Boundary tones indicate turn allocation in Russian

Ludger Paschen

This paper argues that boundary tones (BTs) are an integral building block in the organisation of turn-taking in Russian. Contrary to previous accounts, it is shown that BTs are not redundant with respect to pitch accents. Instead, new data suggests a close correlation between BTs and turn allocation (choice of next speaker) that is stronger for final BTs than for initial BTs. This has direct implications for the status of BTs in autosegmental models of Russian intonation (in particular ToRI) and offers new insights into the debate on whether or not intonational units are functionally decomposable, or just how constructional they really are.

1. Introduction

It is well-established that prosody plays a crucial role in delivering a wide range of communicative meanings (including, but not limited to, functions related to turn-taking) in naturally occurring conversation (e.g. Barth-Weingarten and Szczepek Reed 2014). Both intonational events at prosodically prominent positions and at the edges of prosodic units – represented as *pitch accents* (PAs) and *boundary tones* (BTs) in autosegmental-metrical approaches – have been argued to be relevant in this context, as both contribute to the overall prosodic shape of an utterance. In the case of Russian, research on intonation has been mainly concerned with pitch accents, while boundary tones have received only little attention; this observation holds for several independent veins of research on Russian intonation (Bryzgunova 1980; Yokoyama 2001; Odé 2008b). However, a number of recent studies suggest that speakers of Russian do make systematic use of tonal distinctions at final and initial positions within phrases (IPs) in order to express various communicative meanings such as epistemic modality or (in)completeness (Krause 2007; Rathcke 2009).

Whether boundary tones are relevant for turn-taking in Russian and, if so, what exactly their functional contribution is, however, remains an issue hitherto not addressed in either the phonological or the conversation-analytical literature. For that reason, this paper seeks to make one further step to resolve this question. On the basis of machine-learning classifications carried out on annotated recordings of spontaneous Russian speech, the interaction of PAs, BTs and

Proceedings of ConSOLE XXIII, 2015, 493-508
http://www.sole.leidenuniv.nl

turn-taking categories are discussed. The main findings are a strong involvement of final BTs (FBTs) in turn allocation (speaker choice), a weaker involvement of initial BTs (IBTs), and a predictability of FBTs after one PA (H*M + M%). No evidence of BTs being involved in signalling turn length could be found.

This paper is organised as follows: In section 2, I introduce ToRI, a model of Russian intonation upon which the study of boundary tones will be based. In section 3, I discuss some previous research on the role of prosody in turn-taking and motivate three hypotheses that will be tested later on. In section 4, I outline the data basis, the annotation procedure and the classification algorithms used in the study. In section 5, I present the results of the study, and in section 6, I consider some theoretical implications for the mapping of prosodic form and communicative function in general.

2. *Approaches to Russian Intonation*

The term *intonation* comprises non-lexical phenomena related to the pitch (F0) contour in (larger) prosodic phrases. Intonation has to be distinguished from lexical or morphological tone (Yip 2002), and also from segmentally-driven microintonation (Cohen and ’ t Hart 1967; Hombert 1978; Bradshaw 1999). Intonation is generally assumed to be a universal feature of language, and the constantly growing number of languages for which descriptions of intonational systems or sub-systems are available (Jun 2005, 2014) is feeding that assumption. Intonation has been argued to convey a wide spectrum of functions, including syntactic disambiguation (Odé 2008a), marking of information-structural categories (Krifka and Musan 2014), and expression of subjective-emotional meanings (Bryzgunova 1984).

The traditional view[1] expressed in the reference grammar of Russian (Bryzgunova 1980) is that intonation covers pitch (F0), timbre, intensity and segmental length; the definition thus employs both segmental and suprasegmental features. Bryzgunova postulates seven *intonational constructions* (*intonacionnye konstrukcii*, IK) which are formally described by a combination of the above-mentioned features, with pitch standing out as the most prominent category. Pitch is the sole or principal distinctive feature for six of the seven IKs. Pitch level and pitch movements in both the prosodic centre of a *sintagma* (the most strongly stressed syllable in a phrase) and the regions preceding and following that centre are constitutive of IKs (see table 1). Phrasing of utterances into sintagmas can also be distinctive.

IKs can have a seemingly infinite number of meanings, and no systematic account is given as to how meanings such as *intensified surprise* (“usilennoe udivlenie”), *musing* (“razdum’e”) or *edifying* (“nazidatel’nyj”) (Bryzgunova 1984) are related to each other or what their connection to the formal properties of the IKs is. While it is sometimes argued that each IK has one primary grammatical meaning (cf. section 3.2), the functional dimension appears to be largely approached with ad-hoc and unbounded labels by advocates of the IK system.

The IK system has been subject to substantial criticism from several sides (most rigorously

[1]It must be noted here that, from the large body of work on Russian intonation, only those descriptions and analyses that are directly relevant to the present study are discussed in this paper. For a more extensive overview of the field, the reader is pointed to Svetozarova (1998); Kodzasov (1999); Yokoyama (2001); Kasatkin (2006); Yanko (2008) and references therein.

IK	contour	pitch level before centre	pitch movement in centre	pitch level after centre	additional features
1	¯ ¯ ∖ _	high	falling	low	
2	¯ ¯ ∖ _	(high)	falling	low	stronger accent than IK-1
3	– – ⁄ _	low	rising	low	
4	— – ⁄ ¯	high	falling-rising	high	
5	– ⁄ ¯ ∖ _	low	1. rising, 2. falling	low	longer centres than IK-2
6	— — ⁄ ¯	low	rising	high	
7	— – ⁄ _	low	rising	low	glottal stop at central vowel

Table 1: The seven intonation constructions (IKs) according to Bryzgunova (1980)

exercised in Odé 1992). The main points of criticism regarded the lack of any experimental evidence for the inventory of IKs as proposed in Bryzgunova (1977, 1980) and the more or less arbitrary and sometimes contradictory choice of IK labels for individual phrases (especially in Bryzgunova 1984). On the basis of several acoustic and perceptual studies using recordings representing different registers of Russian (Odé 1989), Odé (2008b,c) develops an alternative to the IK system named ToRI (Transcription of Russian Intonation).[2] The basic units of ToRI are in some cases taken directly from and and in other cases inspired by the autosegmental-metrical tradition and the Dutch school of intonation studies (Gussenhoven et al. 2003; Gussenhoven 2005). ToRI features six PAs, represented by the respective tones and an asterisk following the accent tone (L*, HL*, L*H, H*L, H*M, H*H), three initial BTs, represented by the percentage sign and a tone level symbol (%L, %M, %H), two final BTs (L% and %, a toneless boundary), and two complex patterns (harmonica (/\/) and sawtooth (>>>)) consisting of special PA sequences (ˆHL* and H*M/(H)L*, respectively).[3] PAs are associated with one main and several other (non-principal) communicative functions each, while BTs are not attributed any communicative functions.

