9th International Conference on Speech Prosody 2018
13-16 June 2018, Poznań, Poland

A Preliminary Model of Tatar Intonational Phonology

Adam J. Royer, Sun-Ah Jun

University of California Los Angeles, USA

ajroyer@ucla.edu, jun@humnet.ucla.edu

Abstract

This study is a preliminary report of ongoing research investigating an Autosegmental-Metrical model of the intonational phonology of Kazan Tatar, a Turkic language spoken in Tatarstan, Russia. Tonal patterns of neutral focus utterances were examined by varying the length of words and phrases, the location of stresses, syntactic structures, and sentence types. Results suggest that Tatar has two prosodic units marked by intonation. They are the Intermediate Phrase (ip), and the Intonational Phrase (IP). The stressed syllable of a prominent word is marked with a post-lexical pitch accent, L+H*, which can be realized as H* or L* due to prosodic or tonal contexts. Interestingly, an optional high tone (Hi) can be realized on the initial syllable of a word which does not carry a pitch accent, and this tone can be the only tone in an ip. An ip is marked by a phrase-final boundary tone, H- or L-, realized on an ip-final syllable, which is accompanied by a small degree of final lengthening. Finally, an IP is marked by a phrase-final boundary tone, H% or L%, realized on a substantially lengthened IP-final syllable. This intonation model of Tatar is compared with the intonation model of Turkish and the status of a "head"-less ip is discussed.

Index Terms: intonation, Turkic languages, Tatar, autosegmental-metrical, Initial H tone, head-less prosodic unit.

1. Introduction

Tatar is a Turkic language spoken in Tatarstan, Russia. There are approximately 5 million speakers including L2 speakers. Like many other Turkic languages, Tatar is an agglutinative language with a rich morphology. Some work has been done on the segmental phonology of this language [1], but little work has been done on its suprasegmental properties. Comrie [1] notes that like other Turkic languages, stress in Tatar typically occurs on the last syllable of a morphological word, with some exceptions (e.g., loan words from Russian, question words). No studies, however, have attempted to model the intonation system of Tatar.

The goal of the current study is to build a preliminary model of intonational phonology of Tatar in the AM (Autosegmental-metrical) framework [2, 3, 4]. In particular, we investigated how prominence and prosodic units are marked by intonation in Tatar, and what is the inventory of tones marking prominence (i.e., 'head' tones) and the boundaries of each prosodic unit (i.e., 'edge' tones).

1.1. Data Collection and Analysis Procedure

Data were elicited from our Tatar language consultant, a native female speaker of Kazan, Tatarstan, in her mid 30s. As a child, she spoke Tatar in the home and was educated in Tatar and Russian. She came to the United States after high school and has been living in the Los Angeles area.

In order to find out whether intonational tones in Tatar are marking the stressed syllable of a word or edges of a word, multiple types of declarative sentences were designed by varying in the length of words and the location of stress in a word and the location of a word in a sentence, as well as the type of nouns (e.g., pronouns, proper nouns, common nouns, loan words). Furthermore, to investigate if Tatar has a prosodic unit larger than a word, and if so, whether the edges of the unit are marked by a boundary tone, the structure and the length of the sentence were varied. For example, the Subject or the Object nouns included a single noun, a compound noun, a noun modified by one or more adjectives, a prepositional phrase, and a relative clause, and some sentences included a subordinate or coordinate clause.

The acceptability of the sentences was checked by our consultant, and she was asked to produce the sentences as if she was talking to her friends and family members. Recordings were made in a quiet room with using a Zoom H1 Handy Portable Digital Recorder and a head-mounted microphone. The sampling rate was set to 44.1 kHz and the quantization rate to 24 bit. The pitch contours of utterances (so far 177 sentences) were analyzed by using *Praat* which displays pitch tracks of each sentence together with spectrograms and waveform and four tiers (words, English gloss (for each Tatar word), tones, English gloss of the utterance). On the words tier, the stressed syllable of a word was explicitly marked (with '**"**' before the syllable) only when stress is not on the last syllable of the word (e.g., for loanwords with non-word-final stress).

