Computer Science Journal of Moldova, vol.**14**, no.1(40), 2006

Intonational Structures in Romanian Yes-No Questions

Vasile Apopei, Doina Jitcă, Adrian Turculeţ

Abstract

This paper presents the conclusions resulted from an intonational analysis of Romanian Yes-No questions. The recent analysis results consist in dividing and structuring the F0 curves into intonational units. Each intonational unit is described by a tone sequence using ToBI labels used in annotation of the most important phonetic events: pitch accents and boundary tones. The authors of the present study propose a description of the resulted patterns for F0 contour in terms of intonational units structures described by their tone sequence. We consider this description suitable for the variety of melodic contours resulted from different speakers and different focalizations in their utterances. In paragraph 3, the paper presents the intonational variants resulted from our speech corpus analysis. The conclusions of Yes-No question analysis are important for linguistic studies and in Romanian speech synthesis.

1 Introduction

This paper presents the conclusions resulted from an intonational analysis of Romanian Yes-No questions. This study continues our previous works, the results of which were published in [4], [5], [6]. In the previous study [6] we performed a phonetic and auditory analysis for the "neutral" utterances, only (uttered without an intention to focus certain words) and we presented a description of their intonation based on the ToBI annotation system.

In the present work we are interested both by the utterances "without focus" and those "with focus" on different words of the sentence. The recent analysis results consist in dividing and structuring the F0 curve into intonational units. Each intonational unit is described by a tone sequence using ToBI labels used in annotation of the most important phonetic events: pitch accents and boundary tones [3]. These results are comprised in a series of tables, which contain in each row the description of an intonational variant. The phonological and semantic analysis of Yes-No questions implies the establishment of interrogative emphasis position and, consequently, this information is contained in these tables.

Laurenţia Dascălu-Jinga, in her study on the patterns of Romanian melodic contours, presents conclusions about the patterns of final contours and the positions of interrogative emphasis in the case of Yes-No "neutral" questions [1].

The authors of the present study propose a description of the resulted patterns for F0 contour in terms of intonational units structures, each of them being described by their tone sequence marked by the labels of ToBI annotation system. This description modality of the F0 patterns is valid both for "neutral" and "non-neutral" cases of the Yes-No questions.

We consider this description suitable for the variety of melodic contours resulted from different speakers and different focalisations in their utterances.

The F0 contour was divided into intermediate phrases. In this paper, we denote both intermediate and intonational phrases by the term "intonational unit". In paragraph 3, the paper presents the intonational variants resulted from our speech corpus analysis. The speech corpus is generated by a methodology presented in the next paragraph.

2 The methodology for building the speech corpus

In generating the speech corpus we chose three types of sentences to be uttered, with the final word being oxytone, paroxytone and, respec-

tively, proparoxytone. In the corpus we had the intention to generate utterances with various final F0 contour patterns. The texts are the following:

Ai văzut acest ***afiş****? (Did you see this* ***poster****?)*

Ai văzut afişul ***acesta****? (Did you see* ***this*** *poster?)*

Ai văzut regele? (Did you see the king?)

These texts were uttered by speakers who live in different regions of Romania: from Moldova (speakers AT and LS), south-west Transilvania (speakers LM and IM) and Banat (speaker TM). Each speaker generated 5 "neutral" utterances, without any intention to focus a word, of each text corresponding to oxytone, paroxytone and proparoxytone cases.

Then, speakers uttered each text in different interpretations (five times per each interpretation): with focus on the verb, with focus on the first word of nominal group, and with focus on the second word of the nominal group. We included the focus indications for speakers, in the last column of the tables that contain pattern descriptions.

The results of F0 contour interpretation are presented in the next paragraph.

3 The phonetic and phonologic analysis of the utterances

The purpose of our phonetic and phonologic analysis is to divide the utterances into intonational units (intonational phrases or intermediate phrase) and then to annotate the tonal events on F0 curve, as generating the tone sequence for each of them. In annotation we used the ToBI label system [3].

An intonational unit represents a segment of the F0 curve that contains one primary pitch accent and, possibly other less prominent accents. The primary accents are visible either on the F0 curve, by significant F0 frequency variations, or by the energy level during one stressed syllable within the corresponding unit. The significant F0 frequency variations during the unstressed syllables following a primary

accent can represent an indication for the boundary between two consecutive units. These units can comprise one or many words linked by a meaning or by the syntactic structure.

The identification of the primary tonal accents on F0 curves entails their division into intonational units. The F0 patterns of intermediate phrases resulted from our analysis was grouped into two major categories:

- Units characterised by a F0 contour that begins with a down-stepped trend of the F0 frequency, until a low tone appears, and then finishes with an increasing segment of F0 frequency. We denote this type of unit with "A";
- Units characterised by a F0 contour that begins with a up-stepped trend of the F0 frequency, until a high tone appears, and then finishes with an decreasing segment of F0 frequency. We denote this type of unit with "B".

The definition of "A" and "B" units is a very general one. An "A" unit can be generated by different tone sequences, as the following: L* L-H%, H+!H* H* H-%, L* H-H%. The same for a "B" unit, it may be performed by different tone sequences, as the following: H* L* L-L%, H*+L L-L%.

These two basic prototypes of intonational unit patterns have different variants. In some cases an extra short segment corresponding to the final unaccented syllables (after the last pitch accent) is added. Thus, at the end of an "A" unit a fall of the tone may exist while, at the end of a "B" unit, an increase of the tone may occur. Both extra segments are characterised by a small range of the tone variation and by a short time length.

In the beginning of the sentences, within an "A" unit it can be missed the first down-stepped segment, while, in the end of the sentence, in a "B" unit the first up-stepped one may be missed.