For this study, I will use an extended version of ToRI that I have developed specifically for the purpose of this study. The extended version differs from the original version in having a symmetrical inventory of BTs with 3 IBTs (%L, %M, %H) and 3 FBTs (L%, M%, H%).[4] There are three main reasons for the use of the extended version over the original ToRI system. First, despite having conclusively argued for three perceptually distinctive tone levels (low, mid, high) in Russian herself before (Odé 1989), at no point do Odé (2008b,c) give a motivation for

[2]ToRI offers a free interactive learning module for pitch accents on its website `fon.hum.uva.nl/tori`.

[3]In much recent descriptive literature on prosodic phonology (see e.g. the contributions in Jun 2005 and Jun 2014), only two basic tone levels, L and H, are recognised. Relevant pitch movements situated within or targeting an intermediate or *mid* pitch region, which no doubt exist in a considerable number of languages, are commonly analysed as an upstepped low tone, a downstepped high tone, absence of a tone, or as the result of phonetic adjustment. For example, in recent revisions of Spanish SP_ToBI, the original symbol for a mid final boundary tone M% was replaced by a symbol indicating a downstepped high tone !H% although the existence of a distinctive mid pitch level is still recognised (Estebas-Vilaplana et al. 2015). As far as Russian is concerned, I do not see any benefit that a reanalysis of M, be it in the right half of H*M or as a boundary tone level, could have in ToRI. Therefore, I will be faithful to the original model in Odé (2008b,c) and assume three, as opposed to two, basic distinctive pitch levels.

[4]The symbols for non-accent-lending pitch movements (L, M, H) from the original ToRI model are not discussed in this paper as their theoretical justification is highly questionable and they do not bear any relevance for the current study of boundary tones.

the special status of FBTs (only one tone and a toneless boundary) in ToRI. The three-way distinction is formally implemented in the IBTs and in PAs of the type H*T (H*L, H*M, H*H), but it is mysterious why the ternary distinction is disregarded for FBTs. Bearing in mind the empirical evidence, it seems reasonable to assume three final boundary tones analogously to the three initial boundary tones. Second, while IBTs pattern independently of adjacent PAs, final boundary tones are entirely predictable from preceding PAs in the original ToRI model: if the rightmost element in a PA is low, a following boundary tone is also low; otherwise it is toneless. Third, as my study aims at testing the functional potential of boundary tones, the decision for the revised model also has a practical motivation: it would be a serious impediment to this study if it was construed with an in-built bias in the inventory of the units that it aims to test when there is no a priori reason to assume such an asymmetry in the inventory to begin with.

	%L	**%M**	**%H**	Σ	**Perc.**	**L%**	**%**	Σ	**Perc.**
L*	2	1	2	5	0.04	15	0	15	0.13
HL*	0	2	6	8	0.07	20	2	22	0.19
L*H	3	4	2	9	0.08	0	16	16	0.14
H*L	13	7	0	20	0.17	20	0	20	0.18
H*M	49	8	5	62	0.53	0	25	25	0.22
H*H	13	0	0	13	0.11	0	15	15	0.13
Σ	80	22	15	117		55	58	113	
Perc.	0.68	0.19	0.13		1.00	0.49	0.51		1.00

Table 2: Total count of pairs of adjacent PAs and BTs (without harmonica/sawtooth) as attested in the six "Pitch accents" modules in Odé (2008b). The fact that the number of FBTs is lower than that of IBTs is due to missing FBT tags and pre-boundary non-accent lending pitch movements in isolated cases.

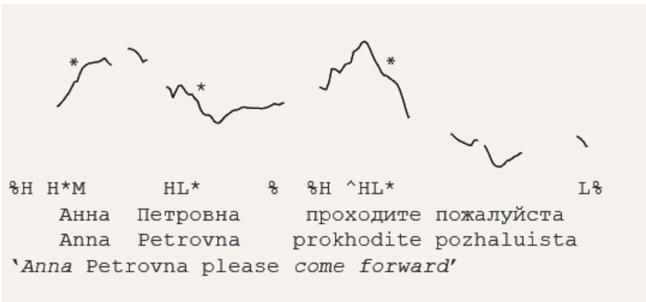

Figure 1: A non-low fall after the first HL* marked as a toneless boundary in ToRI (Odé 2008b, module HL*, example 14).

Table 2 illustrates the non-redundant distribution of IBTs and the redundant distribution of FBTs with respect to adjacent PAs on the basis of the annotated samples available from Odé (2008b). It is worth noting that while HL* is usually followed by L%, we find two instances of HL* followed by %. In one case, this is obviously due to a careless mistake made by the annotator (cf. the well audible final low fall in module H*H, example 10). In the other case, however, the choice of the toneless boundary symbol matches the auditory impression of a non-low fall before a moderately strong boundary (cf. figure 1).[5] It stands to reason that Odé felt the need to somehow account for the fact that the fall in HL* does not reach a low level, and that she therefore chose to deviate from the standard procedure in ToRI by assigning something other than L% after HL*. In what follows, I will show that configurations of low or falling PAs and a non-low pitch level at a final boundary are not exceptions, but a frequently encountered pattern in Russian everyday speech, which underlines the importance of adding non-low M% to the selection of FBTs and strongly supports the choice to allow PAs to combine freely with BTs made in my version of ToRI.