2. Proposal of Tatar Intonation Model

As an overview of the model, the proposed prosodic hierarchy and its tonal affiliations can be seen in Figure1. As shown in the tree diagram, Tatar has two levels of prosodic unit that are marked by intonation: the Intermediate Phrase (ip) and the Intonational Phrase (IP). The head of a prominent word can be marked with a L+H* post-lexical pitch accent, but H* and L* are also found in a certain prosodic context as a variant of L+H*.

Additionally, there is another optional tone, the "initial H (Hi)", which may appear on the initial syllable of a word (though rarely the 2nd syllable is possible when a word is longer than two syllables). This tone does not align with the stressed syllable of the word it docks to, which means that it cannot be considered a pitch accent. The Hi-tone-marked word can be found right after the last pitch accented word in the same ip or as the first word in a multi-word ip. In both cases, the ip is the last ip of an IP and the IP is marked by L% and no pitch accent is found after Hi.

10.21437/SpeechProsody.2018-156

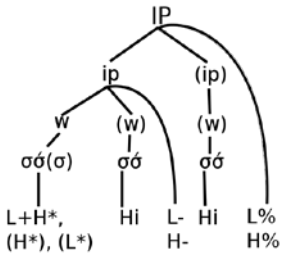

Figure 1*: Proposed prosodic hierarchy for Tatar*

2.1. Pitch Accents and Hi tone

Typically, the stressed syllable of each content word is marked with a L+H* pitch accent, which is realized with f0 peak on the stressed syllable, preceded by f0 valley on the immediately preceding syllable (see the first word in Fig.2). The f0 peak in L+H* is often aligned with the end of the stressed syllable, but can be earlier in the syllable when phrase-medial (see the first word 'my' in Fig.7). However, when the word is monosyllabic, the low tone is not always realized (see the second word 'song' in Fig.2). H* is labeled for such cases to mark the surface f0 shape. Sometimes, H* is also observed on an utterance-initial word even though the word is longer than two syllables. In that case, the low target was on the initial syllable of the word (see the first word in Fig.3). Since the distribution of H* is often predictable and variable, we propose H* is a surface variant of L+H*.

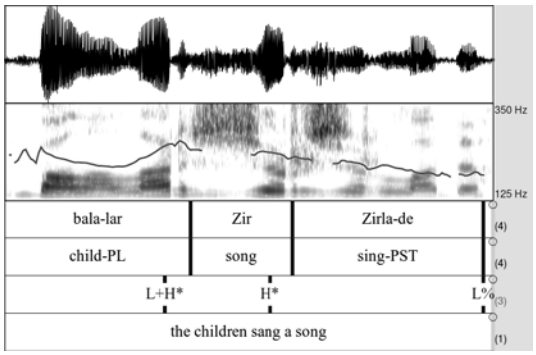

Figure 2: Sample pitch track *showing the realization of the L+H* and H* in the sentence "The children sang a song".*

As shown in Fig.2, the low target of L+H* on the first word 'children' is realized on the second syllable, i.e., immediately before the stressed, final syllable of the word, showing that the L tone is not aligned to the left edge of the word as in Turkish [5, 6]. Furthermore, loanwords with exceptional stress show that the high tone does not mark the right edge of a word, but rather mark the stressed syllable of the word. The f0 contour on the loanword, [matraska] 'mattress', in Fig.3 illustrates this point. The f0 peak is on the second, stressed syllable [tra], with low f0 before and after the stressed syllable.

In addition to H*, as can be seen in Figure 4, L* can occur on the sentence-final verb when the pitch accented syllable also carries a L% IP boundary tone. In general, a verb rarely gets pitch accented, but when it does, the pitch remains low while displaying increased amplitude. Since L* occurs on the syllable carrying L%, thus predictable, we propose L* is also a surface variant of L+H*.

