest a lowered f0 compared to those preceding an L tone. This phenomenon, termed automatic downstep or down-drift [12], stands in contrast to non-automatic downstep, which involves the lowering of an H tone without a conditioning L tone physically present. In the latter case, a floating L tone triggers the downtrend. One of the characterizing features of a downtrend is that a downstepped H tone resets the pitch ceiling for the H tones that will subsequently occur. Moreover, downstep is cumulative as a successive series of downsteps results in a progressively lowering pitch track [3]. According to Rialland [23], downstep has three distinct properties: (i) it does not alter lexical tonal specifications, i.e., a high tone never changes to lower tones due to downstep, (ii) it extends its effect to the entire tonal sequence in its domain, and (iii) the phonetic realization of downstep varies from language to language, and it is defined differently by different authors. Languages like Igbo [24] and Bimoba [25] exhibit both downdrift and non-automatic downstep. However, some tone languages avoid downstep of either type to preserve the lexical tone specifications of syllables, as seen in Mambila, where no instances of downstep occur in sequences of its four tones [26, 27].

Final lowering, a more abrupt and excessive descent of f0 at the end of speech domains like phrases or utterances, has been recognized in various languages [1, 21, 28]. Research on intonation also reports that final lowering exists in languages like English [21], Japanese [14], Dutch [29], Kipare [30], and Yoruba [15, 31]. Welmers [32] notes the presence of final lowering in several discrete-level tone languages of Africa. Gussenhoven [1] opines that final lowering is grammaticalized in many languages where it encodes a post-lexical meaning, i.e., the finality of the utterance. While the earlier impressions of final lowering viewed it as a result of declination, it was later found to be an independent phonological device that may exist as a distinct downtrend feature or category [33].

Despite the abundance of tone languages in India, particularly in the northeastern region, not much work is available on intonation in these languages. Accounts of intonational phonology often include discussions on downtrend patterns, a less explored area in Indian tonal languages. Das [34] reports the presence of downstep, declination, and final lowering in Bodo, a Tibeto Burman tone language. The present study explores how different types of (f0) downtrends impact sentence prosody in Chokri, another lesser-studied Tibeto-Burman language boasting a richer tonal inventory of five lexical tones [35].

Chokri is predominantly spoken in the Phek district of Nagaland in North-East India. The language is spoken by the *Chokri-mi* or *Chokri* subgroup, a segment of the *Chakhesang* tribe. According to the 2011 Census Reports of India, the estimated number of Chokri speakers totals 111,062. Bielenberg and Nienu [36] documented 39 consonants, six vowels, and four tones. They collected data from two speakers residing in the Phek village in Nagaland, another variety of Chokri that is different from the variety reported in this study.

We recorded our data from the native speakers, primarily settled in the Thipuzu village. Our data reveals a segmental phoneme inventory comprising 33 consonants and 7 vowels. The language lends a heavy functional load on tonal phonemes, with every syllable featuring tonal specifications, making them the Tone Bearing Units (TBU) [37, 38]. In a recent study, Gope et al. [35] explored five-way tonal contrasts in Chokri- extra high [EH], high [H], mid [M], low [L], and mid-rising [MR]. Morpho-phonological analysis of tone in Chokri shows that the final syllable carries the contrastive tone in both nominal and verbal roots, while the non-final syllables are realized with a default mid (M) tone. Moreover, all the affixes in the language are specified with tones. This shows a dense syntagmatic distribution of lexical tones in the language [37]. **Figure 1** (reproduced from [35]) shows the 5-way tone contrasts in this language.

The quantitative classification of these contrastive tones is detailed using supervised machine learning algorithms in [35]. The present work is a follow-up research that seeks to address the following research questions:

i) Given Chokri's intricate nature of lexical tones (five-way tonal contrasts), how much are these tones affected by post-lexical downtrends of f0?
ii) What types of downtrends manifest in Chokri? How do

we quantitatively evaluate the presence of the downtrend effects in a language?

iii) What is the precise nature of downward f0 slopes resulting from downtrend effects?

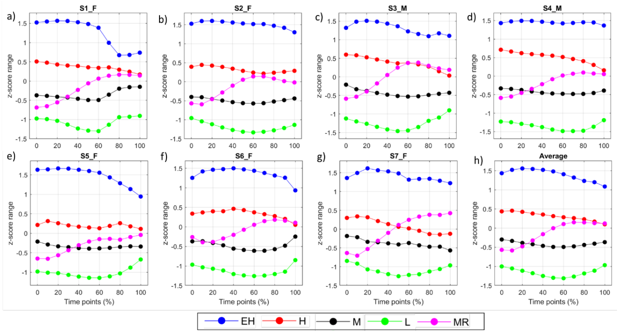

Figure 1. Z-score normalized f0 contours exhibited in a Chokri monosyllabic minimal set (5 way contrast) produced by seven native speakers are shown here- (**a**) S1_F, (**b**) S2_F, (**c**) S3_M, (**d**) S4_M, (**e**) S5_F, (**f**) S6_F, and (**g**) S7_F, and the (**h**) average (averaged across all the tokens by all the subjects). The legends, [S] stands for a subject, followed by the number of the subject and the gender information, where M stands for a male speaker, and F indicates a female speaker. The x-axis indicates the division of the f0 track into eleven equidistant time points, ranging from the onset (0%) to the offset (100%) of the f0 trajectory. This figure is reproduced from [35].

This paper introduces a novel exploration by integrating three complementary methodologies, including visual, statistical, and mathematical modeling, to examine the intrinsic patterns of f0 downtrends in Chokri. The key emphasis is on quantitatively assessing and evaluating the declination, automatic downstep, and final lowering of f0, offering a broader perspective applicable to several languages beyond Chokri.

Following this introduction, Section 2 presents the detailed methodology. Sections 3 and 4 describe the results and discussions. Finally, Section 5 makes the conclusive remarks.