In the case of pitch contours characterised by a small range between *Low* and *High* levels, the "A" and "B" patterns of F0 curve can not be identified because the contour is almost a flat one. In ToBI annotation

system, these segments are characterised by the value *Low* of the parameter HIF0. This kind of contours can spread over an entirely unit (we consider it as a particular case of an "A" unit) or make a segment into an intonational unit of type "A".

We claim that one or more intonational units of one type described before, linked in a sequential structure, can define the intonation of an utterance. In our perspective, the F0 contour shape of one utterance may be divided into a pattern sequence of "A" or "B" type.

The description of intonational variants are based on the following labels:

- H*, L* – for high and respectively, low monotonic pitch accents,
- L+H*, L*+H, H+L* – for bitonal pitch accents with a high or low tone in a middle of accented vowel. The tone is marked by "*";
- H+!H*- for a bitonal accent that marks a decreasing pitch during accented syllable, but not until a low tone;
- H-, L- for the phrase accent tones that mark the end of an intermediate phrase;
- %H, %L for the phrase accent tones that mark the beginning of an intonational phrase;
- H%, L% – for the boundary tones in the end of an intonational phrase;
- HIF0 – for the ToBI parameter that indicates the Topline level used in pitch range modelling. It takes the following values: L-low, H-high;
- H(*) and L(*) – for the accents H* and respectively L* marked in the case of oxytone final word units .

The diacritics "^" and "!" that precede a high tone indicate a higher level, respectively, a lower level than the previous high tone.

To introduce various information in the description of the intonational variants we use the following conventions:

- with bold font we marked the word with a primary accent;

- with bold and underlined letters we marked the word which carries the interrogative emphasis;
- the words with high level of energy are marked between round brackets;
- the length of accented syllables, on which a tonal event is marked, are measured including all voiced phonemes;
- the value *High* (H) of parameter HIF0 is implicit and thus, is missing in annotation of all tonal accents, which contain high level tones. It is explicit in the case of a Low value for the parameter HIF0.

In the tables from the paragraphs 3.1, 3.2 and 3.3 the information concerning the duration of accented syllables is determined by taking into account all their voiced phonemes (vowels and voiced consonants).

In the figures from the paragraphs 3.1, 3.2 and 3.3 some utterances (the wave and the F0 curve) are presented in order to exemplify different intonational variants. On the F0 curve we marked the tonal local events grouped into intonational units.

3.1 Intonational variants generated by the utterances of the oxytone text: *Ai văzut acest afiş?*

The intonational contours generated by the utterances of the oxytone text: "*Ai văzut acest afiş?*" are composed either by one, two or three intonational units.

One intonational unit and two prominent accents characterise the intonational variants from the Table 1: the first accent is on the verb and is generated by a weak tonal event characterised by a low-level topline. Instead a long duration of accented syllable makes prominent the corresponding word. The primary tonal accent is on the final word and generates the interrogative emphasis by a significant rising tone level.

During the first accent a low level limits the tone variation and thus it is not a prominent tonal event. It becomes prominent by increasing its length and the energy. The second accent is a primary tonal event H* (AT7, LS50) or L+H*(TM23, TM30).

Table 1. The "one unit" intonational variants corresponding to the utterances of the text $(Ai\ văzut\ acest\ afiş?)_A$

Intonational unit				Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	
AT7	**(Ai văzut)** acest **afiş**	H* (HIF0=L) $H^{(*)}$ H%	0.190 0.110	any word
LS50	**(Ai văzut)** acest **afiş**	H* (HIF0=L) $H^{(*)}$ H%	0.170 0.090	văzut
TM23	**(Ai văzut)** acest **afiş**	H* (HIF0=L) L+$H^{(*)}$ H%	0.160 0.140	văzut
TM30	**Ai văzut** acest **(afiş)**	H* (HIF0=L) L+$H^{(*)}$ H%	0.120 0.160	afiş

Figures 1 and 2 present the utterances TM23 and LS50. The last accent into utterance TM23 (Figure 1) is stronger (L+H* type) than the corresponding accent of H* type in utterance LS50 (Figure 2). In consequence, the first accent on verb is more prominent, by its duration and energy, in the case of utterances LS50 and AT7, than in the utterances TM23 and TM30.

Dividing the text into two intonational units it is justified by its syntactic structure: *verb+object* or its semantic structure: *thema* (the verb) + *rhema* (object). If both, the thema and rhema are focused, two accent units result. The most prominent accent from those, which generate the focuses, becomes the nuclear accent. In LS46 the nuclear accent is on the object, while in AT10, it is on the verb. The second primary accent is considered post or prenuclear, in respect to the nuclear accent position. The two units can be either in an intonational structure of type "A - A" or "A - B" (Table 2).

The F0 contour of the LS42 utterance is composed by a first unit within which it is accented the verb with an L+H* accent type (Figure 3). The interrogative emphasis is generated within the second unit, on the final word, afiş, by the tonal contrast of the tone sequence L+H* ˆH%. The duration of the accented syllable is of 0.107 msec and the accent becomes more prominent.

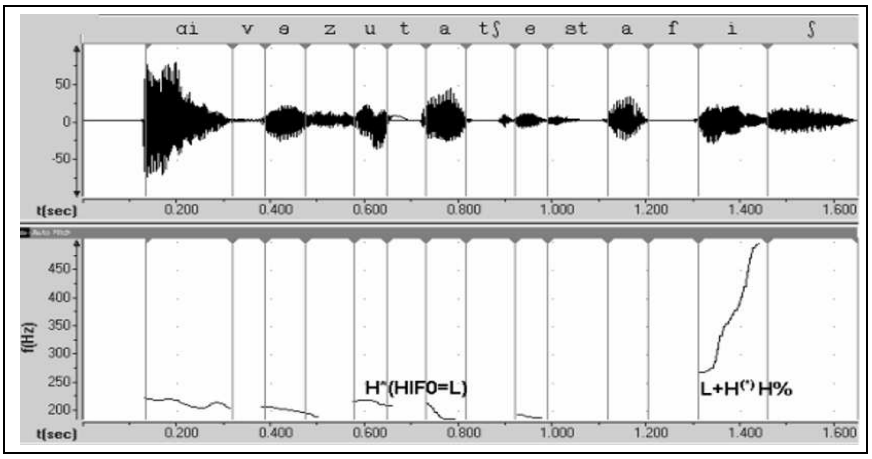

Figure 1. TM23: An utterance of the text *(Ai văzut acest afiş)*$_A$? with the melodic contour composed of one intonational unit