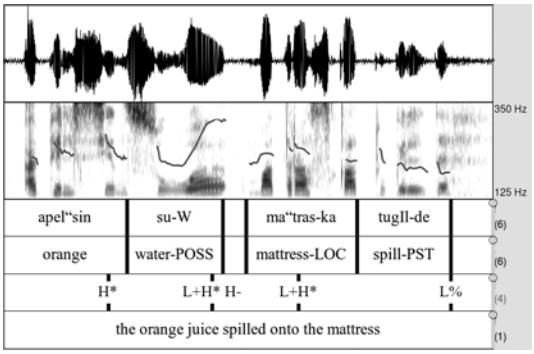

Figure 3: *An example* pitch track showing the Russian *loan "mattress" with penultimate stress.*

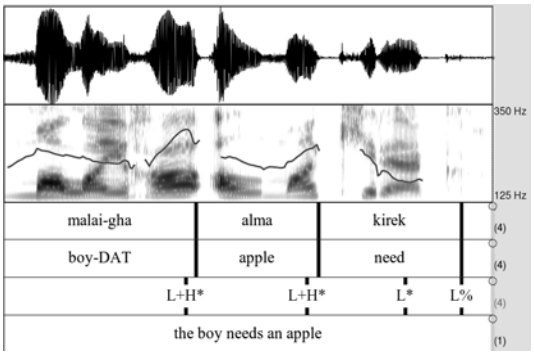

Figure 4: *Sample pitch track showing the realization of a L**

When a word is a compound, consisting of more than one lexical item, the first component of the compound receives a pitch accent in addition to the last (stressed) syllable of the word. This is shown for the second prosodic word, "to the parents", in Fig. 5. In Tatar, the word "parents" is a compound made up of the words for "dad" [eti] and "mom" [eni]. It is clear that the first f0 rise (L+H*) occurs over the first two syllables of the word "dad" and the second f0 rise aligns with the last two syllables of the prosodic word, "to the parents".

Function words tend not to get pitch accented. However, in careful speech, it is possible for them to be accented.

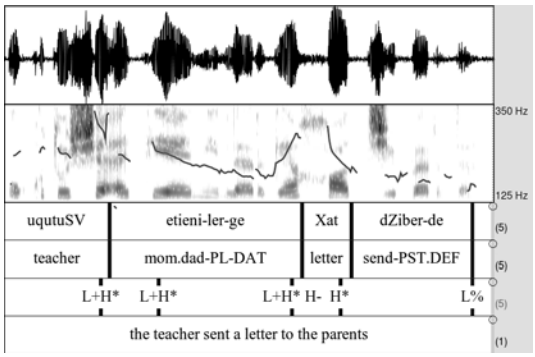

Figure 5: *Sample pitch track showing the first element of a compound word "parent", [eti], receives a pitch accent.*

In addition to pitch accents, Tatar seems to mark some kind of prominence with a high f0 and stronger amplitude on the initial, unstressed, syllable of a word when the word is the last word or second to the last word of an ip/IP, and is preceded by a pitch accented word. We have labeled this tone "Hi" to refer to the "word-initial" preference of the tone (cf. French in [7]). Typically, the Hi tone is realized on the initial syllable of the sentence/IP-final verb or a sentence/IP-final object (before a verb since the default word order in Tatar is SOV), but it can be on any IP-final word. Regardless of the syntactic category of the word(s) or the location of Hi, however, the IP is always marked by L%. That is, the f0 is falling after the Hi-toned syllable until the end of the phrase. Fig. 6 shows an example where Hi occurs on the initial (unstressed) syllable [me] of the Object, "the cat", followed by no pitch accent on the verb.

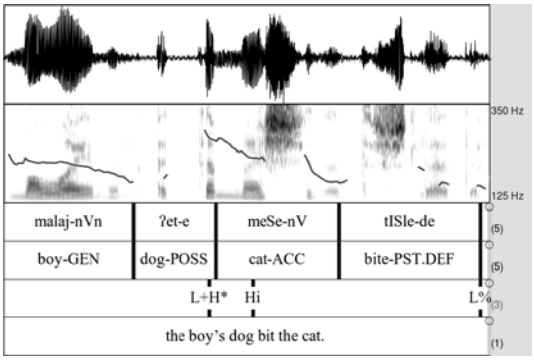

Figure 6: *Sample pitch track showing the realization of Hi on the initial syllable of the object, the cat.*

Furthermore, the Hi tone can occur as the first tone of an ip (intermediate phrase). This happens when the ip is the last ip of an IP and has only one or two words. Since the Hi is not followed by a pitch accent until the end of an ip/IP, this means an ip can have no pitch accent in Tatar. See Fig.7.