2. Materials and Methods

2.1. Participants and Dataset

Speech data for this study were recorded from five native Chokri speakers, comprising three males and two females. All participants, aged between 18 and 55, were residents of Thipuzu village in the Phek district of Nagaland. The data recording task involved scripted sentences displayed on a computer screen, with participants instructed to produce them naturally. The entire dataset was repeated five times to ensure variability, incorporating significant intervals between repetitions. The dataset for this experiment was designed with the assistance of the second author (a native speaker) and cross-verified with two other native informants to ensure the authenticity of the language variety under consideration. The details of the dataset can be found in the supplementary section. The scripted sentences utilized in this study comprised four sets, each containing four distinct types of sentences. The dataset's structure employed for the experiment is outlined as follows:

1a. All high-tone (H) sequences;
1b. All low-tone (L) sequences;
1c. All mid-tone (M) sequences; and
2. H tone sequences with an intervening L– HLHH, HHLH, and HHLHH.

It must be noted that not all the tone sequences (or intervening tonal sequences) form (semantically) meaningful sentence constructions in Chokri. Therefore, this study concentrates only on the tonal sequences that form meaningful and natural sentences in Chokri while preparing the dataset.

2.2. Data Recording and Annotation

Recordings were digitized at a sampling frequency of 44.1 kHz and 32-bit resolution. For detailed analysis, individual sentence files were manually annotated and labeled at the syllable level using *Praat* [39]. Mean f0 and time-normalized f0 values were extracted using *ProsodyPro* [40]. The resulting pitch contours are graphically represented as line charts for visual inspection. This systematic grouping allows for a clear and structured analysis of the f0 patterns in the sentences.

2.3. Statistical and Mathematical Modeling

A one-way repeated measures analysis of variance (RM ANOVA) was performed to assess the statistical significance of observed differences in fundamental frequency (f0) using R (version 4.2.3) (R Core Team et al., 2013) [41]. The analysis utilized the function *aov(response(f0) ∼ factor(time) + Error(factor(syllable)), data = df*) in R. Both mathematical and statistical modeling are incorporated to investigate the phenomena of declination and downstep. Linear and exponential decay fitted models were utilized for sentences with all H, L, and M tone sequences. The mathematical equations were fitted using *Origin* software (version 8.1).

However, it was noted that the range of f0 varied between male and female speakers, primarily due to differences in vocal cord size across genders. To account for potential intra-speaker and inter-speaker anomalies or hidden covariates across all tokens produced by different subjects, a linear mixed-effects regression should be conducted in R using the *lme4* package[42]. The *lm(y ∼ x)* function, representing a basic linear model, was employed, where "y" represents the dependent variable (f0) and "x" signifies the independent variable (time points). Considering the potential influence of syllable count, the model was further adjusted as *lm(f0 ∼ time + syllable + gender).*The *lm()* model was preferred over a mixed-effects model *[lmer()]*since the primary focus of this study was to examine the overall effects of time, syllable count, and gender on f0 across all speakers. While mixed-effects models are valuable for accounting for random effects due to speaker-specific variability, this study specifically aimed to investigate the fixed effects of the independent variables, viz., time and syllable count, while considering gender as a fixed effect on f0. Additionally, our dataset did not demonstrate significant variability attributable to individual speaker differences that would require the inclusion of random effects.

3. Results

3.1. Visual Interpretation and Statistical Analysis

The results of the production experiment were designed to explore the declination trends in Chokri across all H, L, and M tone sequences, which indicate the presence of a phonetic declination process in the language for all three tone sequences. However, the rate of decline varies across the tone sequences. Time-normalized average fundamental frequency (f0) values, calculated for each male and female speaker, revealed that declination is most prominent in the initial syllable of the utterance. The first pitch drop occurs during the production of the first syllable, with the declination rate gradually decreasing as the utterance progresses. Despite the identical phonological specification of all syllables (high), the phonetic f0 value of late-occurring syllables is lower compared to those at the beginning.

Figures 2–4(i,ii) illustrate the time-normalized average f0 trends for male (i) and female (ii) speakers across all H, L, and M tone sequences. The standard error of the aggregated data is indicated with the shaded bars. A visual examination of the pitch tracks across the time-normalized f0 contours for different speakers and tonal sequences (H, M, and L) reveals that the f0 range for females is higher than that of males. The response of f0 as a function of time demonstrates a predominantly linear pattern for H and L tone sequences [see **Figures 2** and **3(i,ii)**]. On the contrary, a curvature is observed for the M tone sequence [see **Figure 4(i,ii)**].

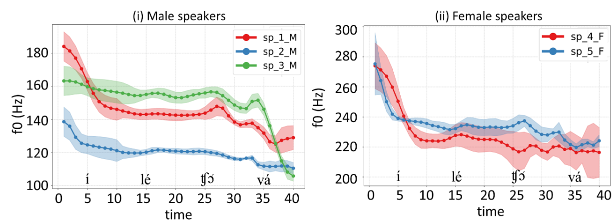

Figure 2. Time-normalized raw f0 contours of all H tone sequences, averaged across all the repetitions produced by each speaker in the utterance [í lé ʧɔ vá] "I am cooking"- (**i**) shows data from 3 male speakers and, (**ii**) shows data from 2 female speakers. The legends, [sp] stands for a subject, followed by the number of the subject and gender information, where M stands for male speaker, and F indicates a female speaker. The shaded error bars indicate the standard error of the aggregated data. Each 10 points in the X-axis represents an individual syllable.

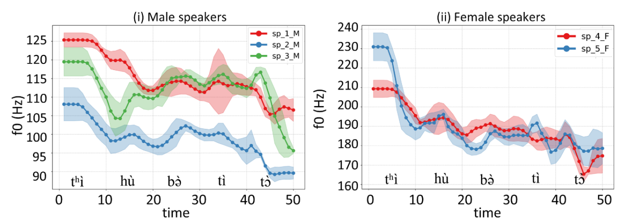

Figure 3. Time-normalized raw f0 contours of all L tone sequences, averaged across all the repetitions produced by each speaker in the utterance, [tʰì hù bə̀ tì tə] "(I) will boil some meat and eat"- (**i**) shows data from 3 male speakers and, (**ii**) shows data from 2 female speakers. The legends, [sp] stands for a subject, followed by the number of the subject and gender information, where M stands for male speaker, and F indicates a female speaker. The shaded error bars indicate the standard error of the aggregated data. Each 10 points in the X-axis represents an individual syllable.

