An Autosegmental-Metrical Analysis of Rising Contours in Standard Albanian

Enkeleida Kapia[1, 2], *Felicitas Kleber*[2], *Jonathan Harrington*[2] [1]

[1]Academy of Albanological Sciences, Institute of Linguistics
[2]Ludwig Maximilian University, Institute of Speech Processing

{ekapia|kleber|jmh}@ phonetik.uni-muenchen.de

Abstract

This study explored basic intonational units for Albanian, an understudied and typologically rare language of the Indo-European branch, through the lens and assumptions of the autosegmental-metrical framework and ToBI conventions. The emphasis was on broad focus sentences with differently stress patterned target words (e.g. trochaic, iambic, etc), in various sentential positions. General observations from recordings of 21 Standard Albanian speakers and a pilot study with acoustic measurements from a subset of 11 of them show that prominence at the word level is marked post-lexically by a low tone pitch accent which docks at the stressed syllable and is followed by what we propose to be an independent high tone at the end of the word; prominence at the phrase level also suggests post-lexical marking via an Accentual Phrase. The findings lend support to the idea that Albanian is a pitch accent language, which marks prominence at the head and at the edge of the phrase, joining thus, the group of a very small number of languages that do so.

Index Terms: Intonational phonology, Autosegmental-metrical framework, ToBI, Accentual phrase, rising, Albanian

1. Introduction

This study is concerned with identifying the basic intonational units of Albanian, a less studied Indo-European language spoken mainly in Albania, in Southeastern Europe. The proposed model adopts elements and assumptions of the auto-segmental metrical framework, and utilizes conventions for transcribing phonological categories within the intonation of a language in the ToBI system (for Tones and Break Indices) [e.g. 1, 2]. ToBI specifies conventions for labelling tonal patterns (Tones) associated with prominences (pitch accents) and phrase boundaries (edge tones), as well as labels for indicating the perceived level of disjuncture between words (the Break Indices), all of which are applied in tiers. It is important to note here that ToBI tone labels do not incorporate a direct record of acoustic measures, but rather categorizations of prosodic aspects of the signal generalized across speakers and contexts. Tone events in ToBI are considered to be targets of either an H (High) and L (Low), or a combination of these two tones. The ToBI conventions state that a full ToBI transcription should include 4 tiers of labels: tones, words, breaks and miscellaneous. However, only the Tone labels will be discussed in this paper.

The AM framework of intonational phonology has enabled the creation of a model of prosodic typology based on the idea that the prosodic properties of an utterance are a combination of prosody at the word level and at the phrase level [e.g. 3, 4].

Prominence marking at the word level categorizes languages by the type of lexical (whether a language has lexical pitch accent, stress, tone, some combination of these, or none of these) or post-lexical prosody (whether prominence is cued by the head of phrase, edge of phrase, both or none). At the phrase level, prominence marking categorizes languages by the type of prosodic units they have at the lexical (morae, syllables, feet, etc.) and post-lexical level (accentual phrase, intermediate phrase, intonational phrase). One of the aims of this study was also to locate Albanian within this established prosodic typology model.

1.1. Previous work on Albanian prosody

Up to 2015 any mention of Albanian prosody consisted of a couple of pages in introductory phonetics books, with pencil drawings of the intonational contours of declarative and interrogative utterances, derived from introspective ear inspections [5, 6, 7]. An important aspect for this study is the fact that Albanian has lexical stress and allows both trochaic and iambic types [8]. In addition, recent research with preliminary experimental evidence from production and perception, has suggested that Albanian marks up prominence acoustically via pitch accents [9, 10]. However, the latter two studies included all types of segmental material (i.e. not just sonorants usually used in prosodic research); in addition, given that the focus of the studies was prosodic realization of information structure correlates, experimental materials only addressed issues of different focus types (see also [11] for an analysis of the prosody of polarity questions in Albanian). By improving on the shortcomings of the previous studies, the aims of the present study are twofold: 1) to discover the basic intonational units of Albanian by providing not just a descriptive analysis, but also empirical evidence through a pilot study and 2) to understand how Albanian prosody is typologically related to (or different from) the prosody of other languages. This study is part of a larger project with focus on addressing issues of the prosody-pragmatics-syntax interface in adult and child grammar.