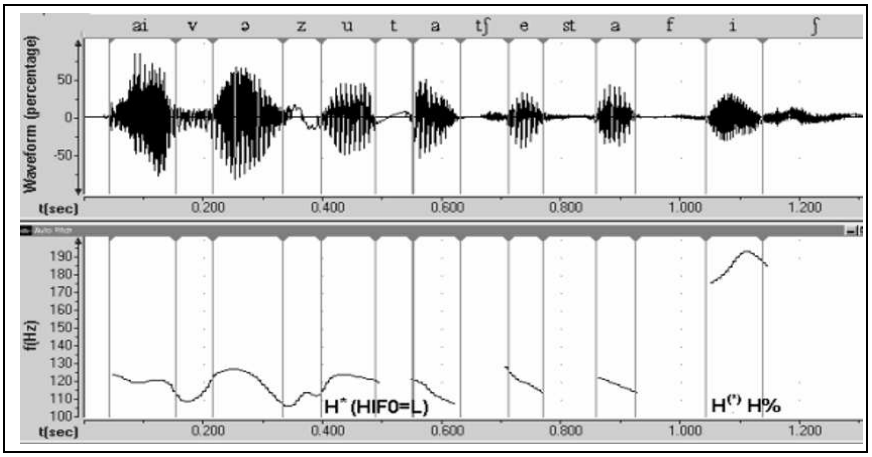

Figure 2. LS50: An utterance of the text (*Ai văzut acest afiş)*$_A$? with the melodic contour composed of one intonational unit

Table 2. The "two units" intonational variants corresponding to the utterances of the text *(Ai văzut)(acest afiş?)*

Intonational unit 1				Intonational unit 2			Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]	
LM 42	Ai văzut	L+H$^{(*)}$	0.146	**acest afiş**	^H* L$^{(*)}$L%	0.088 0.086	văzut
AT 10	**(Ai văzut)**	L+H$^{(*)}$	0.200	**acest afiş**	!H* ^H$^{(*)}$ H%	0.098	văzut
IM38, IM41	**(Ai văzut)**	L+H$^{(*)}$	0.260	**acest afiş**	^H* L$^{(*)}$ L%	0.105 0.120	văzut
LS 42	**(Ai văzut)**	L+H$^{(*)}$	0.113	acest **afiş**	H* L+H$^{(*)}$ ^H%	0.066 0.107	any word
LM 38	**Ai văzut**	L* (HIF0=L)	0.120	**acest afiş**	H* H+L$^{(*)}$ L%	0.058 0.094	any word
IM30, 34	(Ai văzut)	H* (HIF0=L)	0.100	**acest afiş**	H* H+L$^{(*)}$ L%	0.045 0.120	any word
LM 46	Ai văzut	H* (HIF0=L)	0.120	**acest afiş**	H* L$^{(*)}$ L%	0.050 0.096	afiş

In the case of LM42 utterance (Figure 4) the accent of L+H* type, on the verb, generates the most prominent contrast *low-high* and in consequence, the interrogative emphasis. The second unit of type "B" corresponds to the nominal group, with an accent of H* type on the first word and with an accent of L* type on the second.

In Table 2, the last three variants have a first intonational unit without the significant final increasing of pitch (approximately, a flat melodic segment, at a low level), corresponding to the verb *"ai văzut"*. The second unit of type "B" keeps together the words *"acest"* and *"afiş"* (Figure 5). The interrogative emphasis is generated by the tonal contrast between low average level of the first unit and the high average level within the second.

The intonational contour divided in three units is generated in the case of focusing the word "*acest*" placed in the middle of sentence. The modality to focus used by some speakers consists in isolating the word from the left one and from the right one, into a separated unit, within which it is stressed. The intonational contour with three units is generated in the case of jerky utterances, too.

In Table 3 the intonational variants divided in three intonational units are presented and Figure 6 illustrates one of them.

Table 3. The "three units" intonational variants corresponding to the utterances of the text *(Ai văzut)$_A$ (acest)$_A$ (afiş)$_A$*?

Intonational unit 1				**Intonational unit 2**			**Intonational unit 3**		
Utt. code	Text	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]
LS 41	**Ai *văzut***	L+H$^{(*)}$	0.142	**acest**	H$^{(*)}$	0.078	afiş	L$^{(*)}$ H%	0.130
LM 41	**Ai văzut**	L+H$^{(*)}$	0.150	**acest**	ˆH+!H	0.055	afiş	L$^{(*)}$ H%	0.110
LS 54	**Ai văzut**	L+H$^{(*)}$	0.140	**(acest)**	L*+H H-	0.120	afiş	L$^{(*)}$ H-%	0.140
LM 50	**Ai văzut**	L+H$^{(*)}$	0.120	**(acest)**	L*+H H-	0.091	afiş	L*+H H-%	0.120
AT 14	**Ai văzut**	L+H$^{(*)}$	0.200	**(acest)**	ˆH$^{(*)}$	0.120	afiş	ˆH$^{(*)}$H%	0.120
LS 44	Ai văzut	H$^{(*)}$	0.092	**acest**	H$^{(*)}$	0.090	afiş	L$^{(*)}$ H%	0.130

In Figure 6 the LM50 utterance with semantic focus on the word "*acest*" is presented. All three tonal accents in the each unit are prominent but on this word, the length and the energy on accented syllables are higher than other accented syllables in the sentence. The interrogative emphasis (modal focus) is on the final word "*afiş*", and it is generated by a large variation of the F0 frequency between low level into the pitch accent L* and boundary tone H%.