The final lowering in the sentences with tonal sequences of all H and L is not observed across the speakers' repetitions considered in this study, as depicted in **Figures 2** and **3(i,ii)**. Notably, these exhibit a distinctive decreasing pattern in the f0 range towards the end of the normalized time for one male speaker [sp_3_M (in green color)] across both

H and L tone sequences. However, the qualitative analysis failed to accurately classify this trend as a definite "final lowering." It is essential to emphasize that this specific trend observed in a single male speaker's data is not entirely evident in the data of the female speakers or the two other male speakers. In contrast, the f0 trend observed in sentences with the M tone sequences displays a distinct decrease in f0 values towards the end of the normalized time, irrespective of gender. This observation hints at a final lowering trend in Chokri, as illustrated in **Figure 4(i,ii)**.

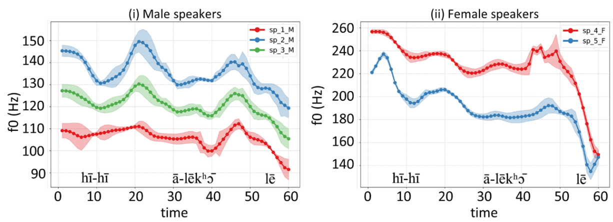

Figure 4. Time-normalized raw f0 contours of all M tone sequences, averaged across all the repetitions produced by each speaker in the utterance, [hīhī ā-lē kʰō lē] "this is my bag"- (**i**) shows data from 3 male speakers and, (**ii**) shows data from 2 female speakers. The legends, [sp] stands for speaker, followed by the number of subjects and gender information, where M stands for male speaker, and F indicates female speaker. The shaded error bars indicate the standard error of the aggregated data. Each 10 points in the X-axis represents an individual syllable.

To rigorously assess the significance of the declination patterns illustrated in **Figures 2–4**, encompassing all H, L, and M tone sequences, a one-way repeated measures analysis of variance (RM ANOVA) was conducted. The results confirm that f0 differences observed in each syllable are indeed statistically significant for the sequences of all H ($F[3, 12] = 66.7$, $p < 0.0001$*), all L ($F[3, 12] = 69.52$, $p < 0.0001$*), and all M ($F[5, 20] = 30.83$, $p < 0.0001$*) tones, respectively. Subsequent post hoc Bonferroni tests further affirm that the sequences of all H, M, and L tones are significantly different for each syllable. The f0 range of all H tones is higher than M and L. Similarly, the f0 range of all M tones is higher than L and lower than H tones. While this statistical analysis offers insights into group differences, it falls short in providing information on the intrinsic patterns or the nature of these tone sequences.

3.2. Statistical Modeling

The primary objective of our statistical modeling is to confirm our visual assessments and gain a deeper understanding of the inherent patterns and characteristics in the sequences of all H, M, and L tones. From a statistical modeling perspective, the choice between additive and interactive models emerges when examining these tone sequences. However, this study opts for an additive modeling approach to understand the impact of each predictor (gender, syllable, and time) on the response variable (f0 for various tone sequences of all H, M, and L tones). These predictors, being independent, facilitate deriving distinct estimations for the effect of each predictor. Given that the f0 range for female speakers is higher than their male counterparts, the primary focus lies in examining individual effects rather than interactive effects. Consequently, the predictor "gender" is intentionally excluded as an interaction term in the modeling adopted in this study.

The results of the generalized statistical modeling are presented in **Table 1**, with the response variable (f0 in Hz) based on various predictor variables (time, syllable, and gender) for different tone sequences (H, L, and M). A smaller p-value (<0.05) indicates statistical significance, and the coefficient estimates (β) provide insights into the strength and direction of the relationships. The intercept, although lacking meaningful interpretation in this context, represents the estimated value of the response variable(f0) when all other predictor variables (time, syllable (S_k), and gender) are set to zero. However, setting these to zero is not meaningful in this scenario as the f0 depends on time and gender. Also, the syllable (S_k) is assigned by grouping the normalized time so that $S_k \in [1 + 10(k - 1), 10k]$, for k = 1, 2, 3, 4, 5, and 6. Therefore, $S_1 = 1$ when the normalized time is from 1 to 10, $S_2 = 2$ when the normalized time is from 11 to 20, and so on. In contrast, the coefficient (β) for time represents the change in the response variable (f0) for a one-unit increase in the time variable while holding all other predictors constant. For H, L, and M tone sequences, the coefficients are −1.1408 Hz, −1.17328 Hz, and −1.4334 Hz, indicating a decrease per unit increase in time, respectively. The "gender" variable (Males or Females) demonstrates coefficients representing the difference in f0 between the two gender groups, with gender M as the reference group.

The coefficients for S_k with $k \in [2, 6]$ denote the change in f0 for the corresponding syllables compared to the reference syllable (S1). For example, in the H tone sequences, S_k with $k \in [2, 4]$ have a significant positive

Table 1. Statistical model output of f0 (Hz) for different tonal sequences - Response $\sim$ time + syllable (S_k) + gender.

Tone	Effects	Estimate	Std. Error	t Value	Pr(>|t|)
H	(Intercept)	240.4563	1.4465	166.233	<0.0001*
	Time	−1.1408	0.1663	−6.858	<0.0001*
	syllable (S_2)	4.7152	2.1431	2.200	0.02*
	syllable (S_3)	15.3410	3.5907	4.272	<0.0001*
	syllable (S_4)	15.3114	5.1698	2.962	0.003*
	gender M	−89.7632	0.9752	−92.043	<0.0001*
L	(Intercept)	204.53248	1.14108	179.245	<0.0001*
	Time	−1.17328	0.13495	−8.694	<0.0001*
	syllable (S_2)	−0.08319	1.69491	−0.049	0.96
	syllable (S_3)	9.60382	2.91739	3.292	0.0015*
	syllable (S_4)	20.15289	4.23875	4.754	<0.0001*
	syllable (S_5)	29.40886	5.63614	5.218	<0.0001*
	gender M	−80.13828	0.79119	−101.288	<0.0001*
M	(Intercept)	236.3181	3.7665	62.742	<0.0001*
	Time	−1.2861	0.3924	−3.277	<0.0001*
	syllable (S_2)	−0.4197	5.5358	−0.076	0.93
	syllable (S_3)	10.4171	8.7660	1.188	0.23
	syllable (S_4)	17.0594	12.4032	1.375	0.17
	syllable (S_5)	35.3654	16.1751	2.186	0.02*
	syllable (S_6)	27.8734	20.0057	1.393	0.16
	gender M	−89.6498	2.3008	−38.965	<0.0001*

effect of 4.7152 Hz, 15.34 Hz, and 15.31 Hz compared to the reference, respectively, with $p < 0.05$. A similar pattern is observed for the L tone sequences for S_k, with k [2, 5]. Nevertheless, significant syllable predictors are notably lacking in the M tone sequences, as evidenced by p-values exceeding the statistical significance threshold ($p > 0.05$) for most syllables in our analysis. Furthermore, for the H and L tone sequences ($F[5, 794] = 1744$, and $F[6, 693] = 1767$, with $p < 0.0001$), the adjusted R squared values are 0.92 and 0.93, respectively.