2. The Corpus

2.1 Speech materials and experimental design

The analysis presented here was developed on the basis of two corpora of spoken Albanian collected for the purpose of this study. 3300 utterances spoken by 21 different speakers were inspected both auditorily and acoustically. The corpus consists of read speech materials collected under laboratory conditions and designed to specifically address issues of stress type, location of target word, sentence type and focus type. Both neutral renditions, as well as renditions in which a particular

---

[1]All authors are equal contributors.

constituent was focalized, were obtained. Broad focused conditions were elicited as out-of-the-blue utterances, while narrow/contrastive/corrective focus conditions were read in response to an elicited context that triggered the focus. All sentences to be read were judged by four independent listeners to be naturally sounding. The language variety looked at here is Standard Albanian (hereafter, SA) as spoken in Tirana.

2.2 Speakers, experimental procedure, and preprocessing

We recorded 20 Standard Albanian (SA) speakers (11 f, 9 m) with mean age 41.14 (sd = 7.4). All speakers were born, educated, and lived all of their lives in Albania; of these, 4 also spoke Geg, the northern Albanian dialect. Subjects were asked to read aloud sentences presented on a laptop computer screen, at their usual speech rate and with natural accentuation. Importantly, they were not instructed which intonation/speech melody they should use. Recordings took place in a quiet room at the Institute of Linguistics in Tirana and were made using the *SpeechRecorder Software* [12] and a *Beyerdynamic Custom One Pro Headset*.

All recordings were labelled semi-automatically with the Munich Automatic Segmentation System [13] using the language independent model and stored as an *EMU speech database* [14]. By means of the Emu-WebApp, word and segment boundaries are currently manually checked and hand-corrected where necessary in all sentences (see 3. below for further details).

2.3 First observations

In the initial analysis, the data were inspected visually and auditorily in the Emu-WebApp token by token by the three authors. These first observations suggested that the smallest prosodic unit at the word level is a pitch accent realized as a low tone target L on the stressed syllable. This L is always followed by a rise that extends over the entire word, ending with a high target on or at the end of the final syllable of the word. Specifically, if stress falls on the ultimate syllable, then the tonal pattern begins with a low target on the ultima and ends with a high target within the ultima. If, on the other hand, stress falls on the penultimate syllable, then the tonal pattern begins with a low target on the penultima, rising from there to reach a high target on or at the final syllable. Figure 1 below shows exemplary pitch tracks of the target words *bari* with an iambic stress pattern i.e. word-final lexical stress and *bari* with a trochaic stress pattern i.e. word-initial lexical stress. Note how the rise begins much earlier for the word with trochaic stress.
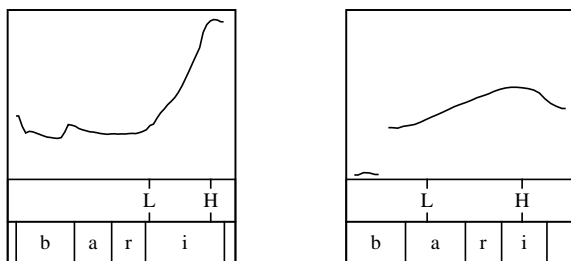

Figure 1: : 'bari' with iambic stress   'bari' with trochaic stress

We propose to model this as two types of tonal events for Albanian: the pitch accent L* which associates with the stressed syllable followed by an H tone which appears to associate with the boundary of the word or phrase. In order to determine whether the sequence L H consist of two independent tones L and H, with H being an independent anchor from L (L* H) or whether the sequence L H is a bitonal pitch accent over the entire word (L* + H; cf. [15]), we investigated in more detail the alignment of the L and H tones in both iambic and trochaic words.

3.   Pilot study on tonal alignment

3.1. Hypotheses

The pilot study investigates the following hypotheses:

(1) The L tone occurs later in words with an iambic stress pattern compared to trochees if the low tone docks onto the primary stressed syllable.

Based on our first observations we further predicted that H is independent of L:

(2) The alignment of the H tone should not differ as a function of a word's stress pattern.
(3) The slope between the L and H tone is flatter in trochees compared to iambic words.

Following [16] we also varied the position of the target word within the sentence so that they occurred in initial or medial position. However, we expected no effect of such a fixed factor on the alignment of H targets and the steepness of the rise because of our hypothesis that H is a boundary tone that aligns with the word boundary. But even if H were a trailing tone its alignment should not be affected by sentence position; only if H were a higher level boundary tone that aligns with the end of a phrase (be it intermediary or intonational), then we should see flatter slopes for words in sentence initial position due to a then much delayed H peak.