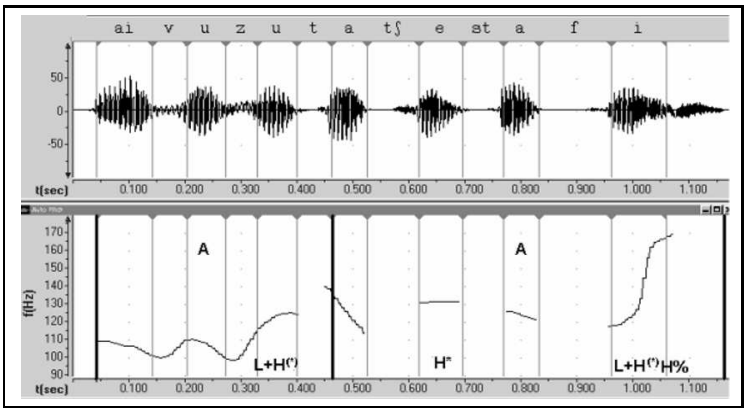

Figure 3. LS42 An utterance of the *text (Ai văzut)$_A$ (acest afiş)$_A$*? with the melodic contour composed of two intonational units

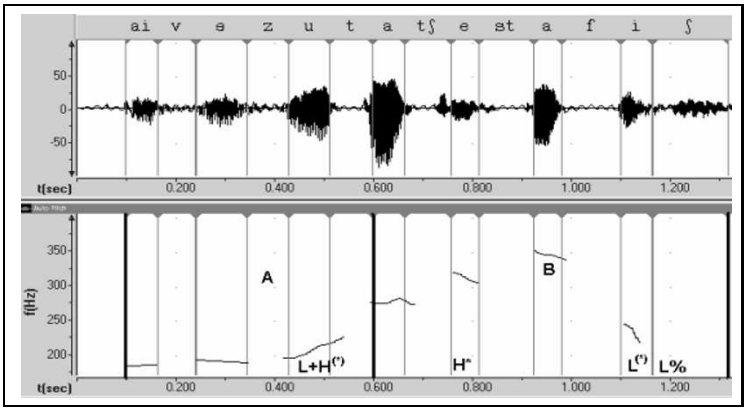

Figure 4. LM42: An utterance of the text *(Ai văzut)$_A$ (acest afiş)$_B$*? with the melodic contour composed of two intonational units

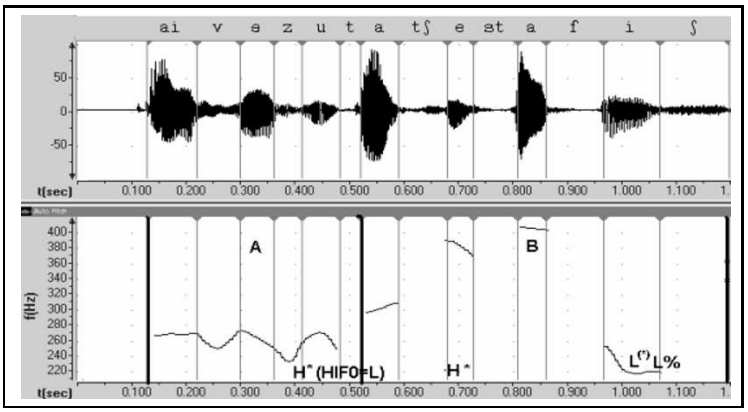

Figure 5. IM30: An utterance of the text $(Ai\ văzut)_A\ (acest\ afiş)_B$? with the melodic contour composed of two intonational units

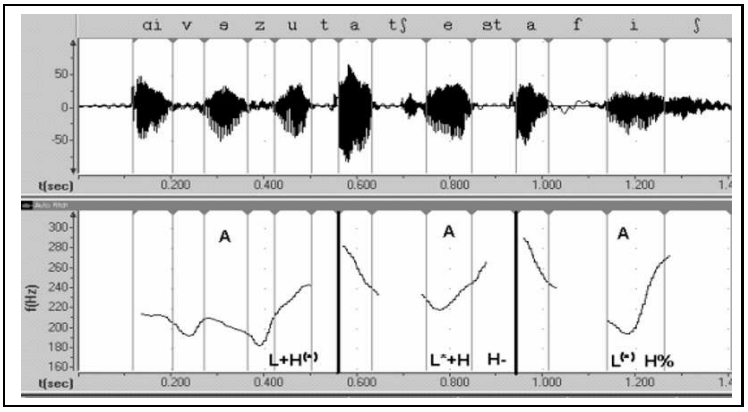

Figure 6. LM50: An utterance of the text $(Ai\ văzut)_A\ (acest\ _A\ (afiş)_A$? with melodic contour composed of three intonational units

3.2 Intonational variants generated by the utterances of the paroxytone text: *Ai văzut afişul acesta?*

The intonational variants corresponding to the text "*Ai văzut afişul acesta/ăsta?*" are presented in Tables 4-6. In Table 4 the variants with one-unit structure are presented. As in the utterance of the text "*Ai văzut acest afiş?*", there are two prominent accents: on verb (H* with HIF0=L) and on the final word (H* or L+H*). In the case of one unit structure, the final contour for this paroxytone case is ascendant-descendent, while in the oxytone case it is ascendant.

Table 4. The "one unit" intonational variants corresponding to the utterances of the text *(Ai văzut afişul acesta)*$_A$?

