Intonational Patterns in Russian Interrogatives

Chapter · January 2006
DOI: 10.1075/ubli.3.11iga

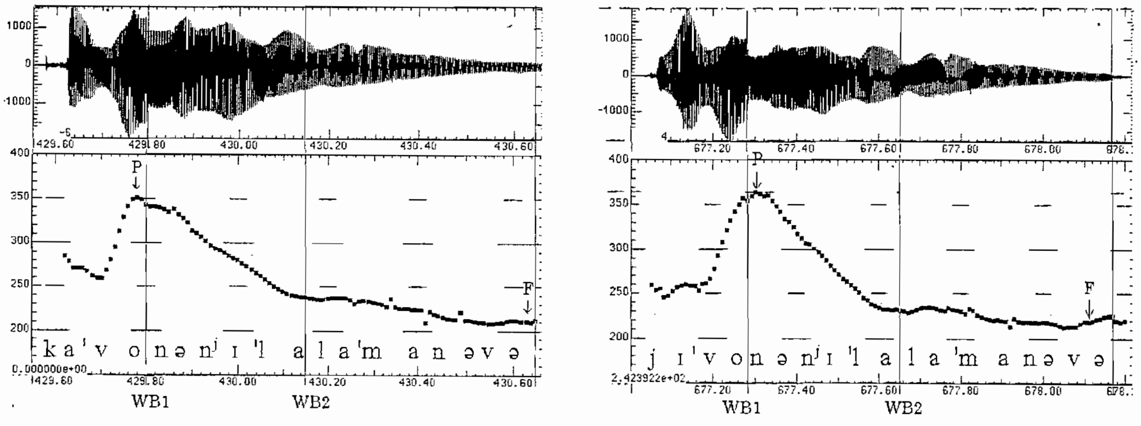
Yosuke Igarashi
National Institute for Japanese Language and Linguistics

Intonational Patterns in Russian Interrogatives — Phonetic Analyses and Phonological Interpretations —

Yosuke IGARASHI

1. Introduction

This article is devoted to empirical and theoretical examinations of intonational patterns which appear in two types of Russian interrogative sentences, wh questions (WHQ) and yes-no questions (YNQ)[1]. In the previous studies, there has been fairly general agreement that the pattern in WHQ and the pattern in YNQ are two categorically different intonational patterns (Bryzgunova 1980; Kodzasov 1996; Svetozarova 1998). It seems, therefore, reasonable to expect that two intonational patterns are phonetically realized as clearly different fundamental frequency (F0) contours. However, visual inspection of the F0 contours, corresponding to the two patterns, reveals that their phonetic differences are by no means obvious. We can



WHQ *Kogó nanjalá Lamánova?*
'Who did Lamanova hire?'

YNQ *Egó nanjalá Lamánova?*
'Is it him that Lamanova hired?'

Figure 1. Waveform and F0 contour for WHQ (left) and YNQ (right), from test sentences of Dataset B in Experiment II, showing the segmental and F0 points measured (see Section 4).

1 Preliminary discussion on the subject of the present article is presented in Igarashi (2004a; 2004b).

notice from Figure 1 that both patterns are realized as strikingly similar contours: they both have a rise-fall configuration in which the rise takes place near the stressed syllable *-gó* in *kogó* 'whom" and *egó* 'him', respectively, and the fall immediately follows the rise.

Curiously enough, there is no reference in literature on the absence of obvious phonetic differences between these two patterns and correspondingly there has been little attempt to examine how they phonetically differ from each other. This leads us to conduct more detailed experimental investigations in order to explore the *phonetic* differences. This is the first goal of this article.

In fact, most of the previous descriptions about the patterns in WHQ and YNQ are quite simple: the patterns have been described, for example, as 'Fall' for WHQ and 'Rise' for YNQ (Bryzgunova 1980; Kodzasov 1999). Obviously, descriptions of this sort do not reflect phonetic properties of the patterns. It would thus be natural to consider them as 'phonological' descriptions of the patterns in which redundant phonetic features are factored out and solely linguistically relevant features are reflected. However, in these studies division of labor between the phonetics and the phonology of intonation is quite confused. In this article, we will propose a new *phonological* interpretation of the intonational patterns in WHQ and YNQ, based on the results of the experiments. This is the second goal of the present article.

The present article is divided into six sections. Chapter 2 will outline the past descriptions on intonational patterns in WHQ and YNQ. In Section 3 and Section 4 two experiments will be performed for the first goal of this article. Section 5 will be for the second goal of the present article: here we will propose a new phonological interpretation of the two patterns on the basis of the experimental results. In Section 6, we will summarize the discussion and raise issues for further research.

2. Description of the pattern in WHQ and YNQ in the previous studies

In this section, we will review how these patterns have been described in the past studies. Before that, a few remarks should be made concerning the fact that some of the previous researchers identified *two* different patterns for WHQ. We will first discuss the patterns for WHQ and define one of them as the subject of the investigation. After that, we will outline the previous description of the patterns for YNQ and WHQ.

2.1. 'Hat pattern' and 'peak pattern' in WHQ

In some of the past works, at least two intonational patterns are reported to appear in WHQ. In what follows, we will refer to them as a 'hat pattern' and a 'peak pattern'. The description of a hat pattern can be found in the works by e.g. Sevetozarova (1998), Nikolaeva (1977) and Kodzasov (1999). This pattern is characterized by the F0 rise in a question word, the F0 fall in the phrase-final word and the high F0 level between the rise and the fall. A peak pattern is described in the studies by e.g. Sevetozarova (1998), Kodzasov (1999) and Bryzgunova (1980). This pattern was illustrated in Figure 1 (left): F0 goes up in a question word and then immediately falls down.

The existence of two types of pattern for WHQ brings us to wonder which of them is more typical, or less marked. In fact, opinions are divided among researchers on this issue. Svetozarova regards the peak pattern as a more marked pattern for WHQ, saying that a peak pattern is a 'rarer' one for WHQ (Svetozarova 1998: 269). Kodzasov agrees with markedness of a peak pattern, arguing that 'a pitch fall on the question word usually occurs only in the case when the following part of the content has already been actualized' (Kodzasov 1996: 96). That is, he claims not only that a peak pattern is marked for WHQ, but also that the choice of each of the two patterns is dependent on the informational structure of WHQ. Bryzgunova (1980), by contrast, appears to regard a peak pattern as almost the only pattern which can be used in WHQ. In Bryzgunova's framework, what we refer to as a hat pattern, accompanied by specific characteristics in duration and pitch range, is classified into the pattern for an exclamation sentence.

While question as to which of the patterns is marked for WHQ and what semantic/ pragmatic differences are distinguished by the two patterns is of theoretical interest, it goes far beyond the scope of the present study. Leaving this question aside, we define a *peak pattern* as the subject of the investigation. We have theoretical and practical reasons for it. First, because our subjects had produced only peak patterns in the experiments, it was impossible for us to examine a hat pattern. Second, since a hat pattern is realized as an obviously distinct F0 contour, it is no use to compare it with the pattern in YNQ.

