

Bistra Andreeva, Institute of Phonetics, University of Saarland
Tania Avgustinova, Department of Computational Linguistics, University of Saarland
William J. Barry, Institute of Phonetics, University of Saarland

Link-Associated and Focus-Associated Accent Patterns in Bulgarian

1 Introduction

On the basis of experimental data, we investigate how the pre-nuclear and nuclear accents are realised in Bulgarian with respect to the type of utterance – question or statement – and the information structure. We use the terminology of the information packaging approach (e.g., Vallduví and Engdahl 1995, 1996) in assuming a partitioning of the utterance into a focus component (rheme) and a ground component.

Within the ground component, a further distinction is made between *link* (topic) and *tail*. The former is a prominent sub-part of the ground that establishes connections to the preceding text / discourse, while the latter functions as a kind of 'filling' specifying some redundant information and, therefore, tends to be 'intonationally elliptical'. In general, the ground contains 'known' data and does not contribute anything further to the context. It is rather context-dependent and presupposed. In the information structure of utterances, the focus can be viewed as a default information state, since the primary goal of any communication is to transmit some information. The focus pushes the communication forward by virtue of being the (new) information that is being added to the context, but the focus itself is context-independent. In this sense, it is the unmarked case. Such neutral utterances are prototypically associated with canonical configurations, both syntactically and prosodically.

Important factors in the realisation of the information structure in Bulgarian utterances¹ are:

- the interaction – association or dissociation – between the actual word order and the canonical obliqueness-based ordering of grammatical functions;
- the well developed mechanism of clitic replication;
- the existence or lack of a morphologically realised category of definiteness / determinedness;
- the plastic / malleable intonation.

We are primarily interested in the contribution of the intonation here. Therefore, in the design of the experiment, we fixed the word order to reflect the canonical one, e.g., *subject < verb < direct object < indirect object < oblique*. This increased the role of the intonation as an information-structuring factor, allowing us to observe more independently the realisation of the link-associated and focus-associated accent patterns in Bulgarian statements and questions. In particular, we considered the following types of utterances in our experiment:

- *non-contrastive broad focus* utterances: these are, of course, potentially ambiguous due to focus-projection possibilities;

¹ Cf. Avgustinova 1997 (pp. 83-155) for details.

- *non-contrastive ground – narrow focus* utterances: here the combinations link-link and link-focus can be observed (in the former combination, the accent patterns turned out to be different for statements and questions);
- *contrastive broad focus* utterances²: similar to the preceding two types of utterances with respect to the information structure; however, the tonal space³, i.e. *inclination* (up-step of the nucleus) is a crucial factor;
- *non-contrastive multiple (double) focus* utterances: here all focus instances have a narrow interpretation;
- *contrastive ground – narrow focus* utterances: here the opposition between the link-link accent pattern and the link-focus accent pattern is neutralised.

2 Experiment Design

The subjects for the experiment were four persons from Sofia (3 female and 1 male, aged 25-45 years). The following 8 test sentences (TS) were recorded three times per task in a random order in a sound treated studio at the Institute of Phonetics (University of Saarland):

1. {/'vče ra/} {/'ma ma/} {/'ma za/} {/'ma sa ta/}⁴
yesterday mama painted the-table
'Yesterday mum painted the table.'
2. {/'vče ra/} {/'ma ma/} {/po 'ma ga/} {/na/ /'ma re to/}
yesterday mama helped to Mareto
'Yesterday mum helped Mareto.'
3. {/'vče ra/} {/'ma ma/} {/ni/ /po 'ma ga/} {/po/ /gra 'ma ti ka/}
yesterday mama us helped in grammar
'Yesterday mum helped us in grammar.'
4. {/'u tre/} {/'ma ma/} {/šte/ /ni/ /po 'ma ga/} {/po/ /ma te 'ma ti ka/}
tomorrow mama will us help in mathematics
'Tomorrow mum will help us in grammar.'

² Here the distinction between a *contrastive broad focus* and a *non-contrastive multiple focus* is mainly international, inasmuch as the phenomenon of contrasting – being, certainly, contextually determined – is orthogonal with respect to the phenomenon of information structuring in terms of ground and focus.