Intonational unit 1				Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	
TM3	**Ai văzut**	H* (HIF0=L)	0.170	any
	afişul		0.064	word
	acesta	L+H*L-L%	0.110	
AT1,	(**Ai văzut**)	H* (HIF0=L)	0.200	any
AT2	afişul		0.092	word
	acesta	H* !H-L%	0.120	
LS9	(**Ai văzut**)	H* (HIF0=L)	0.200	văzut
	afişul		0.059	
	acesta	H*L-%	0.075	
TM7	(**Ai văzut**)	H* (HIF0=L)	0.208	văzut
	afişul		0.606	
	acesta	H*L-L%	0.080	

In Figure 7 the F0 contour, as in all utterances of speaker TM, has a tone variation limited by a low level (200 Hz) from the beginning until the last accented syllable. During the last syllable one observe a tone variation in a large range (250 Hz) that generates the interrogative emphasis. After the last pitch accents the tone falls and returns to the low level.

In Figure 8, the intonational contour generated by a male speaker is characterised by a less prominent accent during the last accented

syllable (a pitch accent of type H*) but it generates the interrogative emphasis as in the case of speaker TM.

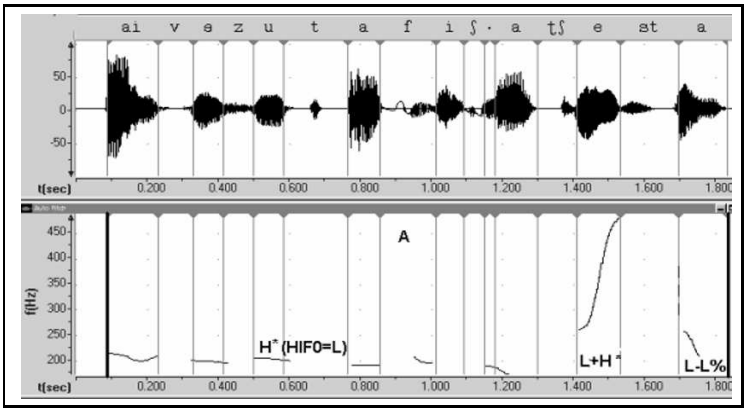

Figure 7. TM3: An utterance of the text *(Ai văzut afişul acesta)* $_A$? with melodic contour composed of one intonational unit

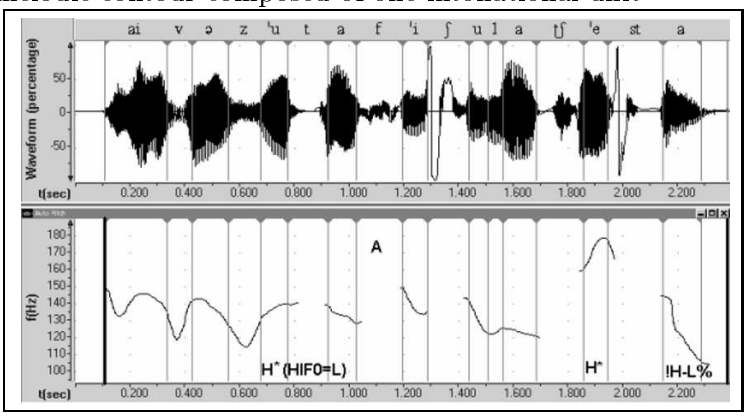

Figure 8. AT1: An utterance of the text *(Ai văzut afişul acesta)* $_A$? with melodic contour composed of one intonational unit

In Table 5 the intonations composed by two units are presented: the verb accented in the first unit and the nominal group in a second unit. The intonational structure and the position of interrogative emphasis divide them in two main types:

- one variant with the intonational structure of type "A"-"A" and the interrogative emphasis in final position, generated by the tonal contrast within the pitch accent of last word "*acesta*" (LS16);
- other variants, with the intonational structure of type "A"-"B" and the interogative emphasis in nonfinal position, generated by a broad phenomena: the tonal contrast between low average level of the first unit and the high average level within the second.

In this first case the emphasis is generated within the final word "*acesta*" based on the tonal contrast between the L* tone accent and phrase and boundary tones, H-H% (Figure 9). The final contour of the intonation is a descendent-ascendant one.

The last five variants in Table 5 represent the second type. They are characterized by a first intonational unit without the significant final increasing of pitch (approximately, a flat melodic segment, at a low level), corresponding to the verb *"ai văzut"*. The interrogative emphasis is in non-final position and is produced during the both words *"afişul acesta"*. The second type is illustrated in Figure 10 by the utterance LM3.

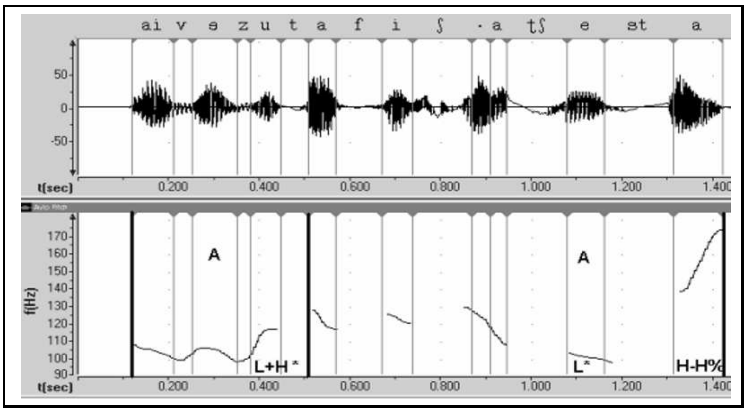

Figure 9. LS16: An utterance of the text *(Ai văzut)$_A$(afişul acesta)$_A$*? with melodic contour composed by two intonational units

Table 5. The "two units" intonational variants corresponding to the utterances of the text *(Ai văzut) (afişul acesta)?*