3. Boundary tones and turn-taking

3.1. The role of prosody in turn-taking

The process of structuring talk-in-interaction commonly referred to as *turn-taking* has been argued to follow a set of ordered principles in Conversation Analysis (CA). The basic units recognised in CA are *turns*, which consist of *turn-constructing units* (TCUs) that are defined as portions of speech between *transition relevance places* (TRPs), i.e. places at which a change of speaker is possible (Sacks et al. 1974; Goodwin 1981). TRPs can be indicated by syntactic, pragmatic, prosodic and non-verbal means, but an indication may also be absent. At a TRP, a current speaker has several choices, including *self-selection* and *other-selection*, while a next speaker may choose to comply or not comply with a previous speaker's selection.

Turn-taking is an interactive and dynamic process, and there is strong evidence suggesting that prosodic cues are a crucial device relied on by speakers for organising turn-taking (Couper-Kuhlen and Selting 1996; Szczepek Reed 2006; Sappok 2010; Barth-Weingarten and Szczepek Reed 2014). However, it is still subject to considerable debate how exactly intonational events at prominent positions (PAs) and at the left and right boundaries of prosodic units (BTs) interact with each other. It is also still an open question if complex intonational constructions comprising both PAs and BTs are decomposable at all (as exercised e.g. in Grice et al. 2005), and if so, how the individual components contribute the the communicative meaning of the whole construction (Steedman 2014). With the aid of two studies from the literature, I will now point out two generalisations about possible turn-taking-related functions of boundary tones. On that basis, I will then develop hypotheses about the use of boundary tones in Russian everyday speech for quantitative testing.

With regards to final boundary tones, Rathcke (2009) presents data from acoustic and perceptual experiments showing that Russian high final boundary tones (H%) are interpreted differently by speakers of Russian and German: while the latter tend to associate high FBTs with *friendliness* , the former tend to perceive them as expressing *incompleteness*. Both interpretations are well in line with the *frequency code* (Ohala 1984, Gussenhoven 2004:95), the generalisation that high or rising F0 is typical of submissiveness and questions, while a low or falling F0 is typical of authority and statements cross-linguistically. It can thus be hypothesised that

[5] In fact, the phrase in question could well be described as an intermediate phrase. ToRI does not provide phrase types other than the IP, and the question of how many phrase types need to be considered in autosegmental models of Russian prosody must be left for future research.

high FBTs (underlying a high level or rising final part of a pitch contour) are used by speakers to indicate a turn's incompleteness, i.e. self-selection, and, conversely, that low FBTs (underlying a low level or falling final part of a pitch contour) indicate a turn's completeness, i.e other-selection or no selection.

As far as initial boundary tones are concerned, I argue that high IBTs correlate with planned long turns. In her study on functions of initial "high onsets" (≈ high IBTs) in a four-hour long corpus of recordings from a British radio programme, Couper-Kuhlen (1998) observes a pattern at the beginning of call-in conversations by which the caller may explain the reason for his call at an anchor position following a short introduction by the host and a greeting sequence. Alternatively, callers may also bring up "some other business requiring priority" (Couper-Kuhlen 1998:17). When that business has been dealt with, callers signal to their host that they now plan to talk about the reason why they called by beginning their turn with a high onset. That way, callers gain the possibility to begin subsequent IPs with successively lower onsets and thus structure their intended multi-unit turn. The high onset thus projects the planned "more to come" (Couper-Kuhlen 1998:15), and at the same time serves to reduce the probability of further interruptions from an interlocutor. I will call the correlation of a high pitch level at a left IP boundary and a planned long turn the HIGH ONSET PRINCIPLE (HOP).

(1) THE HIGH ONSET PRINCIPLE (HOP): High initial boundary tones indicate a speaker's following long turn.

The HOP as phrased in (1) will become relevant in section 3.3 when I formulate three quantitatively testable hypotheses. The question whether or not the HOP, having been developed based on observations of a special and inherently structured type of conversation in English, is also applicable to Russian spontaneous speech will be addressed in section 5.

3.2. Intonation and turn-taking in Russian from the perspectives of the IK and ToRI models

Both the IK model and ToRI include several cases for which intonation can be said to convey meanings related to the organisation of talk-in-interaction (despite the fact that neither model makes explicit reference to Conversation Analysis). For example, the main function of the H*M pitch accent in ToRI, *incompleteness*, easily translates into the conversation-analytical concept of *self-selection*, as H*M can serve as a turn-holding device. An example of a function related to turn-taking in the IK model is "activation of another speaker's attention" (Bryzgunova 1977:206) of IK-6. Such "activation" often implies a request for some kind of backchannel activity and thus roughly translates into *other-selection* with a following planned turn resumption.

In table 3, I give an overview of (tentative) main functions of IKs and corresponding PAs in ToRI. The IK model does not provide single main functions for individual IKs, which makes it difficult to derive unified functional descriptions from the enumerations of functions listed in Bryzgunova (1977, 1980, 1984). For now, I follow Bryzgunova (1977:195–208) and list one function that seems to underlie the most commonly cited "meanings" (*značenija*) for each IK. What becomes clear from table 3 is that, although the right phrase periphery is not discussed with respect to communicative functions in either model, the two models make similar predictions for at least two of their basic units that share contour-final pitch properties. IK-1/L* is described as having a flat low contour at the right edge in a phrase and is used in contexts where S1 leaves the floor to another speaker. IK-6/H*M, on the other hand, is characterised by a non-low final pitch level and is used in contexts where S1 does not yield the floor to S2.