In contrast, for the M tone sequence ($F[7, 292] = 231.6$, with $p < 0.0001$), the adjusted R-squared value is 0.84. The adjusted R-squared value indicates how well the linear regression model fits the data, reflecting the proportion of variance in f0 explained by the predictor variables (time, syllable, and gender). For the H and L tone sequences, the statistical model accounts for 93% of the variance in f0. This suggests that the predictor variables (time, syllable, and gender) collectively exhibit a robust linear explanatory effect on f0. In contrast, the adjusted R-squared value for the M tone sequence is 0.84, signifying that the statistical model explains only 84% of the variance in f0. This is notably less than the explanatory power observed for the H and L tone sequences, indicating a relatively weaker linear association with f0 in the M tone sequences, with less than a 10% reduction in variance compared to H and L tone sequences. This indicates that the linear trends of the H and L tonal sequences are better than the M tonal sequence, suggesting a potential differentiating pattern in the latter.

3.3. Mathematical Modeling

In simpler terms, the statistical modeling adopted in this study indicates that linear models effectively depict the patterns in the H and L tone sequences. However, in the case of M tone sequences, a non-linear trend becomes apparent. Hence, mathematical modeling is adopted to understand and identify the specific nature of this non-linearity and devise a method to assess it. The advantage of employing mathematical models over statistical ones is their ability to provide a more in-depth understanding of non-linear patterns. It is important to note that selecting the best mathematical model or fitting equation depends on the nature of the curves, the theoretical background, and the research context of this study.

Figures 2–4(i,ii) exhibits the f0 trend across the normalized time for both male and female speakers. The syllable (S_k) is assigned to this time such that $S_k \in$ [1 + 10 (k−1), 10 k], for k = 1, 2, 3, 4, 5, and 6, denoting the Syllable no. **Fig-**

ure 5(i,ii) displays the average f0 for sentences containing H tone sequences as a function of S_k, where k ∈ [1, 4]. The length of these sentences varies in terms of the number of syllables they contain. The f0 decreasing trend is consistent across the male and the female speaker data. The f0 values fluctuate between 215–245 Hz for females and 110–170 Hz for males. Similar results are observed for the L tone sequences when the f0 is plotted as a function of S_k, with k ∈ [1, 5]. The f0 values vary within the range of 168–228 Hz for females and 90–126 Hz for males [see **Figure 6(i,ii)**].

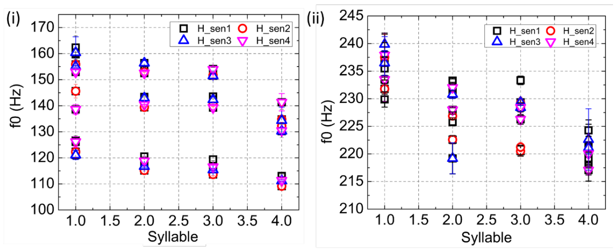

Figure 5. Time-normalized average f0 (in Hertz) for different sentences containing H tone sequences (H_sen1 to H_sen4) for male (**i**) and female (**ii**) speakers. These sentences vary in terms of number of syllables they contain. The x-axis represents the Syllable (S_k, where k = 1, 2, 3, and 4). Error bars depict standard errors aggregated over all the repetitions produced by all the speakers.

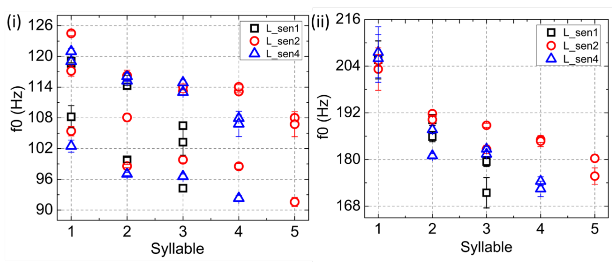

Figure 6. Time-normalized average f0 (in Hertz) for different sentences containing L tone sequences (L_sen1, L_sen2, and L_sen4) for male (**i**) and female (**ii**) speakers. These sentences vary in terms of number of syllables they contain. The x-axis represents the Syllable (S_k, where k = 1, 2, 3, 4, and 5). Error bars depict standard errors aggregated over all the repetitions produced by all the speakers.

The visual observation of the f0(S_k) depicted in **Figures 5** and **6(i,ii)** reveals two distinct fitting equations: linear and exponential decay. The linear equation includes f0(S_k) = $\beta_1 S_k + \beta_0$, where f0(S_k) represents the dependent variable f0 in the y-axis, S_k is the independent variable in the x-axis, and β_1 and β_0 are the coefficients akin to the slope and y-intercept, respectively. The exponential decay equation includes f0(S_k) = A exp($-S_k$/B) + A_0, where A, B, and A_0 are the coefficients. A is the vertical scale of the decay, representing the maximum value that f0(S_k) can reach; B is the decay constant influencing the rate at which the function occurs. f0(S_k) decreases with increasing S_k. The larger value of B leads to a slower decay. Finally, A_0 is the offset representing any constant vertical shift to the entire function. The negative sign in the equation represents the decreasing trend of f0(S_k).