3.2. Methods

We selected all repetitions of two homophones, the iambic /ba'ri/ (*grass*) and the trochaic /'bari/ (*shepherd*) words in sentence-initial and sentence-medial position from all 11 female speakers from the database and labelled these tokens on a separate tone tier in the Emu-WebApp based on visual inspection of the f0 trajectory and according to the following criteria (cf. Fig. 2): The low tone of the lexically stressed syllable was labeled – in accordance with general ToBI conventions – as a starred tone, hence L*. This label was placed at the f0 minimum of the macroprosodically relevant f0 valley and marked the beginning of the rise. The high target of the rising contour was labeled H and placed at f0 maximum or the beginning of a high f0 plateau to mark the offset of the rise. Any misplaced word boundary (resulting from the automatic segmentation, cf. 2.2 above) was hand-corrected so that the word onset boundary always coincided with the beginning of the labial stop's closure phase and the word offset with the end of the stable F2 in /i/.
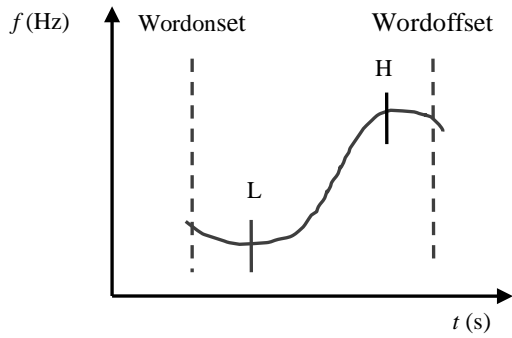
Figure 2: Schematic illustration of tonal target labeling and segmentation.

Since the dataset consisted of only obtained 96 observations and one random factor (speaker) we analyzed in RStudio [17, 18] by means of three separate ANOVAs (using the ezANOVA package, [19]) the effects of the two fixed factors *stress pattern* (trochaic vs. iambic) and *sentence position* (initial vs. medial) on three different dependent variables which are further specified below in (1) to (3):

$$prop.\ Delay\ (L) = \frac{t(L) - t(Word_{onset})}{t(Word_{offset}) - t(Word_{onset})} \qquad (1)$$

$$prop.\ Delay\ (H) = \frac{t(H) - t(Word_{onset})}{t(Word_{offset}) - t(Word_{onset})} \qquad (2)$$

$$Slope\ (LH) = \frac{12 \times log_2\left(\frac{f(H)}{f(L)}\right)}{t(H) - t(L)} \qquad (3)$$

For six out of these 96 tokens the values for *Slope (LH)* were infinite; these tokens were excluded from all three analyses. Additionally, the entire data of four speakers, for whom one pair of observation (e.g. iambic words in initial position) was missing, had to be excluded leaving us with an analysis of seven speakers × four observations (averaged across repetitions) per speaker = 28 data points.
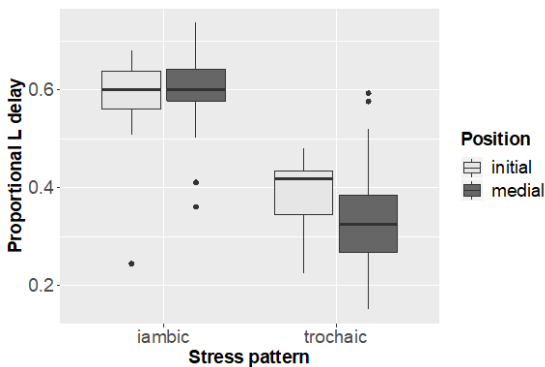
3.3. Results

Figure 3: Prop. Delay (L) in iambic and trochaic words in initial and medial position.

Fig. 3 quite clearly shows that L* occurred later in iambic /baˈri/ than in trochaic /ˈbari/ regardless of its position in the sentence. Commensurate with Fig. 3, the ANOVA with *prop. Delay (L)* as the dependent variable revealed a significant effect for stress pattern ($F[1,6] = 84.3$, $p < 0.001$), but none for sentence position and no significant interaction effect between the two factors.
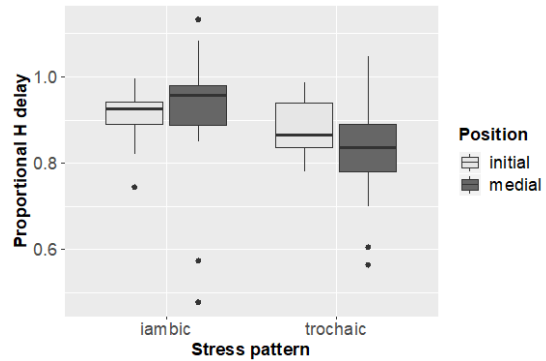

Figure 4: Prop. Delay (H) in iambic and trochaic words in initial and medial position.