IK	**meaning**	**S1 → ...**	**PA**	**main function**	**S1 → ...**	**match?**
1	completeness	S2, Ø	L*	completeness	S2, Ø	yes
2	emphasis	n/a	HL*	completeness with emphasis	S2	no
3	incompleteness	S1	H*L	yes/no question	S2	no
4	incompleteness	S1	L*H	non-first elliptic question	S2	no
5	positive attitude	n/a	H*H	incompleteness	S1	no
6	incompleteness	S1	H*M	incompleteness	S1	yes
7	negative attitude	n/a	n/a	n/a	n/a	n/a

Table 3: Main functions of IKs and corresponding ToRI pitch accents and reanalysis in terms of speaker's choice signals (S1 selects himself/another speaker/unclear).

3.3. Hypotheses

On the basis of the generalisations outlined in the previous sections, the following hypotheses were formulated for quantitative testing:

- Hypothesis 0: Boundary tones are *redundant* with respect to nuclear pitch accents.
- Hypothesis 1: High initial BTs indicate a speaker's *long* turn.
  - Hypothesis 1a: High initial BTs indicate a speaker's *long* turn, as measured in word forms contained in it compared to global average turn length.
  - Hypothesis 1b: High initial BTs indicate a speaker's *long* turn, as measured in word forms contained in it compared to a following turn by a next speaker.
  - Hypothesis 1c: High initial BTs indicate a speaker's *long* TCU, as measured in word forms contained in it compared to global average TCU length.
- Hypothesis 2: High *final* BTs indicate a speaker's *self-selection*.

Hypothesis 0 follows from the distribution of FBTs in the original ToRI system. We expect this hypothesis to be rejected, as BTs can be furnished with independent communicative functions only if they are not redundant with respect to PAs. Hypothesis 1 tests the HOP formulated in (1). Hypotheses 1a and 1b each specify how to measure a turns relative length, as this can be calculated using either a global (average turn length) or a local (length of a next speaker's turn) point of comparison. Hypothesis 1c is a modification of H1 in that the unit it tests is not the turn, but the TCU, the length of which is also relevant for organising turns. Hypothesis 2 probes the perceptual data discussed in Rathcke (2009) and the generalisation over "main" functions in the IK and ToRI systems.

4. Methodology

4.1. Data basis

For the purpose of this study, two recordings of Russian spontaneous speech in a natural environment with an overall length of 278 seconds (801 word tokens) were transcribed and annotated. The first recording (161 seconds, 491 tokens) was taken from Odé (1989) and contains a dialogue of two male speakers talking about a trip to Svijazhsk Island. One speaker starts telling his interlocutor about some of the events that occurred during that trip, while the other speaker interacts with him by backchanneling, commenting, and asking questions. The second recording (117 seconds, 310 tokens) was taken from the ORD corpus (Šerstinova 2009) and contains a conversation in which two female speakers discuss administrative and research-related topics at Saint Petersburg State University. In this recording, the younger of the two speakers is speaking for most of the time. Since the recording quality of the parts spoken by her older interlocutor is rather low, annotations were made only for those parts in which the younger one is speaking.

4.2. Annotation

Recordings were annotated on six tiers for each speaker using Praat (Boersma and Weenik 2014). The six tiers covered turn allocation, turn length, PAs, BTs and voice quality as well as an orthographic transcription (see below). TRPs were determined based on syntactic and pragmatic completeness (since the role of intonation is the subject of this study, pitch-related cues were ignored). IP boundaries were placed at perceptually prominent breaks if they were accompanied by resets; creaky voice, which is typical of IP offsets, was used as a cue as well. FBTs were chosen depending on pitch level (low, non-low) and pitch movement (falling, stable, rising) after the last PA of an IP, whereas IBTs were selected based on where an IP-initial F0-level was situated in an individual speaker's voice range (cf. Nebert 2013).

An example annotation is given in figure 2. The annotated fragment from the first recording depicts S1 summing up a narrative sequence about him and a friend of his not being able to purchase tickets for a boat ride despite getting up at half past five in the morning in order to be first in line at the ticket counter. The first IP is composed of an initial %M, H*M and a final L%, the final portion of the IP is accompanied by creaky voice throughout its final prosodic word. The second IP can be described as a deaccentuated attachment to the first IP, but since both a reset at the beginning (%M) and a PA (HL*) at a local prominence peak can be identified, it is annotated as a separate unit. The L% at the right boundary of the IP coincides with a TRP at which no speaker selection takes place: S1 could continue telling his story, but instead he leaves the floor open for S2 to ask questions or give comments if need be ("." = no selection). S2 latches onto S1's portion and thus immediately accepts the offer ("." = self-selection after no selection); the latching is accompanied by a high onset (%H). S2's yes/no question (H*L) is an echo question that does not necessarily entail other-selection, which is why the TRP is annotated as "no selection", which here coincides with M%. S2's turn ends at that TRP. Tags for PAs and BTs are missing in S1's short second turn ("yes") because the passage is barely intelligible and a F0 contour cannot be reliably reconstructed.