2.2. Phonetic differences between the patterns in WHQ and in YNQ

The important point in the past descriptions for our discussion is that there has been no mention of the striking similarity in F0 contours between the pattern in WHQ and the pattern in YNQ, and thus little experimental investigation has been made on the phonetic differences between the two patterns. In addition, detailed descriptions of a peak pattern are not presented in most of the past studies. In this respect the Bryzgunova's study, where the peak pattern in WHQ is extensively described, can be seen as an exception. It would therefore be worth while reviewing Bryzgunova's description first.

In Bryzgunova's framework the patterns for WHQ and YNQ are called 'IK-2' and 'IK-3', respectively.[2] As a first approximation, Bryzgunova's characterization of IK-2 and IK-3 can be summarized as following. IK-2 has an F0 *fall* in the accented syllable, while IK-3 has an F0 *rise* in the syllable[3]. For the F0 movement before and after the accented syllable, in her description, there are no significant differences: the F0 is mid-level before the stressed syllable and low-level after it (see Figure 2).

According to Bryzgunova, the phonetic difference between the two patterns thus lies in the F0 *shape* in the stressed syllable: 'Fall' for WHQ and 'Rise' for YNQ. Obviously, the phonetic differences between the two patterns are not as simple as in Bryzgunova's description. This makes it clear that Bryzgunova's description is inadequate as a phonetic description of the two intonational patterns.

Вы на каком факультете учитесь?	*Вы были в кино?*
IK-2 *Vy na kakóm fakul'téte účites'?*	**K-3** *Vy býli v kinó?*
'What faculty are you studying at?'	'Were you in a theater?

Figure 2. F0 contours for two types of intonational pattern, IK-2 (the pattern for WHQ) and IK-3 (the pattern for YNQ), schematically represented by Bryzgunova (1980: 190, 111).

It might appear that Bryzgunova's description is not phonetic but *phonological*: a description of the two patterns as a 'Fall' and 'Rise' might have resulted from removing all the redundant features from the description. It seems, however, not to be the case. We should not overlook the fact that we can find descriptions in Bryzgunova's work (Bryzgunova 1980: 98, 111), in which the F0 configuration in the stressed syllable of IK-2 is defined not only as 'Fall' but also as 'Flat' (*rovnoe dviženie*). This means that Bryzgunova offers two different phonological representations to a single intonational pattern. It is obviously a theoretical drawback providing that her description is phonological. Of course, it is not a problem if we see the 'Fall' and the 'Flat' as phonetic variations of a single phonological category. But again, it seems not to be the case.

---

2 She identified seven types of IK in Russian. 'IK' is the abbreviation of *intonacionnye konstrukcii*.

3 IK-2 is accompanied by an additional feature '*usilenie slovesnogo udarenija*', which serves as a distinction from IK-1, another falling pattern. This feature will not be discussed in this article.

We can not find a clear division in Bryzgunova's work between the level of description where phonological categories are defined and the level where phonetic variations of these categories are defined. In fact both IK-2 and IK-3 contain a variation that has a fall-rise-fall contour and in this case, steepness of the rise is claimed to serve as a distinction between the two patterns: when the rise of the fall-rise-fall contour is steep, then it is classified as IK-3. Here, a 'Fall' vs. 'Rise' contrast shown in Figure 2 plays no role at all. After all, it is not clear what 'distinctive feature' differentiates IK-2 from IK-3, and how a set of phonetic realizations is derived from a single phonological category. It seems that, using an analogy of segmental phonology, Bryzgunova offers solely a list of 'allophones' but provides no description about 'phonemes'. In effect, there are considerable discrepancies between the identified patterns and the F0 contours in her description, as have been criticized by a number of researchers (e.g. Keijsper 1992; Odé 1992; Yokoyama 2003).

It can be concluded, from what is said above, that a description of the patterns for WHQ and YNQ as 'Fall' vs. 'Rise' is not adequate either as a phonetic or as a phonological one. We need to make more detailed examinations into the difference.

As noted above, there is little investigation of the phonetic differences between the patterns for WHQ and YNQ. However, there is a good deal of discussion in the previous works on the comparison of the pattern for YNQ with the pattern for declarative sentences. It is useful to review these works for understanding of the pattern for YNQ.

Makarova (2002) constructed a set of speech stimuli with a rise-fall F0 contour, which varied in alignment and height of the F0 peak, and conducted a perceptual experiment to investigate how the stimuli are semantically interpreted (declaratives, interrogatives and exclamations). One of the relevant points of her findings is that the stimuli with the lower F0 peak (270 Hz) yielded less interrogative judgments than stimuli with the higher F0 peak (320 Hz). This result suggests that the pattern for YNQ is characterized by the high F0 peak. Meyer and Mleinek (forthcoming) studied interactions of intonation with focus structure and pragmatic force (declarative vs. YNQ), and found, among other things, that the F0 peak was scaled higher in YNQ than in declarative sentences irrespective of focus structure. High F0 peaks in the pattern for YNQ have apparently been reported in more traditional works in terms of a 'steep' rise in the accented syllable (Bryzgunova 1980; Svetozarova 1998; Fougeron 1999). In the abovementioned study of Makarova (2002), it was also found that the stimuli with the later alignment of the F0 peak were judged as interrogatives while those with the earlier alignment were judged as declaratives. The later alignment of the F0 peak in

YNQ than in declaratives is also reported in the study of Meyer and Mleinek (forthcoming) cited above.

Although the previous descriptions just outlined were comparisons of the pattern for YNQ with the pattern for declaratives, high scaling and late alignment of the F0 peak can be regarded as ones of the characteristics of the pattern for YNQ. In fact, the F0 contours in Figure 1 suggest that the peak is aligned later and scaled higher in YNQ than WHQ. It is fairly possible that the alignment and scaling of the peak are one of the phonetic differences between the two patterns.

3. Experiment I

Experiment I was aimed at a brief understanding of the phonetic differences between the WHQ pattern and the YNQ pattern.

3.1. Methods

Our basic approach in this experiment was to measure the specific points of the F0 contour in prepared short sentences (WHQ and YNQ), read aloud by native speakers of Russian. The points measured were 1) the F0 value of the beginning of the utterance, 2) the F0 value of the end of the utterance, 3) the F0 value of the peak at the end of the F0 rise and 4) the alignment of the peak at the end of the F0 rise. We expected that there should be a significant difference between WHQ and YNQ in some of these points.