3.3.1. Evaluation of Mathematical Fitting and Evidence of Chokri Downtrend

Figure 7(i–iv) shows the averaged f0(S_k) for all sentences featuring H, L, and M tone sequences uttered by all the subjects, denoting k ∈ [1, 4] for the H tone sequences, k ∈ [1, 5] for the L sequences, and k ∈ [1, 6] for the M sequences, respectively. The linear fit is feasible, whereas no convergence is achievable for the exponential decay fit in the case of the H tone sequences [see **Figure 7(i)**]. The blue horizontal line represents the highest f0. This limitation arises not only from the narrow f0 range, which spans from 160 to 180 Hz, i.e., ∼20 Hz, but also from the constraint on the syllable index (S_k), which is limited to 4. Conversely, both equations successfully fit the data for the M tone sequences. The f0 variation is similar for the L sequences, i.e., ∼22 Hz for k ∈ [1, 5] [see **Figure 7(ii)**], whereas f0 variation increases to ∼40 Hz with the k = 6 for M tone sequences [see **Figure 7(iii,iv)**].

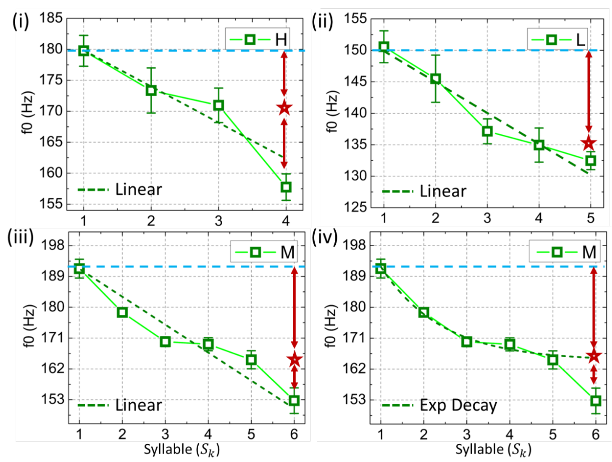

Figure 7. Time-normalized average f0 for all sentences featuring H, L, and M tone sequences produced by all the subjects is displayed in (**i–iv**) as a function of syllables. The corresponding green dashed lines represent the linear and the exponential decay fitted lines. The blue dashed line represents the highest f0. The red lines signify the vertical distance between the reference (calculated from the fit and denoted with a red star) and the actual final point.

This analysis examines the declination slopes and their nature within each tone sequence. To achieve this, each tone sequence is individually fitted in both linear and non-linear (exponentially decay) equations. Notably, the fitting is performed for f0(S_k − 1), excluding the last (final) point at (S_k) to compare it with the reference point later. The reference point, denoted by a red star in **Figure 7(i–iv)**, is calculated using the fitting parameters and the respective fitting equation. For instance, in the H tone sequences, the parameters for the linear fit are β_1= −4.44 ± 0.91 Hz and β_0= 183.83 ± 1.92 Hz, with an adjusted R-square value of 0.91, indicating a well-fitted curve. The reference point is calculated from f0(S_k= 4) = −(4.44 × 4) + 183.83 to assess whether the final point contributes to a lowering effect, resulting in approximately 166 Hz. The deviation is approximately 6 Hz from the actual final point ($\sim$160 Hz). This minimal difference suggests no pronounced final lowering effect in this context.

A similar analysis has been conducted for the other cases, including all L and M tone sequences shown in **Figure 7(ii--iv)**. For sentences with all L tone sequences, the linear fit yields parameters β_1 = −5.77 ± 1.23 Hz and β_0 = 155.60 ± 3.47 Hz, with an adjusted R-square of 0.88, suggesting a relatively weaker fit compared to the H tone sequences [see **Figure 7(ii)**]. These negative values further validate the existence of declination in Chokri, as indicated in the H and L tone sequences [see **Figure 7(i,ii)**].

The linear fit equations of f0 for sentences featuring all M tone sequences give rise to lower R-squared values ($\sim$0.465) [see **Figure 7(iii)**]. The R-squared values increase significantly when an exponential decay equation is employed. The improved R-squared value becomes 0.85. The non-linear fitting is also performed for different orders of polynomial equations f0(S_k) = $a_n S_k^n$+ $a_{n-1} S_k^{n-1}$+ ... +$a_2 S_k^2$+ $a_1 S_k$+ a_0, where a_n, a_{n-1}, ..., a_2, a_1, a_0 are the coefficients of the polynomial. S_k is the variable, and n is the polynomial degree, indicating the highest power of S_k in the equation. It is observed that the R-squared value increases from 0.90 to 0.95 for n = 2 and n = 3, respectively. However, this value is not completely reliable. In evaluating the exponential versus polynomial fit in this context, it's crucial to understand that the exponential fit quantities naturally decrease over S_k, aligning with our data's expected behavior. Additionally, B, the decay constant, affects the rate at which f0(S_k) decreases with increasing S_k. The larger value of B leads to a slower decay. For the individual M tone sequences (M_sen1 and M_sen2) and the averaged M tone sequences, B is 0.75 ± 0.45, 0.87 ± 0.22, and 0.80 ± 0.32, respectively. The R-squared values are 0.81 for M_sen1 and 0.92 for M_sen2, which becomes 0.85 for the average M tone sequences. On the other hand, the polynomial fit provides flexibility with the increased R-squared values; however, it is risky since it may incorporate overfitting with high-degree polynomials, and there is the unreliability of extrapolation beyond the S_k range, which is important in this study. An influence of earlier data points diminishing over S_k visually aligns with the expected behavior of the f0 in sentences featuring all M tone sequences [see **Figures 2–4(i,ii)**]. Despite having a slightly lower R-squared value of 0.85, the exponential decay model demonstrates a better conceptual fit for f0(S_k).

Extrapolating the exponential curve to the reference point reveals that the final point is relatively lower than the fitted point (approximately 15 Hz), implying a final lowering in the syllable sequences featuring all M tones in Chokri, as illustrated in **Figure 7(iii,iv)**. Therefore, the linear fit equation works for both H and L tone sequences, whereas the exponential decay fit equation is valid for the M tone sequence only. Since the k in S_k varies for each tonal sequence by one, it is hard to generalize its effects. Notably, the final lowering is only possible in the M tone sequences. The declination observed in Chokri does not depend on particular tonal sequences (H, M, and L).