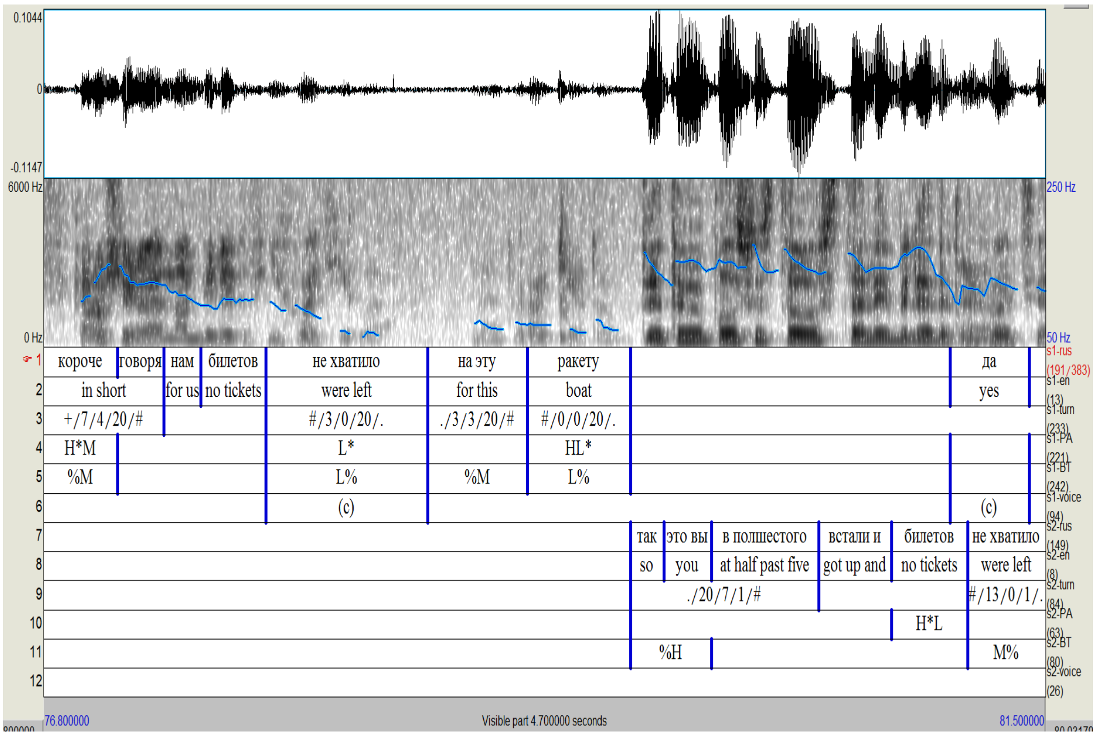

Figure 2: A fragment of a fully annotated TextGrid in the Praat editor window.

The detailed annotation procedure for each tier was as follows:

- *rus*: Orthographic transcription (in Russian), each interval containing one prosodic word
- *en*: English translation (for illustration purposes only)
- *turn*: At each TRP, the following parameters were annotated:
  - *selection at beginning of TCU*: How does a speaker's floor-taking relate to the signal given at the end of the TCU before (options: S1 accepting S0's other-selection, S1 self-selecting after no selection, S1 accepting own self-selection, S1 claiming floor despite of S2's self-selection, n/a)?
  - *turn length*: At a given TRP, how many words are left until the end of the current turn?
  - *TCU length*: At a given TRP, how many words are left until the end of the current TCU?
  - *next turn length*: At a given TRP, how many words does the next speaker's turn contain?
  - *selection at end of TCU*: What signals does a speaker give to his interlocutor (options: S1 self-selecting, S1 selecting S2, no selection, S1 not finishing his TCU or being interrupted, n/a)?

- *PA*: Pitch accents, annotated according to the original ToRI system
- *BT*: Boundary tones (%L, %M, %H, L%, M%, H%)
- *voice*: Voice quality (used for determining boundaries only, not included in the quantitative analysis)

4.3. Machine-learning classification

The fully annotated TextGrids were converted into the arff format and were then processed by four different data-mining algorithms provided by WEKA (Witten et al. 2011). While data-mining techniques are commonly employed when handling larger amounts of data, they have been argued to be amenable to smaller-sized samples as well (Meyerhoff 2015). The algorithms used were *ZeroR*, *OneR*, *LMT* and *SMO*. *ZeroR*, a rule-based classifier assigning all data points to the overall largest class, served as baseline. *OneR* computes simple decision rules based on the single most powerful feature and thus can be helpful in deciding which feature was the overall most important. *LMT* is an algorithm that builds up decision trees, i.e. hierarchical structures reflecting a heterogeneous relevance of individual features. *LMT* outputs factor weights for individual classes. *SMO* is a support vector machine that combines linear models and instance-based learning and computes factor weights for all pairs of classes.

5. Results

Hypothesis 0 stated that BTs are redundant with respect to PAs. As can be seen from table 4, this hypothesis had to be rejected because nearly every combination of PA and BT is attested in the annotated data. L*H and H*H were overall rare, which is why no conclusions can be drawn from the distribution of BTs with these PAs. The same holds for the lack of co-occurrences of H*L and H%, which has to be seen in the light of the overall infrequency of H%. The only pitch accent that consistently combines with only one specific BT (M%) is H*M. Overall, the frequency and distribution of IBTs in my data were comparable to those in Odé (2008b), whereas FBTs patterned in a strikingly different way.

	%L	**%M**	**%H**	Σ	**Perc.**	**L%**	**M%**	**H%**	Σ	**Perc.**
L*	5	6	1	12	0.09	12	10	1	23	0.18
HL*	10	24	14	48	0.36	25	30	1	56	0.44
L*H	0	3	0	3	0.02	0	0	3	3	0.02
HL*	8	14	12	34	0.25	7	17	0	24	0.19
H*M	8	18	5	31	0.23	0	18	0	18	0.14
H*H	1	2	3	6	0.04	0	1	3	4	0.03
Σ	32	67	35	134		44	74	8	128	
Perc.	0.24	0.50	0.26		1.00	0.34	0.59	0.06		1.00

Table 4: Total count of PA/BT combinations attested in the two annotated recordings.

One observation that requires some further discussion is that despite their unbounded distribution, some PAs do seem to exhibit certain preferences with respect to adjacent boundary tones. Thus, we find dispreferences for high IBTs to occur before L* and for low IBTs to occur before HL*. This symmetrical picture is not surprising if we take into account that the acoustic difference between the two falling PAs solely lies in their timing: L* reaches its low goal early, while HL* reaches it late. A late-timed HL* starting from a low level is perceptually close to an early-timed L* starting from a high level, as the latter will reach its low endpoint rather late due to its larger excursion size and the former will reach its low endpoint rather early due to its smaller excursion size. Therefore, we can stipulate that the distribution reflects a strategy of speakers to avoid perceptually disadvantageous IBT/PA combinations in order to preserve the distinctiveness of L* and HL*. The extended version of ToRI thus reveals two aspects that have gone unnoticed in the original version: first, there is no empirical evidence for the implicitly assumed redundancy of FBTs, and second, there are differences in the individual collocation strength of PAs with IBTs and FBTs, yielding prototypical clusters such as %L + L* and H*M + M% (the latter possibly reflecting a more rigid phonological rule, see discussion below).