3.1.1. Speech materials

Six sentences were designed, three for WHQ and three for YNQ[4]. They are presented and underlined in Table 1. The test sentences had identical texts except that the WHQ were given a question word *kakój* 'what (kind of)' and the YNQ were given a pronoun *takój* 'this (kind of)'. All the test sentences were provided with corresponding answers. The answer sentences for YNQ were given, so that the focus should be put on the pronoun. We expected that contexts like these would make speakers put a pitch accent on the stressed syllable of *kakój* for WHQ and of *takój* for YNQ. While the question and answer sentences were both recorded, only the question sentences were measured and analyzed. In order to avoid microprosodic effects on the F0 around the accented syllable and the end of the utterance, we used as many sonorants as possible for consonants after the stressed syllable of the question word or the pronoun.

[4] Though they can not be considered as 'sentences' in a strict syntactic sense, we may be safe in assuming that it will cause no problems for the purpose of the present study.

Table 1. Test sentences for Experiment I. The sentences that were measured and analyzed are the underlined question sentences.

WHQ		YNQ	
– *<u>Kakój nómer?</u>*	'What number?'	– *<u>Takój nómer?</u>*	'This number?'
– *Dvenádcatyj.*	'Twelveth.'	– *Da, takój.*	'Yes, it is.'
– *<u>Kakój limón?</u>*	'What lemon?'	– *<u>Takój limón?</u>*	'This lemonade?'
– *Vot ėto, pažalyjsta.*	'This one, please.'	– *Da, takój.*	'Yes, it is.'
– *<u>Kakój limonád?</u>*	'What lemonade?'	– *<u>Takój limonád?</u>*	'This lemonade?'
– *Sládkij.*	'Sweet one.'	– *Da, takój.*	'Yes, it is.'

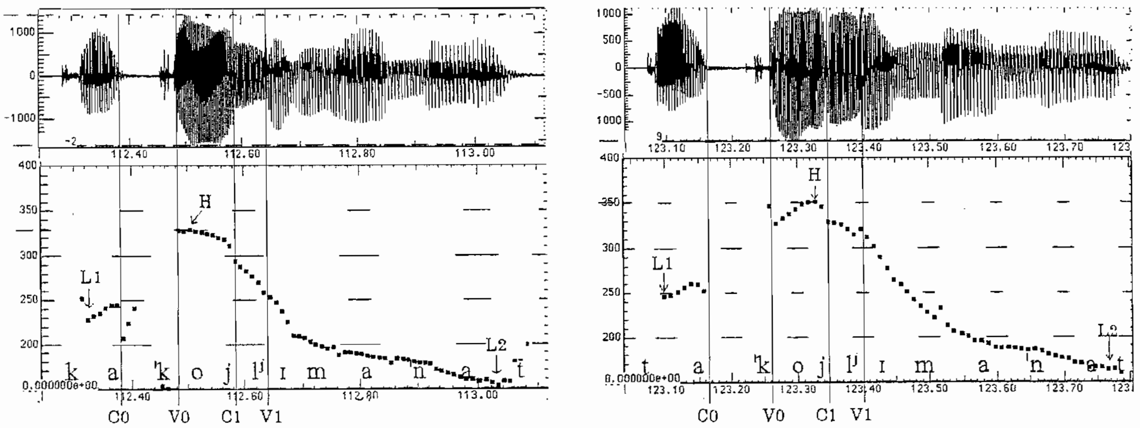
3.1.2. Speakers

The materials were read by four female native speakers of Russian. In what follows, the speakers are identified as A, M, T and Z. M and Z were nineteen years old, and A and M were in their early twenties. They all had been studying at an institute or university in Tokyo at the time of recording. The speakers had no known speech or hearing problems and were naïve as to the purpose of the experiment.

3.1.3. Recording and analysis procedures

The recordings were made on Digital Audio Tape (DAT) in a quiet room at the speakers' home. Speakers read the sentences ten times from a set of cards, each of which had a question sentence (=test sentence) and an answer sentence typed on it. The order of the sentences was random. Speakers were asked to read sentences as naturally as possible. The recorded materials were digitized at a sampling rate of 16 kHz and were analyzed using ESPS Waves+ software.

All measurements were performed manually in a simultaneous display of the waveform, wide-band spectrogram, and F0 track. The utterance beginning and end were measured at corresponding points in the F0 contour and were marked as 'L1' and 'L2', respectively. The peak at the end of the F0 rise was measured at the highest F0 point around the offset of the accented syllable, and was marked as 'H'. The four segmental points measured were the onset of the accented syllable, the onset of the accented vowel, the offset of the accented vowel, and the onset of the following unstressed vowel. These points were marked as 'C0', 'V0', 'C1' and 'V1' respectively (see Figure 3).



WHQ *Kakój limonád?* 'What lemonade?' **YNQ** *Takój limonád?* 'This lemonade?'

Figure 3. Waveform and F0 contour for WHQ (left) and YNQ (right), from test sentences of Experiment I, showing the segmental and F0 points measured.

3.2. Results and discussion

All the data were analyzed on the basis of *t*-tests for each speaker separately, with sentence type (WHQ vs. YNQ) as the independent variable. The dependent variables for each *t*-test were the alignment of the peak (H), the F0 value of the peak (H), the F0 value of the beginning of the utterance (L1) and the end of the utterance (L2). The alignment of the peak was measured by calculating the temporal distance between the onset of the stressed syllable (C0) and the peak (H), divided by the syllable duration (the duration between C0 and C1). Tables 2-5 demonstrate means, standard deviations (SD), *t*-ratios and *p*-values for each point that we measured. Below we will discuss the results in detail.

Table 2. Alignment of H. Means (SD) and individual analyses. *df*=58. Values lower than 1 indicate H is aligned inside the accented syllable.

	WHQ	YNQ	*t*	*p*
A	0.19 (0.3)	1.40 (0.2)	-16.70	<0.001
M	0.39 (0.3)	0.90 (0.1)	-7.46	<0.001
T	0.31 (0.3)	0.84 (0.1)	-6.87	<0.001
Z	0.44 (0.3)	1.15 (0.2)	-9.39	<0.001

Table 3. F0 value of H (in Hz). Means (SD) and individual analyses. *df*=58.

	WHQ	YNQ	*t*	*p*
A	276.86 (9.2)	315.05 (9.7)	-15.50	<0.001
M	332.25 (20.5)	361.88 (11.4)	-6.90	<0.001
T	321.26 (13.5)	389.10 (21.4)	-14.66	<0.001
Z	267.55 (12.9)	316.22 (21.8)	-10.49	<0.001

Table 4. F0 value of L1 (in Hz). Means (SD) and *individual* analyses. df=58.

	WHQ	YNQ	t	p
A	245.32 (10.0)	218.78 (8.1)	11.22	<0.001
M	266.80 (10.0)	258.17 (9.1)	3.47	<0.001
T	251.36 (19.5)	221.39 (12.4)	7.09	<0.001
Z	235.83 (10.2)	222.30 (8.5)	5.54	<0.001

Table 5. F0 value of L2 (in Hz). Means (SD) and *individual* analyses. df=58.