³ Cf. Ladd 1996:73 for this concept.

⁴ ' – word-stress
/word/
{prosodic word}

5. {/'vč̥e ra/} {'m̥a ma/ /li/} {/vi/ /po 'm̥a ga/} {'po/ /gra 'm̥a ti ka/}?
 yesterday mama Q⁵ you helped in grammar
 'Is it mum who helped you yesterday in grammar?'
6. {/'vč̥e ra/} {'m̥a re to/} {'po 'm̥a ga/ /li/} {'po/ /gra 'm̥a ti ka/}?
 yesterday Mareto helped Q in grammar
 'Did Mareto helped in grammar yesterday?'
7. {/'vč̥e ra/} {'m̥a ma/} {'vi/ /po 'm̥a ga/} {'po/ /gra 'm̥a ti ka/ /li/}?
 yesterday mama you helped in grammar Q
 'Is it in grammar that mum helped you yesterday?'
 'Yesterday mum helped you in grammar, didn't she?'
8. {/'vč̥e ra/} {'m̥a re to/} {'po 'm̥a ga/} {'po/ /gra 'm̥a ti ka/}?
 yesterday Mareto helped in grammar
 'Yesterday Mareto helped in grammar?'

There are one to four unaccented syllables between the four accentable syllables. We do not analyse the very first word (*vč̥era* 'yesterday' or *utre* 'tomorrow') but take it as a filling material preceding the second accentable syllable (word-initial in *mama* 'mum' or *Mareto*). The accentable syllables in the relevant material have the same segmental structure (maximally sonorant) in order to avoid micro-prosodic effects.

2.1 Corpus I (Statements)

No explicit instructions were given to the subjects but rather their reaction was driven by the context, i.e. implicitly. We used three methods for inducing the respective focus types:

Method 1: lists of test sentences 1-4 each occurring three times in a random order

- a) reading without context (resulting in a non-contrastive broad focus);

[broad; -contrastive]	(...) Focus
-----------------------	-------------

- b) reply to a wh-query (by the instructor) about the last word, resulting in a non-contrastive narrow focus in the final position;

[narrow; -contrastive]	(...) Link Focus
------------------------	------------------

Method 2: embedding the test sentence 3 in dialogue sequences

- a) a wh-query uttered by the instructor about the initial (with respect to the relevant material), middle or final word with test sentence 3 produced as a reply by the subjects, resulting in a non-contrastive narrow focus in the respective position;

⁵ Q – interrogative particle *li*

[narrow;-contrastive]	(...)	Focus	Tail
		(...)	Link Focus Tail
		(...)	Link Focus

- b) a wh-query about the entire relevant material, resulting in a non-contrastive broad focus;

[broad;-contrastive]	(...)	Focus
----------------------	-------	-------

- c) a wh-query about the first and the last word of the relevant material, resulting in a non-contrastive double focus.

[double;-contrastive]	(...)	Focus	Link/Tail	Fo- cus
-----------------------	-------	-------	-----------	------------

Method 3: embedding of the test sentence 3 in dialogues consisting of correcting sequences

- a) a yes/no query uttered by the instructor about the initial (with respect to the relevant material), middle or final word with the test sentence 3 produced as a correcting reply by the subjects (resulting in a contrastive narrow focus);

[narrow;+contrastive]	(...)	Focus	Tail
		(...)	Link Focus Tail
		(...)	Link Focus

- b) a yes/no query about the entire relevant material and the test sentence 3 produced as a correcting reply by the subjects (resulting in a contrastive broad focus).

[broad;+contrastive]	(...)	Focus
----------------------	-------	-------

2.2 Corpus 2 (Yes/No Questions)

This corpus contains both yes-no questions with the interrogative particle and questions marked only by intonation. Typically *li* belongs to the focal segment. When it occurs in sentence-final position an ambiguous focus interpretation arises (either broad or narrow). A list of test sentences 5-7 (*li*-questions) each occurring three times in a random order was read by the subjects, with the *li*-particle following the initial, the middle and the final part of the relevant material. The test sentence 8 was produced 3 times as a reaction to a described situational context, which was constructed to induce focus on the initial, the middle or the final part of the relevant material.