The alignment of the following high tonal target H was affected to a lesser extent by lexical stress. As shown in Fig. 4, the H was located near the end of the word, particularly in words with an iambic stress pattern. The H tone was reached slightly earlier in trochees, especially when they occurred in sentence-medial position. These descriptive observations are mostly in line with the results from the ANOVA: the stress pattern affected the dependent variable *prop. Delay (H)* significantly ($F[1,6] = 8.8$, $p < 0.05$), but sentence position again turned out to be non-significant, both in the interaction with stress and as the main factor.
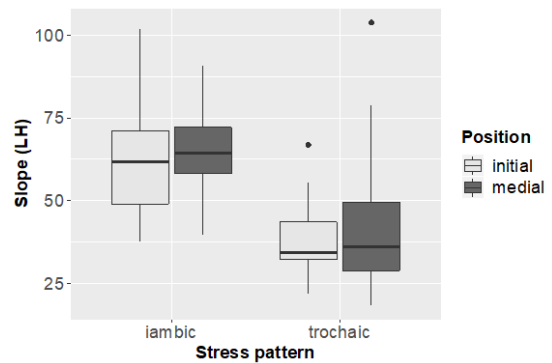

Figure 5: Slope of the rise from L to H in iambic and trochaic words in initial and medial positions

Fig. 5 shows that the rise from L to H was much steeper in iambic than in trochaic words both in initial and medial sentence position. Commensurate with this observation, the ANOVA with *Slope (LH)* as the dependent variable showed again a significant effect of the stress pattern ($F[1,6] = 82.9$, $p < 0.001$). To our surprise, the main factor sentence position turned out significant, too ($F[1,6] = 6.1$, $p = 0.05$), presumably reflecting the small tendency towards steeper rises when a word occurred later in the sentence, i.e. in sentence-medial position,

regardless of its stress pattern (there was no significant interaction effect). This result needs further attention in future analyses that will also include words in final position.

4. Discussion & Conclusions

The results of this study show that Albanian is indeed a language that marks prominence acoustically via pitch accents, as suggested in [9,10]. The main pitch accent described here is L* which docks onto a stressed syllable regardless of its position within the word. Our pilot study of the words 'bari' with stress on the ultimate syllable and 'bari' with stress on the penultimate syllable shows that the position of the L* occurs much later in the iambic words than in trochaic ones, as predicted in Hypothesis (1). Additionally, this analysis shows the existence of a boundary tone H falling on or at the end of the final syllable of the word. Although there was an effect of lexical stress on the alignment of the high tonal target H, this effect was much less pronounced than it was for the L tone; this will be discussed in detail in the following paragraph. Last, but not least, the rise from L* to H is, in line with Hypothesis (3), much steeper in iambic than in trochaic words, whether the words are in initial or in medial sentence positions.

The observed differences in H alignment as a function of the lexical stress pattern was not predicted by Hypothesis (2); however, this is not so surprising given previous evidence for greater variability of H tones compared to L tones in prenuclear rises (e.g. Greek [20], German [21], Dutch [22], English [23], Spanish [24]). Importantly, here we see that the effect of lexical stress on H alignment is not as pronounced as for L alignment and that the slope was much flatter in trochees compared to iambs. We take these combined results to suggest that H is a boundary tone. If H was a trailing tone, i.e. an anchor that depended on L*, we should have seen a similarly strong effect of stress pattern on H alignment to the one that we observed for the L alignment, and, as a result of that, equally steep slopes for both stress patterns [15]. Moreover, the increased variability of H alignment in trochees is likely linked to some variation in the segmental make-up stemming from inflection due to case: /ˈbari/ became /ˈbarin/ in sentence-medial position. Such predictable shifts in alignment, e.g. due to composition of words and phrase, are, however, irrelevant for the intonational phonology [25] of a language.