	WHQ	YNQ	t	p
A	176.34 (8.7)	182.83 (5.8)	-3.38	=0.001
M	215.16 (6.2)	215.75 (8.7)	-0.30	=0.764
T	155.23 (8.5)	163.50 (8.6)	-3.71	<0.001
Z	179.84 (10.6)	184.02 (7.3)	-1.77	=0.082

The peak (H) was aligned significantly later in YNQ than in WHQ for all speakers (see Table 2). Although the mean varied considerably across speakers, the general relation was consistent. The peak was aligned around (just before) the middle of the accented syllable for WHQ, while it was aligned around the offset of the accented syllable for YNQ.

There was a significant difference in the F0 value of the peak (H) as well (see Table 3). For all speakers, it was scaled higher in YNQ than in WHQ. Individual differences in F0 values should be due to the differences in a speaker-specific pitch range.

A significant difference was found in the F0 value in the utterance beginning (L1) for all speakers: it was higher in WHQ than in YNQ (see Table 4). Again, individual differences in F0 values should be a reflection of a speaker-specific pitch range.

The results for the F0 value of the utterance end (L2) were rather difficult to interpret. Table 5 shows that the F0 tends to be higher in YNQ than in WHQ for all the speakers. The difference, however, is significant only for speakers A and T. At the present stage, we conclude that it is not clear whether there is a phonetic difference between the two types of pattern in the F0 value of the end of the utterance.

3.3. Summary of Experiment I and its perspectives

The results of Experiment I can be summarized as follows. First, the alignment of the peak at the end of the F0 rise was significantly later in YNQ than in WHQ. It was aligned around the middle of the accented syllable in WHQ and around the offset of the accented syllable in YNQ. Second, the F0 value of the peak was significantly higher in YNQ than in WHQ. Third, the F0 value of the beginning of the utterance was significantly higher in WHQ than in YNQ. Finally, the F0 value of the end of the utterance tended to be higher in YNQ than in WHQ but the difference was significant only for two of four speakers.

From these results, therefore, it was suggested that the phonetic differences between two types of pattern lay in the three points, which are presented in Figure 4. A question as to whether one of the differences lies in the F0 value of the utterance end is still unclear at this stage.

Having summarized the results, we can turn to consideration of what points should be analyzed for further understanding of phonetic differences. We should notice that the following questions remained unanswered in Experiment I, because speech materials that we used were all short utterances. First, since the test sentences had only one syllable before the accented syllable, it was impossible to observe the F0 movement in a region before the accented syllable ('pre-accentual part', henceforth). It is therefore necessary to make an inquiry into whether there are differences in the pre-accentual part between WHQ and YNQ.

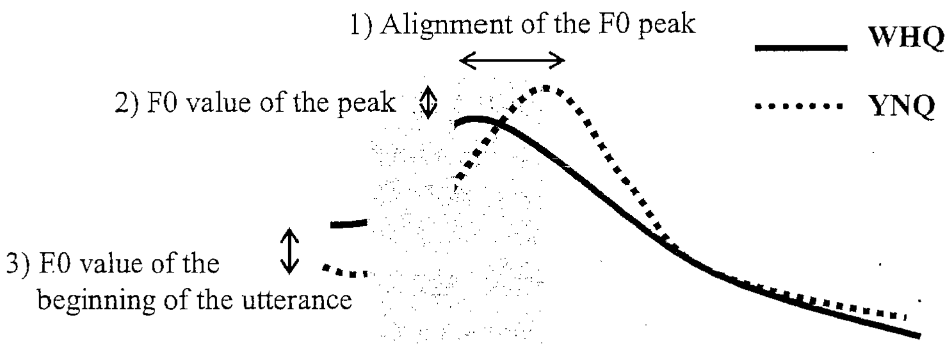


Figure 4. Phonetic differences between the two types of pattern, suggested by the results of Experiment I. Solid and dotted lines represent schematized F0 contours in WHQ and YNQ, respectively. The shaded space indicates the interval of the accented syllable.

Second, the number of syllable following the accented syllable was too small to examine the F0 movement in a region after the accented syllable ('post-accentual part', henceforth). In this part, as we can observe from Figure 1, there is a low turning point at the end of the F0 fall, which the previous researchers have made no close investigation. It is thus feasible to speculate that there are phonetic differences in this point.

In the following section, an experiment will be preformed, in which we will make more detailed examinations of the phonetic differences between the WHQ and YNQ patterns, using speech materials with longer segmental composition, so that the F0 movements in the pre-accentual and post-accentual parts can be satisfactory observed.

4. Experiment II

Experiment II was aimed at examination of the points, which could not be satisfactorily investigated in Experiment I. To be more specific, we investigated whether phonetic differences between the WHQ and YNQ patterns could be found in the F0 movement in the pre-accentual part (interval before the accented syllable) and in the post-accentual part (interval after the accented syllable).

4.1. Methods

Our basic approach in Experiment II was the same as in Experiment I. We asked speakers to read prepared sentences and measured the specific points in the F0 contour.

4.1.1. Speech materials and speakers

We designed two sets of speech materials, one set for the pre-accentual part (Dataset A) and another for the post-accentual part (Dataset B), each of which contained one WHQ and one YNQ. Dataset A is illustrated in Table 6. The test sentences had the identical texts except that the WHQ was given a question word *kogó* 'whom' and the YNQ was given a pronoun *negó* 'him'. Note that the question word and pronoun, both having the lexical stress on the second syllable, were followed by the unstressed preposition *na* 'on', and thus two unstressed syllables preceded the stressed syllable. We expected that this condition allowed us to fully observe the F0 movement in the pre-accentual part.

Table 6. Dataset A. Test sentences are underlined.

WHQ	**YNQ**
– *Na kogó ty smótriš'?*	– *Na negó ty smótriš'?*
'Who are you looking at?'	'Are you looking at him?'
– *Ja smotrjú na Marínu.*	– *Da, ja smotrjú na negó.*
'I'm looking at Marina.'	'Yes, I'm looking at him.'

Table 7 illustrates Dataset B. Again, the test sentences had the identical texts except that the WHQ had a question word and the YNQ had a pronoun. The question word or the pronoun was followed by a three-syllable verb *nanjalá* 'hired' and a four-syllable subject *Lamánova* (woman's surname), so that F0 movement on the post-accentual part could be fully observed. Except one voiced fricative for the final syllable we used sonorant consonants for the syllables following the accented syllable in order to avoid possible microposodic effects.

For both Dataset A and Dataset B, answer sentences were given so that focus would be put on the pronoun in the YNQ. As in Experiment I, question and answer sentences were both recorded, but only the question sentences were measured and analyzed.