Hypothesis 1 stated that high initial BTs indicate a speaker's long turn (or, alternatively, a speaker's long TCU or a following short turn by a next speaker). This hypothesis had to be rejected. Classifiers reached only low F-scores, with the baseline (ZeroR) scoring 50.00% (due to a 50% (67/134) majority of %M in the data) and the best performer, SMO, achieving only a slightly better result of 55.07%. Contrary to expectations, OneR (52.17%) calculated *turn allocation* to be the most powerful indicator and gave the rules in (2). These rules can be interpreted as describing a preference of speakers to start with a higher onset the less one's own turn initiation is in accordance with a preceding speaker's turn allocation: compliance with other-selection correlates with %L, no selection correlates with %M, and interruption with %H.

(2) a. $\text{S1} \rightarrow \text{S2} \approx \%\text{L}$
b. $\text{S1} \rightarrow \{\text{S1}, \text{Ø}\} \approx \%\text{M}$
c. S1 interrupts S2 $\approx$ %H

SMO compares pairs of predictor variables and computes weights for each factor. Adding up the weights of all factors available at a given decision point yields either a negative or a positive value; negative values predict the right entity of the pair, while positive values predict the left entity. The SMO model for IBTs in table 5 shows that *turn allocation* is rated high for all three pairs of BTs, with compliance pointing at lower tones (%L or %M) than non-compliance (%M or %H), mirroring the rules generated by OneR. Turn and TCU length are considered only for the (L, H) pair, with weights contradicting the assumptions made in the hypothesis: longer turns and TCUs were found to tend to start with a low, not a high onset.

To sum up, IBTs were an overall weak predictor for turn-taking-related features, allowing only for the tentative generalisation that compliance with turn allocation is negatively correlated with IBT height. The predictions made by SMO clash with those formulated in the hypotheses, which were originally motivated by the HIGH ONSET PRINCIPLE. The non-applicability of the HOP to Russian is not surprising considering it was developed and tested on the basis of English call-in show conversations and not on spontaneous everyday speech from Russian. The results thus do not challenge the validity of the HOP if it is understood as being tied to a specific (and inherently structured) register of a single language, but they do reveal that it cannot claim to

IBT pair	factor		feature		value
L, M	- 1.5	*	turn allocation	=	S1 → S2
	+ 0.5	*	turn allocation	=	S2 interrupts S1
M, H	+ 1.5	*	turn allocation	=	S2 interrupts S1
	- 0.5	*	turn allocation	=	S1 → S1
	- 0.7				
L, H	- 1.1	*	PA	=	L*
	+ 1.0	*	turn allocation	=	S2 interrupts S1
	- 1.0	*	turn allocation	=	no selection
	- 0.7	*	PA	=	H*M
	+ 0.7	*	PA	=	H*L
	- 0.6	*	length(turn)		
	- 0.5	*	length(TCU)		

Table 5: Model for IBTs based on SMO. Predictions with PA = T*H and factors < 0.5 have not been included in this table due to their weak overall conclusiveness.

extend to a broader array of communicative situations in other languages, and should therefore not be misunderstood as a cross-linguistic generalisation.

Hypothesis 2 stated that high FBTs indicate a speaker's self-selection. This hypothesis could be confirmed. The best classifier was LMT with an F-score of 73.85%, which was considerably higher than the baseline score of 58.46%. OneR identified PA as best predictor (the rules are given in (3)), reaching a score of 60.77%.

(3) a. L* → L%
b. {HL*, H*L, H*M} → M%
c. {L*H, H*H} → H%

However, as can be seen from the more fine-grained LMT model in table 6, the reason for the power of OneR owes a great deal to close tie between H*M and M% combined with the overall high frequency of M%. Nevertheless, *turn allocation* does play a role as predicted by the hypothesis once one takes a closer look at the factors and their attributed weights. The model makes use of *turn allocation* for the L% and M% classes (due to sparseness of H%, no solid statements can be made about it, and results for H% will therefore not be discussed here) and clearly predicts self-selection to correlate with M%, not L%. The model's predictions are thus fully in accordance with the hypothesis. In conclusion, FBTs were a solid predictor for turn allocation, with self-selection frequently co-occurring with M%. Also, a strong preference for M% to occur after H*M was confirmed.

FBT class	factor		feature		value
L	- 1.95	*	PA	=	H*M
	- 1.08	*	turn allocation	=	S1 → S1
	+ 0.94				
M	+ 1.90	*	PA	=	H*M
	+ 1.07	*	turn allocation	=	S1 → S1
	+ 0.80	*	PA	=	H*L

Table 6: Model for FBTs based on LMT, predictions for the H class and factors < 0.5 omitted.

The functional load and the non-redundancy of BTs that have become evident so far necessitate a reformulation of the basic rules describing licit IPs in the intonational grammar of Russian. The rules in (4) constitute an intonational grammar of Russian (a simplified version, to be precise, as non-accent-lending pitch movements are not taken into account). They have been deduced from the body of annotated data in (Odé 2008b), though a comparable formalisation is missing in ToRI.

(4) IP → %T + T'
%T → {%L, %M, %H}
T' → T*$^{x}_{0}$ + {L*, HL*, /\/, H*L} + L%
T' → T*$^{x}_{0}$ + {L*H, H*M, >>>, H*H} + %
T' → T*$^{x}_{1}$ + L + L%
T' → T*$^{x}_{1}$ + {M, H} + %
T* → {L*, HL*, /\/, H*L, L*H, H*M, >>>, H*H}

My revised version of ToRI based on the results of this study allows for freer combinations of BTs and PAs and limits redundancy to H*M + M%; it can be described by the grammar in (5).