3.3.2. The Effect of An Intervening L in All H Tone Sequences

This study further examines the effect of an intervening L tone in the H tone sequences to explore the potential occurrence of automatic downsteps in Chokri. This exploration is achieved through the sentences containing different combinations of syllables- HLH, HHLH, HLHH, HHLHH. Observing the time normalized f0 contours suggests that the intervening L can influence the preceding (H_p and H_2p) and the following syllable (H f and H 2f) containing an H tone. The legends, H_2p and H_2f, represent the location of a second preceding and following syllable featuring an H tone from the position of an intervening L tone (and not necessarily the immediate preceding syllable containing an H tone), respectively. In **Figure 8(i)**, the time-normalized average f0 is presented, averaging across all repetitions produced by all the subjects for the sentences containing the syllable sequence of HLHH, HHLH, and HHLHH, respectively.

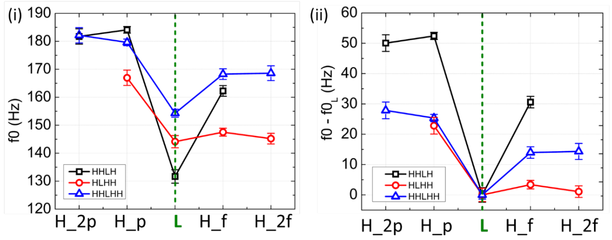

Figure 8. Time-normalized average f0 contours showing the effect of an intervening L between consecutive H tone in the syllable sequences featuring HHLH, HLHH, and HHLHH: (**i**) depict the time normalized f0 contours, and (**ii**) showcases the transformed f0 contours ($f0_L$). The transformed value is obtained by subtracting the f0 of the intervening L from each f0. The green dashed line is a reference point, distinguishing the preceding (H_p and H_2p) and following (H_f and H_2f) tone syllables.

The raw f0 values depicted in **Figure 8(i)** confirm that the intervening L tone predominantly affects the f0 of the immediately preceding and following H tone syllable (H_p and H_f) in the syllable sequence of HLHH, HHLH, and HHLHH, respectively. A reduction of f0 values in the immediate preceding H tone syllable of an intervening L tone syllable (depicted as H_p) is attributed to a possible declination effect. However, the second preceding H syllables- H_2p and H_2f, drawn in the sentences containing the syllable sequences of HHLH and HHLHH, are the least affected. Notably, a stiff fall in pitch occurs at the intervening L syllable, followed by a linear f0 contour. This phenomenon is distinctly observable in the HLHH and HHLH tonal sequences, confirming an automatic downstep in Chokri.

However, the raw f0 alone does not reveal the impact of an intervening L tone on these syllable sequences. The f0 is mathematically transformed to quantify this impact. This transformation ($f0_L$) is obtained by subtracting the f0 of the intervening L (∼130 Hz, ∼145 Hz, and ∼155 Hz) from all the f0 ranges for HHLH, HLHH, and HHLHH, respectively. The y-axis of **Figure 8(ii)**, marked as f0 - $f0_L$, depicts the calculated values and indicates the impact of the intervening L tone on the neighboring H tone syllables in each sentence. The hierarchy of this intervening L tone's impact on the following H tone syllable is observed to be in the following order– HHLH (the impact is approximately 30 Hz) > HHLHH (the impact is approximately 15 Hz) > HLHH (the impact is approximately 5 Hz). On the other hand, the hierarchical order of the intervening L tone's impact on the preceding H tone syllable is observed to be in the following order– HHLH (the impact is approximately 50 Hz) > HHLHH HLHH (the impact is approximately 30 Hz for each syllable type). This trend approves the presence of automatic downstep borne out due to the influence of an intervening L tone.

4. Discussion

This study effectively addresses all three research questions: (i) Given Chokri's intricate nature of lexical tones (five-way tonal contrasts), to what extent are these tones affected by post-lexical downtrends of f0?- The experimental results discussed above confirm that the post-lexical intonational effects like declination and automatic downstep are present in this language. However, the resulting gradual reduction of f0 is not present to the extent that would obliterate the lexical tones. (ii) What types of downtrends manifest in Chokri? How do we quantitatively evaluate the presence of the downtrend effects in a language?- The findings confirm that Chokri exhibits both declination and automatic downstep. Notably, the final lowering is restricted to only utterances featuring all M tone sequences. The integration of different mathematical fitting equations confirms this finding. Finally, (iii) What is the precise nature of downward f0 slopes resulting from downtrend effects?- The distribution of f0 contours affirmed that the declination manifests as an initial steep fall in f0, particularly observable in the first syllable, followed by a gradual slowing of the adjacent syllables. A linear model effectively fits the trend for the sentences with all H and L tone sequences. On the other hand, an exponential decay model establishes the declination nature of the sentences featuring all M tone sequences. The application of linear effects for H and L tone sequences and non-linear effects for the M tone sequence is substantiated through statistical and mathematical modeling. Consequently, this study elucidates how utterance-level pitch influences the realization of lexical pitch.

The outcome of this work seeks to fill in the need for quantitative modeling for enhanced verification of downtrend properties in natural languages. Devising apt mathematical models that best describe the global f0 slopes has complemented our initial postulations about the downtrend types in Chokri by establishing the precise nature of the f0 contours. Such an approach is also instrumental in bringing intonational phonology closer to speech technology, as our findings are expected to facilitate better systems for intonational aspects of speech synthesis and recognition.

Most works on downtrends primarily explore the phonetic and phonological properties of various downtrend types. They are based on the measurement of f0 on syllables at different positions across an utterance viz., Mambila [3], Igbo [24], Cantonese [22], and Tswana [43, 44], to name a few. However, drawing mathematical modeling to establish the nature and types of the downtrend is still relatively scarce. Liberman and Pierrehumbert [21] showed that downdrift in English is realized as a decaying exponential, wherein the f0 slope becomes progressively less steep towards the later part of the phrase. Myers [45] demonstrated that the downtrend of f0 in Chichewa could be evaluated using additive linear modeling and did not require exponential decay. However, the analysis is performed from a statistical modeling perspective, not the actual mathematical fitting. Furthermore, in other languages, such as Mandarin, it is exhibited that the declining f0 of sequences of Tone1 (H) is steepest at the start of the utterance and is modeled with an exponential decay [8]. However, this argument is primarily based on least square modeling, where the correlation is obtained in the utterances based on the correlation between the proceeding and following tone sequences.