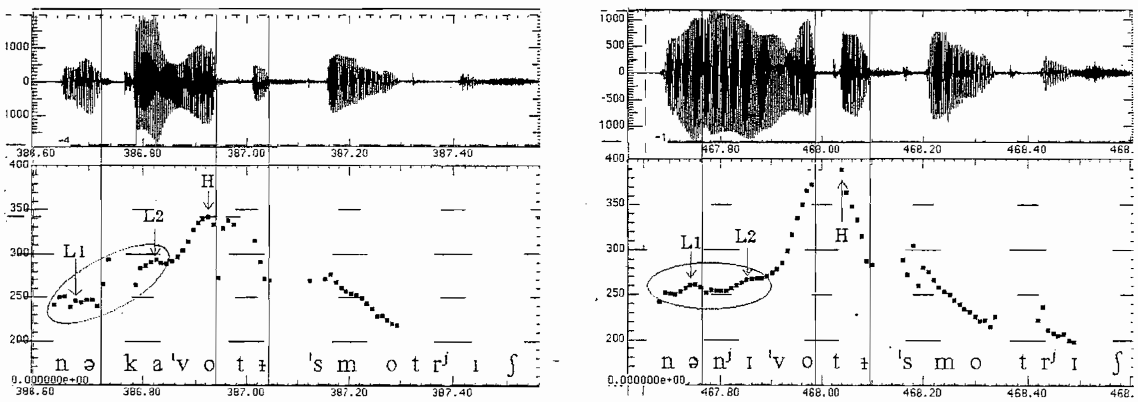
Table 7. Dataset B. Test sentences are underlined

WHQ		YNQ	
– *Kogó nanjalá Lamánova?*		– *Egó nanjalá Lamánova?*	
	'Who did Lamanova hire?'		'Is it him that Lamanova hired?'
– *Marínu*	'Marina.'	– *Da, egó.*	'Yes, (it's) him.'

The speakers were two female native speakers of Russian, A and M, who took part in Experiment I.

4.1.2. Measurement and analysis procedure

Figure 5 shows F0 contours for the test sentences in Dataset A. We can notice that there seems to be a difference in the way in which F0 moves in the pre-accentual part. Thus, for Dataset A, we measured the relatively stable F0 point in the first syllable (L1) and the second syllable (L2). Also, the peak at the end of the rise (H) was measured.



WHQ *Na kogó ty smótriš'?*
'Who are you looking at?'

YNQ *Na negó ty smótriš'?*
'Are you looking at him?'

Figure 5. Waveform and F0 contour for WHQ (left) and YNQ (right), from test sentences of Dataset A in Experiment II, showing the segmental and F0 points measured. For circles emphasizes possible difference in the F0 movement in the pre-accentual part.

For Dataset B, we measured the valley at the end of the F0 fall, which can be observed between the F0 peak and the utterance end. For the measurement of this point, we first measured the F0 peak and the utterance end and marked them as 'P' and 'F', respectively (see Figure 1). Second, two straight lines were fitted to the F0 contour going from P to F. The parameters of the two linear models were estimated by means of the linear least-squares method. To estimate the F0 valley, i.e., the intersection of the two fitted lines, two linear regressions were computed for each possible location (from 1 to *n*, where n is the number of samples within the F0 contour). The location eventually selected as the F0 valley (marked as 'Elbow') was the one leading to the smallest total modeling error (see Figure 6). To investigate the alignment of the valley, the boundary between the first and second words and the boundary between the second and the third words were also measured and were marked as 'WB1' and 'WB2', respectively (see Figure 1).

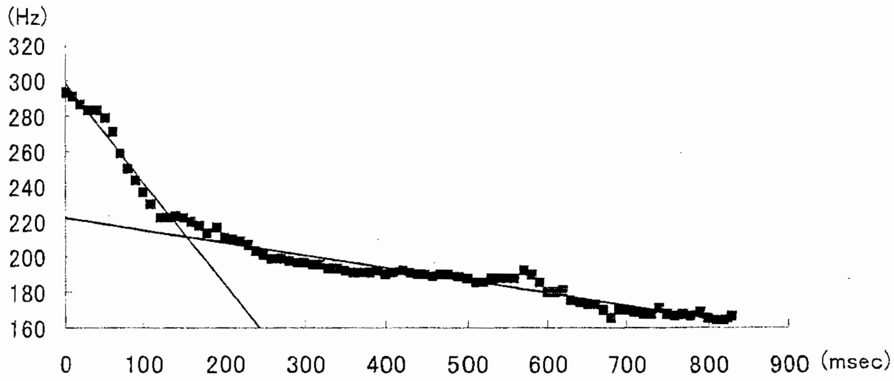


Figure 6. The F0 valley at the end of the fall (Elbow) measured by the least-squares method.

4.2. Results and discussion

First, we will discuss the results of the analyses of the pre-accentual part. Figure 7 illustrates means for F0 values of the first syllable (L1), the second syllable (L2) and the peak (H). We can see that there is an obvious difference in how F0 moves from the utterance beginning to the accented syllable. In WHQ F0 simply rises towards the peak, while in YNQ it remains relatively flat in the first and second syllable and then sharply rises towards the peak.

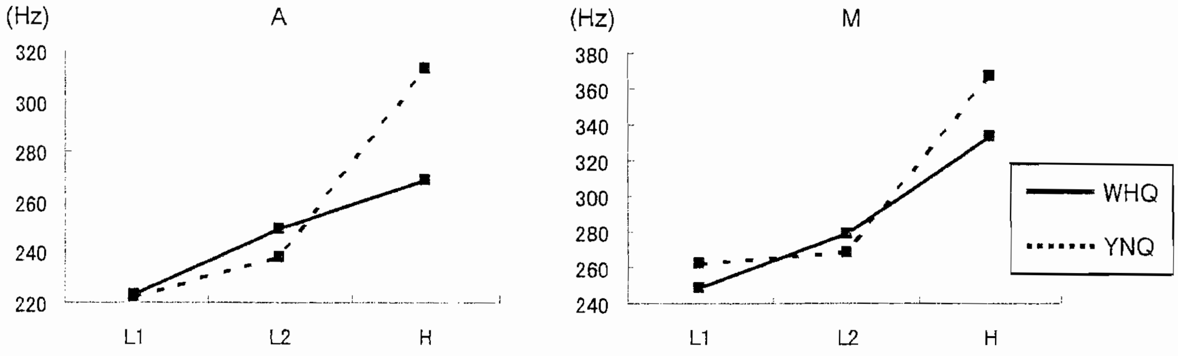


Figure 7. Means for F0 value (in Hz) of L1, L2 and H.

In Experiment I, we found that the F0 value of the utterance beginning was significantly higher in WHQ than in YNQ, and we interpreted this result as indicating one of the phonetic differences in the two intonational patterns. A more detailed examination of the pre-accentual part, however, suggests that it is not true. The F0 height difference found in Experiment I might not be the property of the utterance beginning *per se*. As Figure 7 reveals, the F0 value of the first syllable (L1) presents an inconsistent tendency: for speaker M, L1 is higher in YNQ than in WHQ, while for speaker A, there is almost no difference. The *t*-tests with the F0 value of L1 as the dependent variable revealed a significant effect only for speaker M [df=18 t=-4.31 p<0.001]. The tendency of the F0 value of the second syllable (L2), on the other hand, is consistent across both speakers. L2 is scaled higher in WHQ than in YNQ and the difference was significant for both of the speakers [df=18: A t=2.95 p=0.009; M A t=2.86 p=0.010].