Intonational unit 1				**Intonational unit 2**			Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]	
LS 16	**Ai văzut**	L+H$^{(*)}$	0.110	afişul **acesta**	H* L*^H-H%	0.048 0.096	acesta
LM 3	(**Ai văzut**)	H$^{(*)}$ (HIF0=L)	0.126	afişu' **ăsta**	H* H+L$^{(*)}$L%	0.048 0.085	any word
IM 22	Ai văzut	L$^{(*)}$ (HIF0=L)	0.211	**afişul ăsta**	H* H+L$^{(*)}$!H%	0.077 0.1324	any word
LM 10	Ai văzut	L+H$^{(*)}$ (HIF0=L)	0.131	**afişu ăsta**	H* H+L$^{(*)}$L%	0.038 0.100	văzut
IM 23	(**Ai văzut**)	L+H$^{(*)}$ (HIF0=L)	0.315	**afişul ăsta**	H* H+L*H%	0.098 0.135	văzut
IM 28	**Ai văzut**	L+H$^{(*)}$ (HIF0=L)	0.190	**afişu' (ăsta)**	H* H+L*!H%	0.110 0.220	ăsta

The intonational variants in Table 6 are composed by three units generated by the fact that each word is uttered in a separate intermediate phrase with a pause between them. It is a hesitation of the speaker and it isn't an intention to focus on the word being in a middle position (Figure 11).

From a comparative analysis between oxytone and paroxytone cases of utterances of the same speaker, we conclude that speakers keep unchanged the melodic contour, except the final contour that changes like in the following:

Table 6. The "three units" intonational variants of the utterance *(Ai văzut)$_A$(afişul)$_A$ (acesta)$_A$*?

Intonational unit 1				**Intonational unit 2**			**Intonational unit 3**		
Utt. code	**Text**	Tone sequence	Duration [sec.]	**Text**	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]
LS 17	Ai văzut	H*	0.100	**(afişul)**	!H*	0.120	acesta	L*!H-H%	0.100
LS 5	**(Ai văzut)**	L+H*	0.164	afişul	!H*	0.055	acesta	L*!H-H%	0.0750

- from ascendant to ascendant-descendent one, at TM or AT speaker,
- from ascendant to descendent-ascendant at LS.

The IM and LM speaker, from south-west Transilvania region, don't change the final melodic contour.

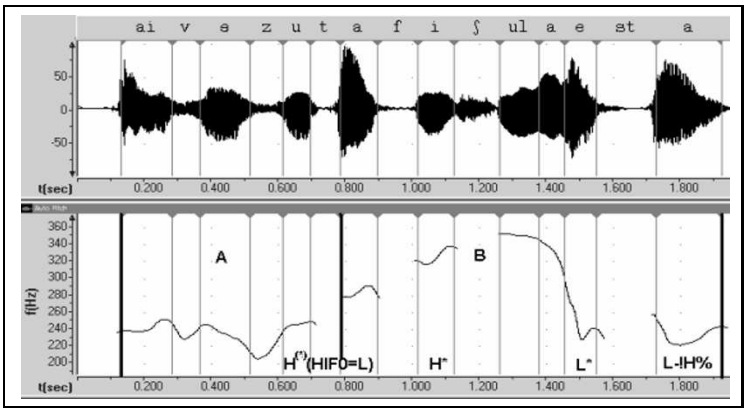

Figure 10. IM28: An utterance of the text *(Ai văzut)$_A$ (afişul ăsta)$_B$*? with melodic contour composed by two intonational units

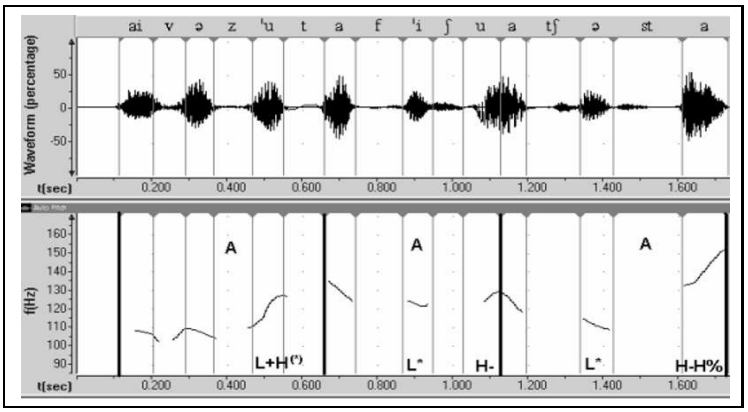

Figure 11. LS5: An utterance of the text *(Ai văzut)$_A$(afişul)$_A$ (acesta)$_A$*? with melodic contour composed of three intonational units

3.3 Intonational variants generated by the utterances of the proparoxytone text: *Ai văzut regele?*

In Table 7 and in Figure 12 one utterance of speaker LS with the intonational contour consisting of one unit is presented. The verb is uttered with a low limit tone and the emphasis is generated on the word "*regele*" beginning with the accented syllable. The highest tone of the final contour corresponds to the phrase accent H-.

Table 7. The "one unit" intonational variants corresponding to the utterances of the text *(Ai văzut regele?)$_A$*

Intonational unit 1				Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	
LS 23	*Ai* **văzu regele**	H* (HIF0=L) L+H* ˆH-L%	0.120 0.133	any word

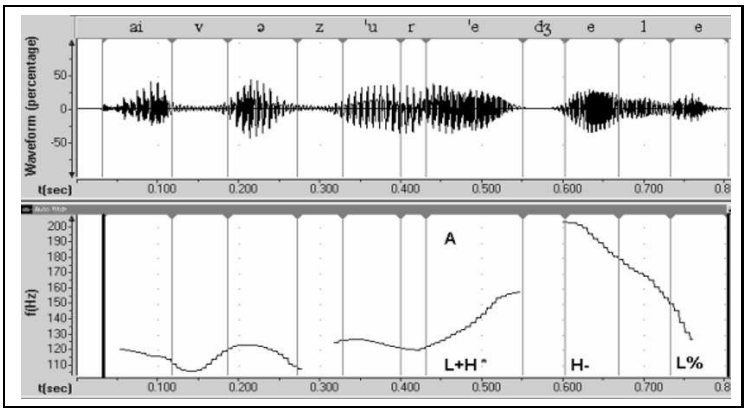

Figure 12. LS23: An utterance of the text *(Ai văzut regele?)*$_A$ with a melodic contour composed of one intonational unit