But what kind of a boundary tone is then the high tonal target H at the boundary of each unit? Given that the effect of sentence position on slope was so small, in our opinion, an interpretation of H as a higher level boundary (e.g. intonational phrase) tone is not justified. We propose that, since the domain of this tone seems to be the lexical word, this H tone marks the end of an Accentual Phrase (AP), as suggested for other languages, such as Georgian [26], French [27], Korean [28], Japanese [29], Bengali [30], and Farsi [31], etc. Accentual phrases in Albanian, similarly to the ones in the aforementioned languages, seem to be the smallest tonally demarcated phrases in the language. In general, our data suggests that the Albanian AP may include one word (a content word), but it may also include more than one word. In the interest of space, we provide below in Fig. 6 one example of an AP with more two words. The stressed syllable of the AP is the prominent one, marked by a low tonal target L, while the final syllable of the AP is marked by a high tonal target H.
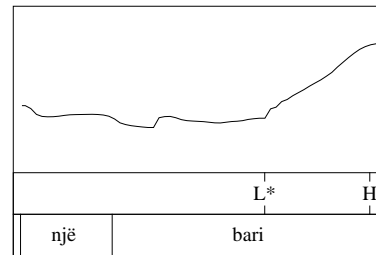

Figure 6: AP with two words: 'një bari' (*a shepherd*)

Based on the AM model of intonational phonology, wherein prominence and phrasing are the two major properties marked by intonation, we also propose here that Albanian is both a head-prominence and an edge prominence language. In Albanian, as observed for head/edge-prominence languages (such as French [32], Kiche [33], Bengali [34], Georgian [25], etc.), prominence at the phrase level seems both to be marked by the head and the edge of the phrase; the pitch accent L, marking the head, and the boundary tone H, marking the edge of the phrase, are opposite tonal types, which, in the case of Albanian, realize the rising tonal pattern over a word or an entire AP. This being said, the notion of an Accentual Phrase in Albanian needs to be further checked against empirical evidence considering the amplitude and/or the duration of the phrase's prominent syllable.

5. Future Directions

While the results of this study have brought us one step further into the investigation of Albanian intonation by establishing the smallest prosodic units for Albanian at the word and phrasal level, a natural question that arises now is to what extent can these results be generalized to speech spoken in natural settings. In order to answer this, we are already preparing our methods and procedures for a naturalistic corpus, as well as for a perception experiment that will look further into the tonal units uncovered here, in both given and new speech contexts. Importantly, additional empirical analyses of both duration and alignment measures of these proposed tonal categories are necessary in varying focus and more complex sentence types and structures.

6. Acknowledgements

7. References

[1] Pierrehumbert, J.B. *The phonology and phonetics of English Intonation*. Ph.D. thesis, MIT, 1980.

[2] Beckmann, M. E., Hirschberg, J., Shattuck-Hufnagel, S. "The original ToBI system and the evolution of the ToBI framework ", in S.-A. Jun [Ed], *Prosodic typology*, Oxford, University Press, 2005, pp. 9-54.

[3] Jun, S.-A. *Prosodic typology: the phonology of intonation and phrasing*. Oxford, University Press, 2005.

[4] Jun, S.-A. *Prosodic typology II: the phonology of intonation and phrasing*. Vol. 2. Oxford University Press, 2014.

[5] Boriçi, L. "Intonacioni dhe funksionet e tij në gjuhën shqipe". *Studime Filologjike*, 2, 1987, pp. 74– 87.

[6] Memushaj, R. *Fonetika e shqipes standarde*. Tiranë: Botimet Toena, 2009.

[7] Beci, B. *Fonetika e gjuhës shqipe*. EDFA. Tiranë, 2004.

[8] Topalli, K. *Theksi në gjuhën shqipe*. Instituti i Gjuhësisë dhe i Letërsisë. Akademia e Shkencave, Tiranë, 1995.

[9] Kapia, E., Brugos, A. "Information structure-prosody interface: Towards a Model of Albanian Intonational Phonology". *Speech Prosody 2016*, 2016, pp. 946-950.

[10] Kapia. E., Brugos, A. "Perception of prosodic correlates of Topic and Focus in Albanian". ICPhS Satellite Workshop: *Intonational Phonology of Typologically Rare or Understudied Languages.* University of Melbourne, Australia, 2019.