(5) IP → %T + T'
%T → {%L, %M, %H}
T% → {L%, M%, H%}
T' → T*$^{x}_{0}$ + {L*, HL*, /\/, L*H, H*L, H*H} + T%
T' → T*$^{x}_{0}$ + {H*M, >>>} + M%
T* → {L*, HL*, /\/, H*L, L*H, H*M, >>>, H*H}

6. The benefits of functional decomposition

Enumerative approaches to intonational functions such as Bryzgunova (1984) risk finding themselves tempted with ad-hoc and unsystematic labelling of usage domains (see Prieto 2014 for a recent case in point). Since what I have argued for so far in this paper is effectively decomposing larger complex entities (IPs or intonational constructions) into smaller meaningful elements (PAs and BTs) and identifying particular functions for BTs that complement the previously known functions of PAs, it appears promising to try and further apply such a decomposition to the next smaller complex unit, the PA, as well. Five of the six PAs in ToRI consist of a tone associated with an accentuated syllable (L*, H*) and another tone to the left or to the right of that tone (L, M, H).

The question then arises what functional domains can be ascribed to the starred and the non-starred tones . A proposal offering an account for the former is laid out in Paschen and Sappok (2012). The authors put forward the idea that starred tones link the prosodically highlighted (accentuated) referent with another referent located in the current discourse model. The choice of T* thereby reflects the direction of linking (dubbed *discourse orientation* by the authors): L* (as in L*, HL*, L*H) is leftward linking, while H* (as in H*L, H*M, H*H) means rightward linking. Y/n-questions, for example, are commonly marked with H*L on the element for which the current speaker expects an update from the next speaker in the following turn (anticipation =

temporal rightward linking). In contrast, non-first elliptic questions are often marked with L*H on the element that the current speaker marks as being connected to a previously mentioned proposition (re-activation = temporal leftward linking).

The independent contribution of BTs to communicative meanings conveyed by prosody argued for in this paper is supported by similar recent proposals in formal semantics. In particular, Steedman (2014) presents a surface-compositional account of intonational semantics in which he analyses intonational sequences consisting of a complex PA and an FBT in English. His claim is that those sequences can be decomposed into three meaningful atomic elements, each of which corresponds to a different dimension of information-structural meaning. On his account, the overall shape of PAs (falling/rising) is associated with the dimension supposition/update, alignment or headedness of PAs is associated with the success/failure dimension, while BTs are associated with the speaker/hearer dimension, as summarised in tables 7 and 8. The meaning of an IP as a whole therefore composed of the three meanings introduced by its two components in a fully transparent fashion.

	success	**failure**
thematic supposition	L+H*	L*+H
rhematic update	H*(+L)	(H+)L*

Table 7: Atomic semantics of PAs in English (Steedman 2014).

speaker	(H, L)L%
hearer	(H, L)H%

Table 8: Atomic semantics of BTs in English (Steedman 2014).

On a final note, there is an interesting tension between the decompositional approach and holistic or construction-based approaches such as the IK model Bryzgunova (1980). The core idea of construction-based models is the inseparability of complex form-meaning pairs, and I hope to have offered support in this paper for claiming that intonational constructions in Russian are an illusion.This does not mean, however, that certain indivisible constructions cannot exist in a grammar in which some meaning is also built up compositionally. It is up to future research to determine which elements contributing to the construction of interactional meaning – lexical items, prosodic devices such as voice quality and rhythm, and paralinguistic events such as clicks, to name but a few – are best described as belonging to a specific construction, and which can readily be analysed as having a well-defined meaning on its own combining compositionally with other elements.

7. *Conclusion*

In this paper, I have argued that functional decomposition of intonational patterns into PAs and BTs, each contributing a certain communicative meaning, is a viable approach to intonation as used in Russian everyday speech. I have shown that boundary tones play an integral part in the organisation of talk-in-interaction (*turn-taking*) and that they are overall not redundant with respect to pitch accents, contrary to claims in the literature. Speakers of Russian make systematic use of final boundary tones to manage turn allocation (speaker selection), with the probability of turn transition increasing with higher pitch levels at a right IP boundary. Initial boundary tones seem to be only marginally involved in turn-taking, leaving the question of whether they can be linked to other functional domains for future research.

References

Barth-Weingarten, Dagmar, and Beatrice Szczepek Reed. 2014. Prosodie und Phonetik in der Interaktion – Prosody and phonetics in interaction. In *Prosodie und Phonetik in der Interaktion. Prosody and phonetics in interaction*, ed. Dagmar Barth-Weingarten and Beatrice Szczepek Reed, 4–19. Mannheim: Verlag für Gesprächsforschung.

Boersma, Paul, and David Weenik. 2014. Praat: Doing phonetics by computer. URL `http://www.praat.org/`.

Bradshaw, Mary. 1999. A crosslinguistic study of consonant-tone interaction. Doctoral Dissertation, Ohio State University, Columbus.

Bryzgunova, Elena Andreevna. 1977. *Zvuki i intonacija russkoj reči*. Moscow: Russkij Jazyk, 3 edition.

Bryzgunova, Elena Andreevna. 1980. Intonacija. In *Fonetika, fonologija, udarenie, intonacija, slovoobrazovanie, morfologija*, ed. N. Ju. Švedova, number 1 in Russkaja grammatika, 96–122. Moskva: Nauka.

Bryzgunova, Elena Andreevna. 1984. *Ėmocial'nyo-stilističeskie različija russkoj zvučaščej reči*. Moscow: MGU.

Cohen, A., and Johan ' t Hart. 1967. On the anatomy of intonation. *Lingua* 19:173–192.

Couper-Kuhlen, Elizabeth. 1998. On high onsets and their absence in conversational interaction. *InLiSt* 8.

Couper-Kuhlen, Elizabeth, and Margret Selting. 1996. Towards an interactional perspective on prosody and a prosodic perspective on interaction. In *Prosody in conversation*, ed. Elizabeth Couper-Kuhlen and Margret Selting, 11–56. Cambridge: University Press.

Estebas-Vilaplana, Eva, Yurena M. Gutiérrez, Francisco Vizcaíno, and Mercedes Cabrera. 2015. Boundary tones in Spanish declaratives: Modelling sustained pitch. *Proceedings of the 18th ICPhS, Glasgow* .