What is noteworthy is that the point measured in Experiment I as the 'utterance beginning' was at the same time the point at the syllable immediately preceding the accented syllable (the point measured as 'L2' in the present experiment). It is thus plausible to assume that the higher F0 value at the 'utterance beginning' of WHQ found in Experiment I was actually the F0 value at the syllable immediately preceding the accented syllable.

We further assume that the difference in the F0 value at the syllable preceding the accented syllable is a reflection of the difference in the way in which F0 moves in the pre-accentual part. To be more specific, we assume the difference to be the presence (YNQ) or absence (WHQ) of the low turning point at the onset of the accented syllable. In YNQ, F0 moves through the low turning point at the onset of the accented syllable, towards the accented syllable, and thus the resultant F0 configuration becomes a combination of the low plateau in the pre-accentual part and the sharp rise in the accented syllable. In WHQ, on the other hand, since there is no low turning point, F0 simply rises towards the accented syllable.

If this assumption is correct, there should be a significant difference in the F0 excursion in the pre-accentual part: the excursion should be much larger in WHQ than in YNQ. In order to examine this, we conducted *t*-tests for each speaker separately, with the F0 difference in Hz between L1 and L2 as the dependent variable. The results were shown in Table 8. Although F0 rises in the pre-accentual part in both types of sentences, the excursion is very small for WHQ, while it is much larger in YNQ, and the difference was significant for both speakers.

In sum, it was suggested that one of the phonetic differences between the two patterns lay not in the F0 value of the utterance beginning, but in the presence (YNQ) or absence (WHQ) of the low turning point at the onset of the accented syllable.

Now let us turn to the discussion on the results of the analyses of the post-accentual part. In order to examine whether there is a difference in the alignment of the F0 valley at the end of the fall (Elbow), we conducted *t*-tests with the temporal distance between the peak (P) and the valley (Elbow) as the independent variable. The results of the analyses are presented in Table 9. We can notice that there is no significant difference for both speakers. It is thus suggested that the valley is aligned by a given distance from the peak and there is no difference in the distance between the two patterns. The result, however, should be treated with caution, because the standard deviations are large. This may imply variability, rather than stability, of the alignment of the valley. We also measured the interval between the end of the first word (WB1) and the valley (Elbow) as well as the interval between the end of the second word (WB2) and the valley (Elbow), but we could not find a consistent tendency in the alignment of the valley.

To examine whether the two patterns differ in F0 height of the valley (Elbow), we conducted *t*-tests with the F0 value of the valley (Elbow) as the independent variable. Table 10 shows the results of the analyses. Again, the tendency is inconsistent. The F0 value was larger in the YNQ for speaker A, while it was larger in the WHQ for speaker M. Deference was not significant for both speakers.

Obliviously, further research is necessary on the alignment and scaling of the F0 valley. At the present stage, we conclude that evidence, showing that there is a phonetic difference in the alignment and scaling of the valley, failed to be found.

Finally, we examined whether there is a phonetic difference in the F0 value of the end of the utterance (F). Table 11 shows the results of the analyses. It can be seen that there is no significant difference for both speakers. We can thus confirm that the phonetic differences between the WHQ and YNQ patterns are not found in the F0 value of the utterance end.

Table 8. Excursion from L1 to L2 (in Hz). Means (SD) and individual analyses. df=18. Positive values mean that F0 rises in the pre-accentual part.

	WHQ	YNQ	t	p
A	26.14 (5.6)	15.54 (6.5)	3.87	=0.001
M	31.02 (8.7)	6.84 (7.1)	6.76	<0.001

Table 9. Temporal distance between P and Elbow (in msec). Means (SD) and individual analyses. df=18.

	WHQ	YNQ	t	t
A	241.07 (41.0)	218.09 (33.7)	1.36	=0.189
M	291.74 (43.8)	273.85 (38.0)	0.97	=0.343

Table 10. F0 value of Elbow (in Hz). Means (SD) and individual analyses. df=18.

	WHQ	YNQ	*t*	*p*
A	198.4 (4.9)	203.88 (6.3)	-2.14	=0.047
M	250.83 (11.4)	241.30 (7.3)	2.21	=0.042

Table 11. F0 value of F (in Hz). Means (SD) and individual analyses. df=18.

	WHQ	YNQ	*t*	*p*
A	176.82 (5.4)	175.10 (8.7)	0.52	=0.607
M	208.14 (6.1)	208.33 (5.9)	-0.06	=0.947

4.3. Summary of the experiments

The results of Experiment II can be summarized as follows. First, there was a significant difference in the way in which F0 moves in the pre-accentual part: in the WHQ F0 simply rises towards the peak, while in YNQ it remains relatively flat until the onset of the accented syllable and then sharply rises towards the peak. Second, we have not found evidence, indicating a possible difference in the alignment and scaling of the F0 valley at the end of the fall.

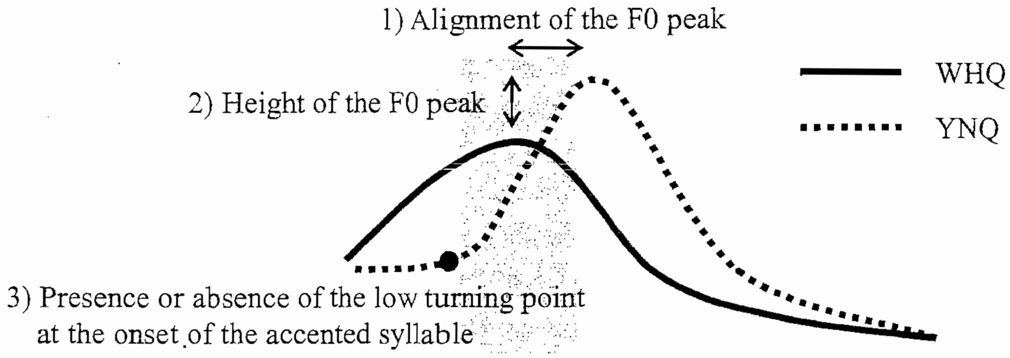


Figure 8. Phonetic differences between the two types of pattern, suggested by the results of Experiment I and Experiment II. Solid and dotted lines represent schematized F0 contours in WHQ and YNQ, respectively. The shaded space indicates the interval of the accented syllable.

From the results of two experiments conducted in the present study, it can be concluded that the phonetic differences between the two patterns are found in 1) the alignment of the peak, 2) the height of the peak, 3) the presence or absence of the low turning point at the onset of the accented syllable (see also Figure 8).

Let us now look at relevant descriptions in the previous works and see how they differ from the results of our experiments. First, the result of Experiment I, showing that the F0 peak in the YNQ pattern was aligned around the end of the accented syllable, is compatible with the studies by Makarova (2002) and Meyer and Mleinek (forthcoming), in which it was shown that the F0 peak was aligned later in YNQ than in declaratives.