Table 8 contains several interesting variants for intonation of this proparoxyton text. From the speaker LS we extracted two intonation contours. In LS26 utterance, the intonational contour can be defined by a unit sequence "A-B" (Figure 13). Within the unit "A" a large duration and a high level of energy accent the verb. Instead the object is stressed by a prominent pitch accents of type L+H*. The final contour is ascendant-descendent and the interrogative emphasis is on the object. In LS37 utterance the verb is stressed by an increasing accent of type L+H*, while the object – by a decreasing accent of type L* (Figure 14). The interrogative emphasis is on the word "*regele*", with the low segment on the first two syllables and the high prominence on the last unaccented syllable. The final contour is descendent-ascendant.

The speaker LM generates two intonational variants presented in Table 8. Both of them have an increasing accent L+H* on a verb and the emphasis on the object *"regele"* generated in the second unit. In LM 23 utterance (Figure 15), in the second unit of type B, the word is focused with an accent of type H*+L on stressed syllable, and a tone sequence L-L% on the next unaccented syllables. The final contour is

ascendant-descendent.

Table 8. The "two units" intonational variants corresponding to the utterances of the text *(Ai văzut)(regele?)*

Intonational unit 1				**Intonational unit 2**			Word with focus intention
Utterance code	Text	Tone sequence	Duration [sec.]	Text	Tone sequence	Duration [sec.]	
LM 23	Ai **văzu'**	L+H*	0.152	**(regele)**	H* L- L%	0.172	any word
LM 22	Ai **văzu'**	L+H*	0.160	**(regele)**	H+L* H-L%	0.157	any word
LS 26	**(Ai văzut)**	H* HIF0=L	0.337	**regele**	L+^H* H-L%	0.215	văzut
LM 34	**(Ai văzu')**	L+H*	0.158	**regele**	H* L-L%	0.131	văzut
AT 32	**(Ai văzut)**	L+H$^{(*)}$	0.250	**regele**	H* L-L%	0.220	văzut
LS 37	Ai văzut	L+H$^{(*)}$	0.129	**(regele)**	L* L- H%	0.138	regele
LM 27	*Ai* ***văzut***	L+H$^{(*)}$	0.171	***(regele)***	L* H- L%	0.155	regele
AT 36	*Ai* ***văzut***	L+H$^{(*)}$	0.197	***(regele)***	L+H* H-L%	0.270	regele

In LM 22 utterance (Figure 16), in the second unit of type A, the word is focused by an accent of type L* on stressed syllable followed by a tone sequence H-L% on the next unaccented syllables. The final contour is descendent-ascendant-descendent.

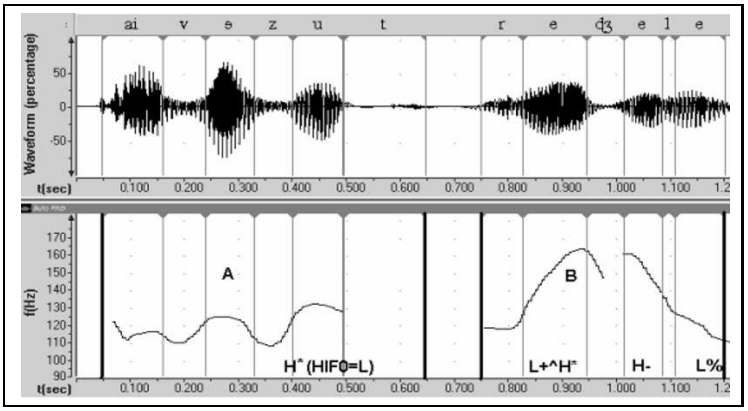

Figure 13. LS26: An utterance of the text *(Ai văzut)$_A$ (regele?)$_B$* with a melodic contour composed of two intonational units

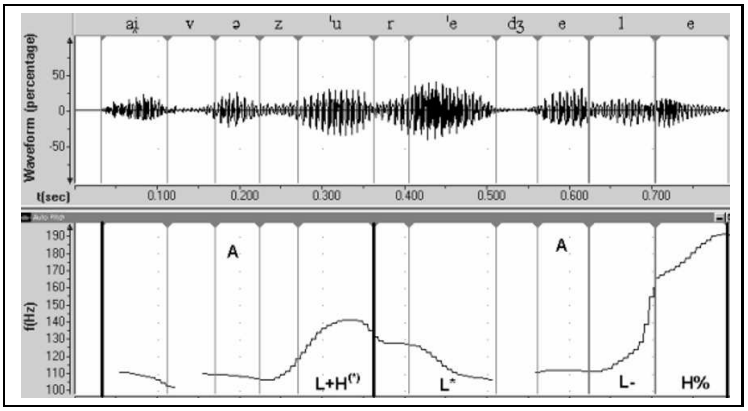

Figure 14. LS37: An utterance of the text *(Ai văzut)$_A$ (regele?)$_A$* with a melodic contour composed of two intonational units