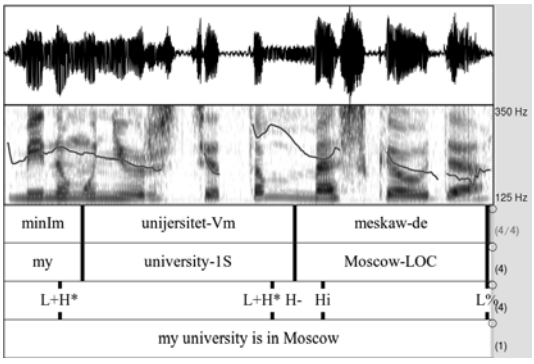

Figure 7: *Sample pitch track showing the Hi on the initial syllable of the last word "in Moscow" as the only tone in the last intermediate phrase.*

Figure 7 shows that Hi occurs on the initial syllable of the last word, "in Moscow", but this word starts a new ip because the preceding word ("university-1S") marks the end of an ip (marked by an H- boundary tone, which is higher than the preceding H target; see the next section for more detail).

The occurrence of Hi is optional. Of the 177 declarative utterances examined in this study, Hi tone was found about 30% of the time, and the distribution of the two types of Hi (i.e., occurring after a preceding pitch accent or alone in an ip) was about equal, and it is not yet clear what conditions the presence or absence of the Hi.

2.2. Intermediate phrase (ip)

The intermediate phrase can have more than one prosodic word and its right edge is marked by either a high (H-) or a low (L-) boundary tone, realized on its last syllable. The H- boundary tone is more common than L-. The ip-final syllable tends to show a small amount of lengthening. There are some cases, however, where the only discernible cue to the right edge of an ip is the f0 rise or fall, with no lengthening at all. Typically, the f0 height of H- is higher than that of the preceding H of pitch accent, and the f0 of L- is lower than that of the preceding L of pitch accent. But if the syllable carrying L- is also stressed, it is realized with a falling f0 (because L- is appended after the L+H* pitch accent).

For the ip boundary in Fig.7, the last syllable of the word "university-1S" is slightly lengthened and its f0 is much higher than the H on the preceding pitch accent, whose f0 peak is realized on the 2nd syllable of the 1st word, "my". We take this higher H- tone to be the result of an H- being realized on the same syllable that carries the H of the pitch accent.

When there are more than two pitch accents in an ip, a sequence of L+H* pitch accents shows downtrend of f0 peaks, following the f0 declination slope. So, when the f0 peak of L+H* is higher than that of the preceding H, we interpret the H tone as the H- boundary tone of an ip. An example is shown in Fig.8, which includes a complex NP object (*the photo of the monkey boy*). Here, the first word (*the mother*) forms its own ip, cued by final lengthening and super-high f0. Starting from the accent on "monkey", the accent's peak on "boy-GEN" is lower than that of "monkey", but the peak of the accent on "picture" is clearly higher than that of "boy-GEN", violating the downtrend. This suggests that the H tone at the end of the word "picture" marks a boundary of a larger prosodic unit, which corresponds to the complex NP object. This larger prosodic unit is an ip, and its right edge is marked by H-. An ip often marks the end of a large syntactic constituent (e.g. a relative clause, a subordinate clause, a heavy NP, PP). Fig. 9 shows an example of L- ip boundary, marking the end of a relative clause, "that scolds me". Here, the f0 on the last syllable of the word "come.out" shows a falling f0, with the L tone lower than the L of the preceding word.

When L+H* and H- do not appear on the same syllable (i.e., when the ip-final syllable is not stressed), f0 from the H* syllable stays high, maintaining a high plateau till the H- syllable, or shows a small rise to the end of the H- syllable. This happens often when the unstressed syllable is a function word. An example is illustrated in Fig.10. Here, the end of Prepositional Phrase "with Guzel" is marked by H- boundary tone (and final lengthening), but the last pitch accent of this ip is realized on the second syllable of "Guzel", and the high f0 from [zel] keeps rising till the end of the function word which is unaccented. [Note that the second ip in this utterance begins with an Hi tone on the initial syllable of the object noun, "chess", followed by no pitch accent till the L% syllable.]