The result of Experiment I, concerning the height of the F0 peak in the YNQ pattern also confirms the two studies cited above (Makarova 2002; Meyer and Mleinek forthcoming), where higher scaling of the F0 peak in YNQ than declaratives was reported. In addition, the high peak scaling in the YNQ pattern would give justification to the traditional works (Bryzgunova 1980; Svetozarova 1998; Fougeron 1999), where the F0 rise at the accents syllable is described as 'sharp'.

Finally, we should notice that the results of Experiment II do not bear out Bryzgunova's description about the pre-accentual part. While Bryzgunova (1980) claims that the F0 movement in the pre-accentual part is the same for WHQ and YNQ, our experiments revealed that there was a significant difference in this part. Our results concerning the pre-accentual part should be regarded as a new finding, which could not be accomplished by describing the patterns simply as 'Fall' vs. 'Rise'.

5. Phonological interpretation of the patterns

In the present section, we will propose a phonological interpretation of the WHQ pattern and the YNQ pattern on the basis of the results of the experiments.

First, we assume that an intonational contour is a string of phonological level tones, Highs (H), and Lows (L), occurring at a specific point in the utterance. This is the assumption that underlies the theoretical frameworks in recent phonological research of intonation (e.g. Pierrehumbert 1980; Ladd 1996; Gussenhoven 2004). In this theory (often called Autosegmental-Metrical (AM) theory), movements such as rises and falls *per se* are regarded not as the intonational primitives but merely as transitions from one tone to another. The rise, for example, is taken to be a transition from L to H.

The tonal string involves 'pitch accents', 'phrase tone' and 'boundary tone'. Pitch accents are tones that appear at the stressed syllables. Phrase tone is a tone that spreads over the entire region from the phrase-final pitch accent to the end of the phrase[5]. Boundary tone is a tone that appears at the edge of the phrase. Pitch accent can consist of a single tone ('monotonal') or of a sequence of two different tones ('bitonal'), whereas phrase tone and boundary tone are always monotonal.

Following a conventional notation in the standard AM theory, a tone forming a pitch accent is marked by an asterisk or a *star* '*' (e.g. 'H*') and a boundary tone is marked by a percent sign '%' (e.g. 'H%'). When a pitch accent is bitonal, i.e. composed of a HL or LH sequence, then the tones are

---

5 What we call 'phrase tone' is referred to as 'phrase accent' in Pierrehumbert (1980) and as 'phrasal tone' in some of the AM literature.

connected by a plus sign "+", and one of them is marked by a star (e.g., 'H*+L'). Here, the star notation indicates timing difference of the tones: a tone marked by a star ('starred tone') represents temporal co-occurrence, or *association* of the tone with the stressed syllable. A tone without a star ('unstarred tone') is not associated with any syllable, but it appears by a given time interval before or after the starred tone. The notation 'H*+L', for example, indicates that the H is associated with the stressed syllable and the L appears later by a given time than the H.

Naturally enough, the definition of these phonological tones requires a phonetic basis. Although various sorts of conceptualization of tones have been proposed in the AM literature, we will consider turning points in the F0 contour as phonetic realization of tones: the F0 peak and F0 valley are regarded as a reflection of H and L, respectively. The equation of tone with an F0 turning point might be considered as the most clear and simple way of abstraction of phonological tones (see Ladd 1996: 103-107).

Having outlined the framework that will be applied, we will now turn to propose phonological representations of the WHQ pattern and the YNQ.

Boundary tones. The F0 point at the end of the utterance can be taken as a phonetic realization of a boundary tone. Since the results of the experiments showed that there was no significant difference between the WHQ pattern and the YNQ pattern in the F0 height of the end of the utterance, we propose that both patterns have the same boundary tone, that is, L%[6].

Phrase tones. Readers might recall that for both patterns, we could observe the low F0 region between the F0 peak and the end of the utterance. Insofar as phrase tone is a tone spreading over the entire region from the final pitch accent to the end of the phrase, we can regard this F0 region as a reflection of a phrase tone. Since the F0 value of this region was fairly low,

---

6 As far as we know, there are no published studies that explicitly show that the Russian intonational system has pitch movements localized at the right edge of the phrase (for example, the pitch rise at the phrase-final syllable observable in the yes-no question in English (see e.g. Pierrehumbert 1980). If Russian did not have local pitch movements characterizing the edge of the phrase, then we should not define boundary tones for this language. (In fact, we have not introduced boundary tones to Russian in our previous study (Igarashi 2002). However, in the recording of Russian speech, we often observe intonational patterns with the pitch rise precisely localized at the phrase-final syllable. According to the intuition of our Russian informants, these patterns are "fairly natural". Although we can not exclude the possibility that the patterns with the final rise are specific for a given group of speakers (generation, gender, for instance), we believe that it is necessary to take into account these patterns as well for the comprehensive description of Russian intonation. The definition of boundary tones allows us to represent, for example, a pattern with the final rise, which appear in Russian YNQ, as L+H* L H%.

we can propose that both patterns have a phrase tone consisting of L. Moreover, we measured the alignment and scaling of the valley at the beginning of the low F0 region, and we failed to find evidence that they significantly differed between the two patterns. As noted in 2.2, much work remains to be done on the scaling and alignment of this point, but if further evidence for the identical behaviors of this turning point between WHQ and YNQ is found, then it will give strong support for the claim that both patterns have the same phrase tone, that is, L.

Pitch accents. The F0 peak and valley in the vicinity of the stressed syllable can be regarded as a reflection of a pitch accent. We will first consider the pitch accent in the YNQ pattern. The experimental results showed that for the YNQ pattern, the peak at the end of the F0 rise was aligned near the offset of the accented syllable, and that there was a low turning point at the onset of the accented syllable. Considering the peak and the valley appearing near the accented syllable as a realization of the H tone and L tone, respectively, we propose that the YNQ pattern has a bitonal pitch accent, consisting of a LH tonal sequence. A problem arises as to the star notation that indicates the timing difference of the tones. Since the valley and peak are aligned in the periphery of the stressed syllable (near the onset and offset, respectively), it is difficult to decide which of the two tones (L or H) is associated with the syllable[7]. More research is necessary to make a crucial decision over star notation. Specifically, we should examine whether Russian has two rising pitch accents (L+H* vs. L*+H) which contrast in the timing difference of the tones[8]. Acknowledging that there is likely to be some dissatisfaction, we propose that the H tone is the starred tone and that the pitch accent in the YNQ pattern is L+H*.