The approach adopted in this study shares foundational concepts with those presented in [21]. The uniqueness of the current approach lies in systematically evaluating fitting equations for all H, L, and M tone sequences by comparing the linear and exponential decay modeling. Notably, a lower final point in the f0 vs. syllable or time plot does not necessarily provide evidence of the final lowering. Instead, this determination relies on factors such as total time, syllable count, and the nature of the fitted mathematical equation. The convergence of exponential decay may vary across data, and linear fits may not universally perform well. Judgment calls depend on domain knowledge and context.

Nonetheless, this study outlines a general methodology to uncover the intrinsic patterns of f0 trends in any language. The complexity of tone sequences and hidden covariates may pose challenges in identifying the appropriate mathematical equation. Furthermore, the quantitative analysis highlights the substantial influence of intervening L tones on both preceding and following H tone sequences. The impact of intervening L tone on the preceding and following H tone syllables is most significantly observed in the syllable sequence of HHLH. The study successfully bridges the two-fold objectives of testing the phonetic properties of downtrend patterns and developing mathematical models to establish these observations.

5. Conclusions

This study draws the characteristics of f0 downtrends in Chokri. The analysis of the results affirms the presence of temporal declination in utterances featuring all H, L, and M tone sequences. The f0 scaling at the onset of the utterance is higher than at the end, even when all constituent syllables are specified for the same lexical tone. The observed differences in f0 for each syllable are statistically significant, as derived through the one-way repeated measures ANOVA. Statistical modeling supports the visual inspection observed in the data and aids in formulating mathematical modeling. It confirms that H and L tone sequences follow a linear trend, whereas the M tone sequences follow a non-linear trend. The mathematical modeling evaluates and assesses the H, L, and M tone sequences with both linear and exponential decay equations. The adopted model quantitatively concludes that exponential decay fits the M tone sequences well. This model further confirms the presence of a final lowering in the sentences containing M tone sequences. The presence of an automatic downstep is depicted in the sentences containing HLH, HLHH, HHLH, and HHLHH tonal sequences by a consistently lowered f0 in the successive syllable containing an H tone. In addition to identifying the phonological downtrend categories, the quantitative analysis also reveals that the intervening L tone has a pronounced effect on the preceding and the following H tone sequence, with the most significant effect observed on the sentences containing the HHLH tonal sequences. In contrast, a relatively weaker effect surfaced in the second preceding and the following H sequence, irrespective of the number of syllables present in a given sentence. While this study offers valuable insights into the independent dynamics of downtrends, further research exploring their interaction with other intonational functions of f0, such as focus and phrasing, will contribute to a more comprehensive understanding of this phenomenon.

Author Contributions

Conceptualization: A.G. Dataset Preparation: S.T. and T.G., Experimental Design: T.G., S.T., and A.G., Data Recording: T.G., and S.T., Data annotation: T.G. Data curation: A.G., and T.G. Formal analysis: A.G., T.G., and S.T. Investigation: A.G. and T.G. Methodology: A.G., and T.G. Project administration and Supervision: A.G., Validation: T.G., S.T., and A.G., Visualization: T.G., and A.G.. Writing—original draft: T.G. and A.G. Writing (review and editing)—A.G., Final Draft—A.G., and T.G. Funding acquisition: A.G. All authors reviewed and approved the manuscript.

Funding

This research is partially funded by the National Language Translation Mission (NLTM): BHASHINI, sponsored by the Ministry of Electronics & Information Technology, Government of India, for the project "Speech Datasets and Models for Tibeto-Burman Languages" (SpeeD-TB), Project No: DoRD/LLT/AG/20-521/826-A.

Institutional Review Board Statement

Approval was obtained from the ethics committee of Tezpur University, comprising of the Member Secretary, Dean of Research and Development, and two internal experts, granted under order No. DORD/TUEC/10-14 (Vol-III)/1528-4. The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Informed Consent Statement

The informed consent was obtained from all participants for this study prior to the data collection. The participants were explained the recording procedure, and a demo was shown before the recording started. The participants signed an agreed consent form. They were also offered a token amount for participating in the recording experiment.

Data Availability Statement

The data that support the findings of this study are available with the corresponding author (Amalesh Gope, amaleshtezu@gmail.com), and those can be made available upon reasonable request. Representative sound files are available in the supplementary section.

Acknowledgments

The current paper is a revised and updated version of "The Phonetics of Downtrends in Chokri," presented and subsequently published in the Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS), Prague 2023 [38]. The content of this paper is a rework of that paper. The present paper has been updated and upgraded following the feedback and suggestions received from the blind reviewers of ICPhS 2023 and the general audience and experts of that conference. We sincerely acknowledge all the feedback we received from several experts attending the 20th ICPhS in Prague.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Gussenhoven, C., 2004. The Phonology of Tone and Intonation. Cambridge University Press: Cambridge.
[2] Ladd, R., 1984. Declination: A review and some hypotheses. Phonology Yearbook. 1, 53–74.
[3] Connell, B., 2001. Downdrift, Downstep, and Declination. Typology of African Prosodic Systems Workshop Bielefeld University: Germany.
[4] Maeda, S., 1976. A Characterization of American English Intonation [Ph.D. Thesis]. MIT: Cambridge.
[5] Pierrehumbert, J., 1979. The Perception of Fundamental Frequency Declination. Journal of the Acoustical Society of America. 66(2), 363–369.
[6] Cooper, W.E., Sorensen, J.M., 1981. Fundamental Frequency in Sentence Production. Springer: Berlin.
[7] Appels, R., 1985. Een productiemodel voor declinatie en onderzoek naar declinatie-resets binnen spraakuitingen. Internal Report no. 498. Institute for Perception Research: Eindhoven.
[8] Shih, C., 1998. Declination in Mandarin. Intonation: Theory, Models and Applications: Proceedings of an ESCA Workshop; Athens, Greece; 18–20 September 1998.
[9] Xu, Y., 1999. Effects of tone and focus on the formation and alignment of f0 contours. Journal of Phonetics. 27, 55–105. DOI: https://doi.org/10.1006/jpho.1999.0086

[10] Wang, B., Kügler, F, Genzel, S., 2022. The interaction of focus and phrasing with downstep and post-low-bouncing in Mandarin Chinese. Frontiers in Psychology. 13, 884102. DOI: https://doi.org/10.3389/fpsyg.2022.884102

[11] Ishihara, S., 2007. Major phrase, focus intonation, multiple spell-out (MaP, FI, MSO). Linguistic Review. 24, 137–167. DOI: https://doi.org/10.1515/TLR.2007.006

[12] Stewart, J.M., 1983. Downstep and Floating Low Tones in Adioukrou. Journal of African Languages and Linguistics. 5, 57–78.