Goodwin, Charles. 1981. *Conversational organization: Interaction between speakers and hearers*. Language, thought, and culture. New York: Academic Press.

Grice, M., S. Baumann, and R. Benzmüller. 2005. German intonation in autosegmental-metrical phonology. In *Prosodic typology*, ed. Sun-Ah Jun, 55–83. Oxford: University Press.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Research surveys in linguistics. Cambridge: University Press.

Gussenhoven, Carlos. 2005. Transcription of Dutch intonation. In *Prosodic typology*, ed. Sun-Ah Jun, 118–145. Oxford: University Press.

Gussenhoven, Carlos, Toni Rietveld, Joop Kerkhoff, and Jacques Terken. 2003. Transcription of Dutch intonation. URL `http://todi.let.kun.nl/ToDI/home.htm`.

Hombert, Jean-Marie. 1978. Consonant types, vowel quality, and tone. In *Tone: A linguistic survey*, ed. Victoria Fromkin, 77–111. New York: Academic Press.

Jun, Sun-Ah, ed. 2005. *Prosodic typology*. Oxford: University Press.

Jun, Sun-Ah, ed. 2014. *Prosodic typology II: The phonology of intonation and phrasing*. Oxford linguistics. Oxford: University Press.

Kasatkin, Leonid Leonidovič. 2006. *Sovremennyj russkij jazyk: Fonetika*. Vysšee professional'noe obrazovanie. Moscow: Academia.

Kodzasov, S.V. 1999. Urovni, edinicy i processy v intonacii. In *Problemy fonetiki III*, ed. Leonid Leonidovič Kasatkin, 197–216. Moscow: Nauka.

Krause, Marion. 2007. *Epistemische Modalität*. Wiesbaden: Harrassowitz.

Krifka, Manfred, and Renate Musan. 2014. Information structure: Overview and linguistic issues. In *The expression of information structure*, ed. Manfred Krifka and Renate Musan, volume 5 of *The expression of cognitive categories*, 1–43. Berlin: DeGruyter.

Meyerhoff, Miriam. 2015. The large and the small of it: Big issues with smaller samples in the study of language variation. Talk given at ICLAVE 8, Leipzig, on 28 May 2015.

Nebert, Augustin Ulrich. 2013. *Der Tonhöhenumfang der deutschen und russischen Sprechstimme*, volume 46 of *Hallesche Schriften zur Sprechwissenschaft und Phonetik*. Frankfurt am Main: Lang.

Odé, Cecilia. 1989. *Russian intonation: A perceptual description*. Amsterdam: Rodopi. Leiden, Univ. Diss.

Odé, Cecilia. 1992. Perceptivnaja ėkvivalentnost' realizacij tipov intonacionnych konstrukcij E.A. Bryzgunovoj. In *Studies in Russian linguistics*, ed. A.A. Barentsen, B.M. Groen, and R. Spenger, volume 17 of *Studies in Slavic and General Linguistics*, 227–284. Amsterdam: Rodopi.

Odé, Cecilia. 2008a. Communicative functions and prosodic labelling of three Russian pitch accents. In *Evidence and counter-evidence*, ed. Alexander Lubotsky, J. Schaeken, and J. Wiedenhof, volume 32 of *Studies in Slavic and general linguistics*, 279–296. Amsterdam: Rodopi.

Odé, Cecilia. 2008b. Tori: Transcription of Russian Intonation. A free interactive research tool and learning module. URL `http://www.fon.hum.uva.nl/tori/`.

Odé, Cecilia. 2008c. Transcription of Russian Intonation, ToRI, an interactive research tool and learning module on the internet. In *Dutch Contributions to the Fourteenth International Congress of Slavists, Ohrid: Linguistics*, volume 34 of *Studies in Slavic and General Linguistics*, 431–449.

Ohala, J. J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41:1–16.

Paschen, Ludger, and Christian Sappok. 2012. Drei Konzeptionen der russischen Intonation und ihre diskursorientierte Synthese. *Wiener Slawistischer Almanach* 70:247–293.

Prieto, Pilar. 2014. The intonational phonology of Catalan. In *Prosodic typology II*, ed. Sun-Ah Jun, Oxford linguistics, 43–80. Oxford: University Press.

Rathcke, Tamara. 2009. *Komparative Phonetik und Phonologie der Intonationssysteme des Deutschen und Russischen*, volume 29 of *Sprach- und Literaturwissenschaften*. München: Utz.

Sacks, H., E. A. Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language* 50:696–735.

Sappok, Christian. 2010. Russische regionale Varietäten und Dialekte - eine akustische Datenbank mit diskursiven Annotationen. *Wiener Slawistischer Almanach* 65:163–190.

Šerstinova, Tat'jana Jur'jevna. 2009. The structure of the ORD speech corpus of Russian everyday communication. In *Text, speech and dialogue*, ed. Václav Matoušek and Pavel Mautner, 258–265. Springer.

Steedman, Mark. 2014. The surface-compositional semantics of English intonation. *Language* 90:2–57.

Svetozarova, Natalija D. 1998. Intonation in Russian. In *Intonation systems*, ed. Daniel Hirst and Albert Di Cristo, 264–277. Cambridge: University Press.

Szczepek Reed, Beatrice. 2006. *Prosodic orientation in English conversation*. Basingstoke: Palgrave Macmillan.

Witten, I. H., Eibe Frank, and Mark A. Hall. 2011. *Data mining: Practical machine learning tools and techniques*. Morgan Kaufmann series in data management systems. Burlington, MA: Morgan Kaufmann, 3 edition.

Yanko, T. E. 2008. *Intonacionnye strategii russkoj reči v sopostavitel'nom aspekte*. Moscow: Yazyki Slavjanskich Kultur.

Yip, M. 2002. *Tone*. Cambridge: University Press.

Yokoyama, Olga Tsuneko. 2001. Neutral and non-neutral intonation in Russian: A reinterpretation of the IK system. *Die Welt der Slaven* XLVI:1–26.