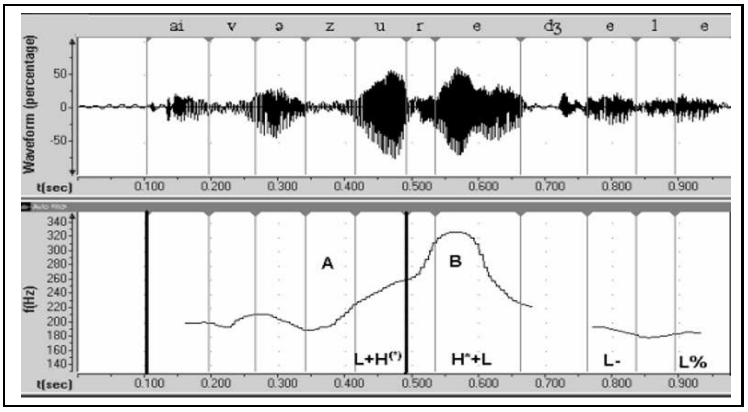

Figure 15. LM23: An utterance of the text *(Ai văzut)$_A$(regele?)$_B$* with melodic contour composed by two intonational units

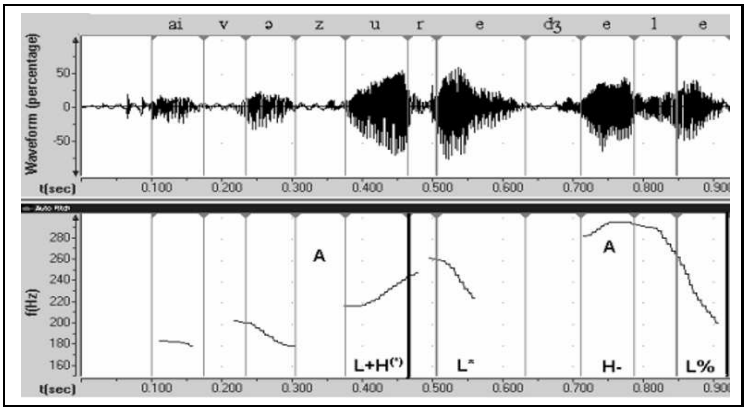

Figure 16. LM22: An utterance of the text (Ai văzut)$_A$(regele?)$_A$ with a melodic contour composed of two intonational units

4 Conclusions

From this study results that the intonation in interrogative utterances is characterised by an interrogative emphasis generated by a tonal contrast within one of the intonational units that compose the F0 contour or between the average tone of two consecutive units.

In first case the contrast implies two syllables (Figure 3) or only one accented syllable of the same word (Figure 13). In the second case the contrast between the average tone of two consecutive units is illustrated in Figure 11.

Within one intonational unit there is a prominent accent, generated by tonal variation or by duration and energy. The nuclear accent of the sentence is the most prominent accent from all units. The interrogative emphasis needs a tonal contrast in all cases. The most prominent accents that can generate an emphasis are the bitonal ones L+H* or L*+H, or H+L*. Both monotone and bitonal accents become prominent by increasing their energy and duration.

Some general rules for Yes-No questions speech synthesis, without any indication of intonation, are the following:

- Implicit accents must be provided on verb and final word. The final accent must have a more prominent tonal contrast and the verb can be accented specially by duration and energy;
- One or two intonational units can compose the intonational contour in one of intonational variants presented in this paper;
- The indications for semantic focus must drive the synthesis in increasing prominence of implicit accents or in generating others, by dividing the F0 contour in many intonational units.

The author D.R. Ladd characterised in [2] the Romanian intonation for the Yes-No questions by the sequence L* HL in one variant and H L*HL in the second. By "*" he marked the nuclear accent. He illustrates his presentation by the neutral utterances of the text "*Ai văzut afişul acesta?*" and "*Ai văzut regele?*"

The first sequence L* HL corresponds to the utterances characterised by one-unit intonational contour (Table 1, 4). The tones from D.R.Ladd's description correspond to our interpretation in following manner:

- the tone L* corresponds to the low average tone that limits the tone variations during pitch accents (H*), before the last accented syllable (in ToBI annotation system, parameter HIF0=L);
- the tone H corresponds to the increasing accent (H* or L+H*) on last accented syllable;
- the last tone L corresponds to the boundary tone L% only in the paroxytone cases.

Comparing the second sequence with the intonational variants resulted from our analysis we conclude that this corresponds to the variant composed by two intonational units: in the first, the verb is accented by a L+H* type accent and in the second, the emphasis is generated by a sequence L*H-H%. The last tone "L" from D. R. Ladd's description can be found only in proparoxytone text cases, uttered by speakers from south-west Transilvania and Banat (like as LM22 utterance, here illustrated in Figure 16).

The conclusions of Yes-No question analysis are important for linguistic studies and in Romanian speech synthesis.

References

[1] Laurenţia Dascălu-Jinga, *Melodia vorbirii în limba română*, Editura Univers enciclopedic, Bucureşti, 2001.

[2] D.R. Ladd, *Intonational Phonology*, Cambridge University Press, 1996

[3] M.Beckman, G.Ayers, *Guidelines for ToBI Labelling* (version 3, March 1997), http://www.ling.ohio-state.edu/research/phonetics/E_ToBI, 1997.

[4] V.Apopei, D.Jitca, H.N.Teodorescu, *Implementation of stress and emotion rendering rules in synthesized speech*, Trends in speech technology, Romanian Academy Publishing, pg. 67–72, 2005.

[5] V.Apopei, D.Jitca, *Romanian Intonational Annotation Based on Tone Sequence Model*, SASM 2005, Iaşi, Romania, May 5-7, 2005.

[6] A.Turculeţ, V.Apopei, D.Jitcă, *Aspecte ale intonaţiei propoziţiilor interogative totale cu structura* VO(Adj), Anuar de Lingvistică şi Istorie Literară, Iaşi.

V.Apopei, D.Jitcă, A.Turculeţ,

Received February 1, 2006

Institute for Computer Science,
Romanian Academy, Iaşi Branch
B-dul Carol I, nr. 22A, 6600 IAŞI, ROMANIA
E–mail: *vapopei@iit.tuiasi.ro*