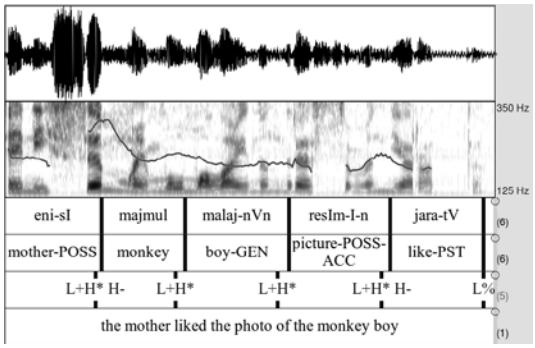

Figure 8: *Sample pitch track of H- occurring after the subject of the sentence, "mother", and also after the complex object NP "photo of the monkey boy".*

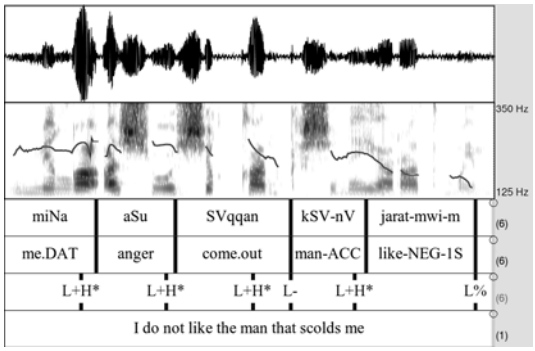

Figure 9: *Sample pitch track showing L-, marking the end of a relative clause, before the head noun "man-ACC".*

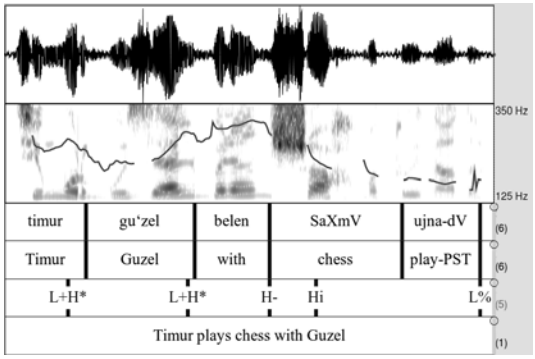

Figure 10: *Sample pitch track showing a small f0 rise over the function word "with", marking the H- intermediate phrase boundary tone.*

2.3. Intonational phrase (IP)

An IP can have one or more ip's and is marked by a boundary tone on its final syllable, which is substantially lengthened, and an optional pause. The IP final syllable is notably longer than that of an ip final syllable. So far, two types of an IP boundary tones are observed: The L% tone at the ends of declaratives and the H% in the middle of an utterance, typically signaling a continuation rise.

Like the ip boundary tone, an IP-final syllable with H% shows an f0 peak higher than the preceding high target, above the declination line. Fig.11 shows both the H%, after the conjunction "because", and the L% at the end of the utterance.

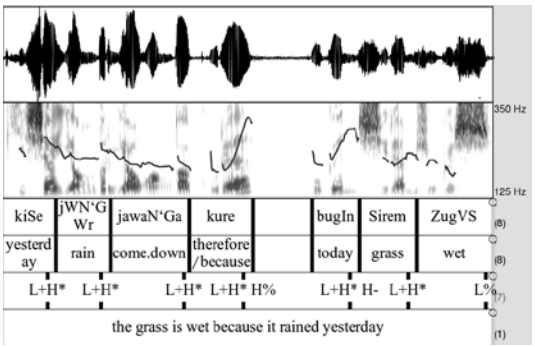

Figure 11: *Sample pitch track of L% and H% realization.*

3. Discussion and Conclusions

As a Turkic language, Tatar is similar to Turkish in that stress is typically realized word-finally ([5, 6, 8, 9, 10, 11]) and is marked by pitch accent ([5, 6, 12]). Both languages have two prosodic units above a word, i.e., IP & ip/PhP ([5, 6, 12]), and the end of a syntactic constituent is marked by an ip boundary tone, though more regularly in Turkish than in Tatar. A prosodic word typically shows a rising f0 contour in both languages, though the Low tone is aligned with the pretonic syllable in Tatar while with the left edge of a word in Turkish.