Let us now move onto a consideration of the pitch accent in the WHQ pattern. One alternative would be H*, a monotonal high pitch accent. Interpretation of H as a starred tone is justified by the results showing that the F0 peak, a reflection of H, is aligned around the middle of the accented syllable. Also it is entirely fair to say that the pitch accent does not consist of the LH tonal sequence albeit the F0 contour in WHQ presents a rising

7 The same problem was addressed firstly in Arvaniti et al. (1998) for the study of rising prenuclear pitch accents in Modern Greek. In this language, the rising pitch accent is characterized by a clear F0 valley and peak, both of which are aligned well *outside* of the stressed syllable. The theoretical problems of the treatment of this phenomenon within the AM theory are exhaustively discussed in Arvaniti et al. (2000).

8 We proposed elsewhere (Igarashi 2002) that Russian had a L+H* vs. L*+H contrast, but without empirical evidence for it. In that work we also proposed that Russian has two falling pitch accents H+L* and H*+L. The timing contrast for Russian falling pitch accents was firstly proposed by Odé (1989) and was confirmed in our experimental study (Igarashi 2005).

configuration near the accented syllable, because the results revealed that the WHQ pattern lacks a low turning point, a reflection of L, at the onset of the accented syllable. Note that in the H* representation, the F0 fall after the accented syllable is regarded as a transition between H* pitch accent and L phrase tone. Support for the H* representation might be found in Bryzgunova's description of the F0 at the accented syllable being 'flat' (Bryzgunova 1980: 98, 111).

Another alternative for the pitch accent in WHQ would be H*+L, a bitonal falling pitch accent. This representation would receive plausibility from the 'orthodox' description in the previous works in which distinction between the WHQ and YNQ patterns are described as 'Fall' vs. 'Rise' (Bryzgunova 1980; Kodzasov 1999). What has to be noticed, however, is that there is no phonetic evidence for L of a HL tonal sequence. If the WHQ pattern is represented as H*+L L L%, then there will be one more L tone than in the representation of the YNQ pattern, i.e. L+H* L L%. Evidence for the existence of the 'additional' L would be, for example, that the F0 valley after the F0 peak were aligned earlier or scaled lower in WHQ than in YNQ. Such evidence, however, has not been found in the present study.

More evidence and discussion are necessary to establish the representation of the pitch accent in WHQ. One could imagine further findings about the F0 valley after the F0 peak might be used to decide which of the alternatives is valid. At this stage, H* would be preferred to H*+L, because of the lack of phonetic evidence for L of H*+L, though the possibility of H*+L should not be completely ruled out.

In sum, the phonological representations for the WHQ and YNQ patterns proposed in the present study are as follows.

WHQ	H*	L	L%
YNQ	L+H*	L	L%

In the representations proposed here, two patterns contrast in the *shape* of the pitch accent, one has a monotonal high pitch accent and the other has a bitonal rising pitch accent. One might notice that parts of the phonetic differences found in the experiments, namely alignment and scaling of the F0 peak are not reflected in the representations. Insofar as the differences are phonologically interpreted as the *shape* of the pitch accent, we might be safe in assuming that alignment and scaling are redundant features, assigned in phonetic implementation.

6. Conclusion

In the present study, experiments were performed to investigate phonetic differences between the pattern for WHQ and the pattern for YNQ, both of which are realized as similar rise-fall F0 contours. Since in the previous works on Russian intonation there has been little attempt to examine the phonetic properties of the two intonational patterns, we believe that our findings will contribute to Russian intonational research in general. More research is needed to investigate the F0 valley at the end of the fall, which would give empirical and theoretical implications for the study of Russian intonational patterns.

Also, the experimental results were phonologically interpreted, and phonological representation of the patterns was proposed, applying the autosegmental-phonological theory. A question remained as to whether pitch accent in the WHQ pattern should be H* or H*+L. One of the issues for further research would be whether differences in F0 alignment and scaling, which we considered as redundant features, function as a distinction for Russian pitch accents in light of other intonational patterns. We hope the present study would stipulate research on the phonology of Russian intonation.

References

Arvaniti, A., D.R. Ladd and I. Mennen 1998. "Stability of tonal alignment: the case of Greek prenuclear accents". *Journal of phonetics* 26. 3-25.

Arvaniti, A., D.R. Ladd and I. Mennen 2000. "What is a starred tone? Evidence from Greek". *Papers in laboratory phonology* V. Cambridge: Cambridge University Press. 119-131.

Bryzgunova, E.A. 1980. "Intonation (in Russian)". *Russkaja grammatika*, Tom I. 96-122. Moscow: Akademija Nauk SSSR.

Fougeron, Irina 1999. "Intonation and unity of text (In Russian)". *Problemy fonetiki* III. Moscow. 216-227.

Gussenhoven, Carlos 2004. *The Phonology of tone and intonation.* Cambridge: Cambridge University Press.

Igarashi, Y. 2002. "So-called "neutralization of intonation" –Phonological description of Russian intonation (In Russian)". *Bulletin of the Japanese Association of Russian Scholars* 34. 15-21.

Igarashi, Y. 2004a. "Phonetics and phonology of intonation in interrogative sentences in Russian (in Russian)". *Bulletin of the Japanese Association of Russian Scholars* 36. 85-92.

Igarashi, Y. 2004b. "Experimental Phonological studies on Russian question intonation (In Japanese)". *Gengo johogaku kenkyu hokoku* 4. 255-276.

Igarashi, Y. 2005. "How many falling intonation patterns in Russian?: categories of F0 alignment". *Programme & Book of Abstracts: Between and Stress and Tone*: 32-33.

Keijsper, C.E. 1992. "Recent intonation research and its implications for teaching Russian". *Studies in Slavic and general linguistics* 19.

Amsterdam: Rodopi. 151-212.

Kodzasov, S.V. 1996. "Combinational model of phrasal prosody (In Russian)". *Prosodičeskij stroj russkoj reči*. Moscow: Rossijskaja Akademija Nauk. 85-123.

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

Makarova, V. 2002. "Perception of F0 peak alignment differences by Japanese and Russian subjects". *Proceedings of the 16th General Meeting of the Phonetic Society of Japan*. 181-186.

Meyer, R. and I. Mleinek forthcommig. "How prosody signals force and focus - a study of pitch accents in Russian yes-no questions". To appear in: *Journal of Pragmatics*.

Nikolaeva, T.M. *Phrasal intonation of Slavic languages* (In Russian). Moscow.

Odé, C. 1989. *Russian Intonation: A perceptual description*. Amsterdam: Rodopi.

Odé, C. 1992. "Perceptual equivalence of realization of types of intonational constructions of E.A. Bryzgunova (In Russian)". *Studies in Slavic and general linguistics* 19, Amsterdam: Rodopi. 227-284.

Pierrehumbert, J. 1980. *The phonology and phonetics of English intonation*. Doctoral dissertation, MIT.

Svetozarova N.D. 1998. "Intonation in Russian". *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press. 261-274.

Yokoyama, O.T. 2003. "Neutral and non-neutral intonation in Russian: A reinterpretation of the system of intonational constructions (In Russian). *Voprosy jazykoznanija* 5. 99-122.