[11] Themistocleus, C., Mueller, D. "The Intonational of Albanian Polar Questions and Statements". In Botinis, A. (ed.), *Proceedings of the 6th ISEL Conference of Experimental Linguistics*, Athens, Greece, 2015.

[12] Draxler, C. & Jänsch, K. SpeechRecorder – A universal platform independent multichannel audio recording software. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, 2004, pp. 559–562.

[13] Kisler, T., Reichel, U., Schiel, F. Multilingual processing of speech via web services. *Computer Speech & Language* 45, 2017, pp. 326–347.

[14] Winkelmann, R., Harrington, J., Jänsch, K. EMU-SDMS: Advanced speech database management and analysis in R. *Computer Speech & Language* 45, 2017, pp. 392– 410.

[15] Grice, M., Arvaniti, A., & Ladd, D. R. "On the Place of Phrase Accents in Intonational Phonology", *Phonology* 17(2), 2000, pp. 143–85.

[16] Jun, S.-A. & Fletcher, J. Methodology of studying intonation: From data collection to data analysis. In *Prosodic Typology II*, ed. by S.-A. Jun, 2014, pp. 493-519.

[17] Bates, D., Chambers, J., Dalgaard, P., Gentleman, R., Hornik, K., Iacus, S., Ihaka, R., Leisch, F., Lumley, T., Maechler, M., Murdoch, D., Murrell, P., Plummer, M., Ripley, B., Temple Lang, D., Tierney, L., Urbanek, S. *The R project for statistical computing*. URL: http://www.rproject.org/.

[18] RStudio Team. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

[19] Bakeman, R. "Recommended effect size statistics for repeated measures designs". *Behavior Research Methods*, 37 (3), 2005, pp. 379-384.

[20] Arvaniti, A., Ladd, D.R.; Mennen, I. "Stability of tonal alignment: The case of Greek prenuclear accents". Journal of Phonetics 26, 1998, pp. 3-25.

[21] Atterer, M., Ladd, D.R. "On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German". Journal of Phonetics, 32, 2004, pp. 177-197.

[22] Ladd, D.R., Mennen, I., Schepman, A. "Phonological conditioning of peak alignment in rising pitch accents in Dutch". Journal of the Acoustical Society of America 107(5), 2000, pp. 2685-2696.

[23] Ladd, D.R., Faulkner, D., Faulkner, H., Schepman, A. "Constant "segmental anchoring" of F0 movements under changes in speech rate". Journal of the Acoustical Society of America 106(3), 1999, pp. 1543-1554.

[24] Prieto, P., van Santen, J., Hirschberg, J. "Tonal alignment patterns in Spanish". *Journal of Phonetics* 23, 1995, pp. 429-451.

[25] Ladd, R. Intonational Phonology. Cambridge University Press, 2008.

[26] Vicenik, C., Jun, S.A. "An autosegmental-metrical analysis of Georgian intonation". In *Prosodic Typology II*: The Phonology of Intonation and Phrasing, Jun. S.A. ed. Oxford, OUP, 2014, pp. 154-186.

[27] Jun, S.A., Fougeron, C. "A Phonological model of French intonation". In *Intonation: Analysis, Modeling and Technology*. Botinis, A. (ed.), Dordrecht: Kluwer Academic Publishers, 2000, pp. 209-242.

[28] Jun. S. A. "The Accentual Phrase in the Korean prosodic hierarchy". *Phonology* 15 (2), 1998, pp.189–226.

[29] Beckman, M., Pierrehumbert, J. "Intonational structure in Japanese and English". *Phonology Yearbook* 3, 1986, pp. 255–309

[30] Khan, S. D. *Intonational Phonology and Focus Prosody of Bengali*, Ph.D. dissertation, University of California, Los Angeles, 2008.

[31] Sadat-Tehrani, N. S. "The Structure of Persian Intonation", *Proceedings of Speech Prosody*, Campinas, Brazil, May 6–9, 2008.

[32] Jun, S.-A. & Fougeron, C. "Realizations of Accentual Phrase in French Intonation", *Probus* 14, 2002, pp. 147–72.

[33] Nielsen, K. "Kiche Intonation", *UCLA Working Papers in Phonetics* 104, 45–60, Los Angeles: University of California, 2005.

[34] Hayes, B. & Lahiri, A. "Bengali Intonational Phonology", *Natural Language and Linguistic Theory* 9, 1991, pp. 47–96.