However, these two languages' intonation systems differ in many ways. The most obvious one is that, in Turkish, the nuclear pitch accent plays an important role in syntax and is consistently marked by a High tone on the left edge of the nuclear accented word while the word itself is produced in a reduced pitch range [5, 6]. In Tatar, however, there does not seem to be any evidence for nuclear pitch accent marking. Further work on focus prosody in Tatar might shed some light on the nuclear pitch accent.

Another difference that sets Tatar apart from Turkish is the presence of the Hi. In Tatar, the Hi tone seems to mark prominence of a word even though the head of word is available. When the word with the Hi is the last word of an ip, it may make sense to reason that the prominence of the word is shifted to its initial syllable because its final syllable is carrying an ip/IP boundary tone. However, this interpretation would not work when the word is the first word of a multi-word ip. Furthermore, an ip without any pitch accent is theoretically challenging because, in the AM model, languages where prominence is marked by pitch accents associated with a stressed syllable ("head"), a prosodic unit larger than a word is expected to have a prosodic head. For example, in English, an ip should have at least one pitch accent [2, 3, 4, 13]. Though now we have seen a few languages where a stressed syllable does not contribute to the formation of intonation contour by marking postlexical prominence through a pitch accent ([14, 15]), in Tatar this seems to happen only to the words in the IP-final ip (or the last word in an ip). This is therefore also typologically interesting. Finally, it is also possible that the Hi on the ip-initial word may not mark the prominence of the word, but the prominence of the phrase. More work is needed to figure out the nature of Hi and the factors contributing to the appearance of the Hi. More data should be examined to confirm the current findings and to include focus prosody and prosody of various sentence types.

4. References

[1] B. Comrie, Tatar (Volga Tatar, Kazan Tatar) phonology. *Phonologies of Asia and Africa*, 2, pp. 899-911. Chicago, 1997.

[2] J. Pierrehumbert, *The phonology and phonetics of English intonation*, Ph.D. dissertation, MIT, 1980.

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

[4] D. R. Ladd, *Intonational Phonology*, Cambridge: Cambridge University Press, 1996/2008.

[5] C. Ipek, *The Phonology and Phonetics of Turkish Intonation*. PhD dissertation, Univ. of Sothern California. 2015.

[6] C. Ipek, and S.-A. Jun, "Towards a Model of Intonational Phonology of Turkish: Neutral Intonation", in the Proc of Meeting on Acoustics (POMA), vol. 9, pp. 060230-069238. 2013.

[7] S.-A. Jun and C. Fougeron, The realization of the Accentual Phrase in French intonation. *Probus* 14: 147-172, 2002.

[8] E. Kaisse, Some theoretical consequences of stress rules in Turkish. *CLS* 21:1. 199–209, 1985

[9] S. Inkelas & C. O. Orgun, Level (non)ordering in recursive morphology: evidence from Turkish. In Steven Lapointe, Diane Brentari & Patrick Farrell (eds.) *Morphology and its relation to phonology and syntax*, Stanford: CSLI, pp. 360–392., 1998.

[10] S. Inkelas, Exceptional stress-attracting suffixes in Turkish: representations versus the grammar, In Rene Kager, Harry van der Hulst & Wim Zonneveld (eds.) *The prosody–morphology interface*. Cambridge: Cambridge University Press, pp. 134–187, 1999.

[11] B. Kabak & I. Vogel, The phonological word and stress assignment in Turkish. *Phonology* 18: 315-360, 2001.

[12] S. Kan, *Prosodic domains and the syntax-prosody mapping Turkish*. Unpublished MA thesis. Boğaziçi University. 2009.

[13] S.-A. Jun, Prosodic typology: by prominence type, word prosody, and macro-rhythm. S.-A. Jun (ed.) *Prosodic Typology II: The phonology of intonation and phrasing*, Oxford University Press, pp. 520-539, 2014.

[14] E. Lindstrom and B. Remijsen, "Aspects of the prosody of Kuot, a language where intonation ignores stress," *Linguistics*, vol. 43(4), pp. 839–870, 2005.

[15] C. W. Kisseberth and M. I. Abasheikh, "Chimwiini phonological phrasing revisited," *Lingua*, vol. 121, pp. 1987–2013, 2011.