[13] Cohen, A., Hart, J.T., 1967. On the anatomy of intonation. Lingua. 19, 177–192.

[14] Pierrehumbert, J., Beckman, M., 1988. Japanese Tone Structure. The MIT Press: Cambridge, MA.

[15] Connell, B., Ladd, R., 1990. Aspects of pitch realization in Yoruba. Phonology. 7, 1–30.

[16] Lieberman, P., 1967. Intonation, perception, and language. MIT Press: Cambridge, MA.

[17] Titze, I.R., 1988. The physics of small-amplitude oscillation of the vocal folds. The Journal of the Acoustical Society of America. 83, 1536–1552.

[18] Strik, H., Boves, L., 1995. Downtrend in f0 and Psb Journal of Phonetics. 23, 203–220.

[19] Ohala, J., Ewan, W.G., 1973. Speed of pitch change. Journal of the Acoustical Society of America. 53(1), 345. DOI: https://doi.org/10.1121/1.1912878

[20] Ohala, J., 1981. The listener as a source of sound change. In Papers from the Parasession on Language and Behavior. Chicago Linguistic Society: Chicago, IL. pp. 178–204.

[21] Liberman, M., Pierrehumbert, J., 1984. Intonational invariance under changes in pitch range and length. In: Aronoff, M., Oehrle, R. (Eds.). Language Sound Structure. MIT Press: Cambridge, MA. pp. 157–233.

[22] Johnson, K., 1986. Tone and intonation in Cantonese. Journal of the Acoustical Society of America. 79 (S1), S37. DOI: https://doi.org/10.1121/1.2023194

[23] Rialland, A., 1997. Le parcours du "Downstep" ou l'evolution d'une notion. Bulletin de la Société de Linguistique de Paris. 92, 207–243.

[24] Laniran, Y.d., 1990. Intonation in tone languages: The Yoruba example [Ph.D. thesis]. Cornell University: Ithaca, NY.

[25] Snider, K.L., 1998. Phonetic Realisation of Downstep in Bimoba. Phonology. 15(1), 77–101.

[26] Connell, B., 2006. Four tones and downtrend: a preliminary report on pitch realization in Mambila, a language with four level tones. In Proceedings of the 27th ACAL, University of Florida, Gainesville, 29–31 March 1996 and the LAGB, University of Sussex, Brighton, 11–13 April 1996.

[27] Connell, B., 2017. Tone and Intonation in Mambila. In: Downing, L.J., Rialland, A. (Eds.). Intonation in African Tone Languages. De Gruyter Mouton.

[28] Truckenbrodt, H., 2004. Final lowering in non-final position. Journal of Phonetics. 32, 313–348. DOI: https://doi.org/10.1016/j.wocn.2003.11.001

[29] Gussenhoven, C., Rietveld, A., 1988. Fundamental frequency declination in Dutch: Testing three hypotheses. Journal of Phonetics. 16, 355–369.

[30] Herman, R., 1996. Final lowering in Kipare. Phonology. 13(2), 171–196. DOI: https://doi.org/10.1017/S0952675700002098

[31] Laniran, Y., 1992. Phonetic aspects of tone realisation in Igbo. Progress Reports from Oxford Phonetics. 5, 35–51.

[32] Welmers, W., 1973. African Language Structures. University of California Press: Berkeley.

[33] Arvaniti, A., 2004. Final lowering: Fact, artifact or dialectal variation? The Journal of the Acoustical Society of America. 116(4), 2644. DOI: https://doi.org/10.1121/1.4785546

[34] Das, K., 2017. Tone and Intonation in Bodo [Ph.D. Dissertation]. Indian Institue of Technology: Guwahati.

[35] Gope, A., Pal, A., Tetseo, S., et al., 2024. Multi-class identification of tonal contrasts in Chokri using supervised machine learning algorithms. Humanities and Social Sciences Communications. 11(1), 1–13. DOI: https://doi.org/10.1057/s41599-024-03113-2

[36] Bielenberg, B., Nienu, Z., 2001. Chokri (Phek Dialect): Phonetics and Phonology. Linguistics of the Tibeto Burman Area. 24(2), 85–122.

[37] Tetseo, S., Gope, A., 2021. Tonal Contrast and Grammatical Tone in Chokri. In Proceedings of the 43rd International Conference of Linguistic Society of India. CIIL-Mysore: Mysuru, KA.

[38] Gogoi, T., Tetseo, S., Gope, A., 2023. The Phonetics of Downtrends in Chokri. In Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS); Prague, Czech Republic; 7–11 August 2023.

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

[40] Xu, Y., 2013. ProsodyPro—A tool for large-scale systematic prosody analysis. In Proceedings of the TRASP Conference, Aix-en-Provence, France. pp. 7–10.

[41] R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available from: https://www.R-project.org/

[42] Bates, D., Maechler, M., Bolker, B., et al., 2009. lme4: Linear mixed-effects models using Eigen and S4. Available from: http://lme4.r-forge.r-project.org

[43] Zerbian, S., Kügler, F., 2015. Downstep in Tswana (Southern Bantu). In Proceedings of the ICPhS, Glasgow.

[44] Zerbian, S., Kügler, F., 2021. Sequences of high tones across word boundaries in Tswana. Journal of the International Phonetic Association. 53(2), 361–382. DOI: https://doi.org/10.1017/S0025100321000141

[45] Myers, S., Selkirk, E., Fainleib, Y., 2018. Phonetic implementation of high-tone spans in Luganda. Laboratory Phonology. 9(1), 19. DOI: https://doi.org/10.5334/labphon.101