3 Discussion of the Results

3.1 Focus-Associated Accent Patterns in Statements

The framework adopted in the present study is the autosegmental-metrical model of intonational phonology (Pierrehumbert, 1980)⁶. The phonological correlate of focus is at least one pitch accent which is realised on one of the syllables in the focussed material.

(1) non-contrastive narrow focus: H*

Statements with non-contrastive narrow focus have a falling nuclear pitch movement, i.e. a H(igh) accent followed by a L(ow) boundary tone. Within an autosegmental approach, we may analyse this focus-associated accent as H*, where the H target is a local peak aligned around the beginning of the accented syllable (cf. Fig. 1). The tonal movement from the high target to the low boundary target is not phonologically specified. It is realised as a linear interpolation, i.e. a transition between tonal targets.

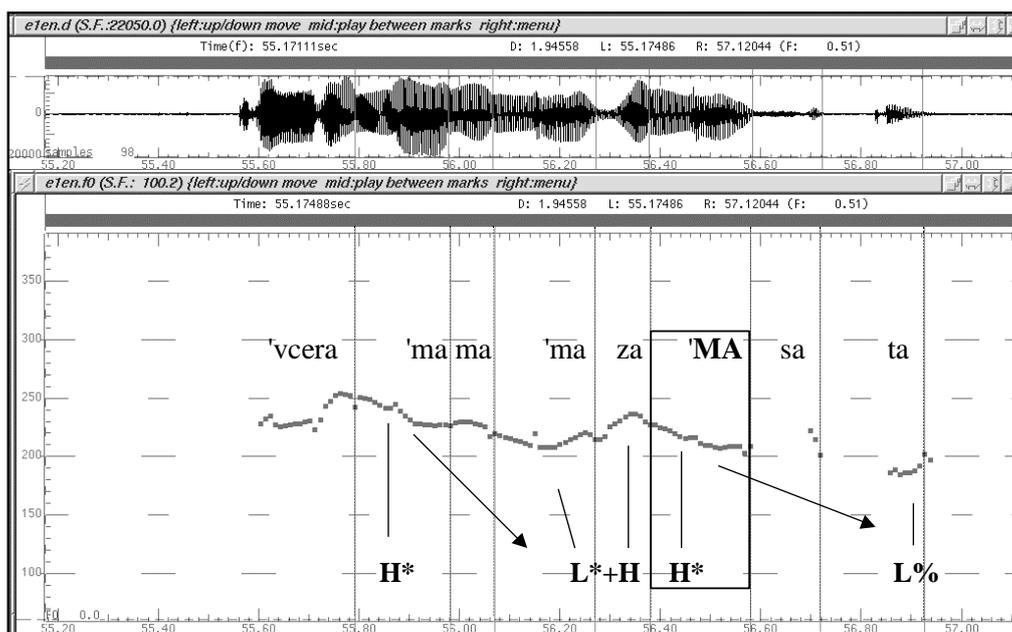


Figure 1: Speaker EK (female); TS 1
Context: *Kakvo maza mama včera?* ('What did mum paint yesterday?')

⁶ Previous work in Bulgarian sentence intonation and information structure can not be compared directly with the approach used here. That research either uses the four level American structuralist tradition of intonation description (e.g., Penčev, 1980) or has, despite some observational similarities with the present work, not captured the observations within a formal system (e.g., Miševa 1991).

(2) contrastive narrow focus: H*>

Following Ladd (1983) and Gussenhoven (1984) we assume that the tonal composition of the underlying pattern for narrow [\pm contrastive] focus is the same, namely a monotonal H* accent. The accent has different phonetic realisation, though. The alignment⁷ of the high tone with the accented syllable is different for contrastive versus non-contrastive focus. With non-contrastive focus the peak is aligned early in the syllable, whereas with contrastive narrow focus the peak is aligned later (cf. Fig.2), for some of the subjects into the following unstressed syllable. This modification of the position of the peak (delayed peak) signals that the utterance is not neutral.

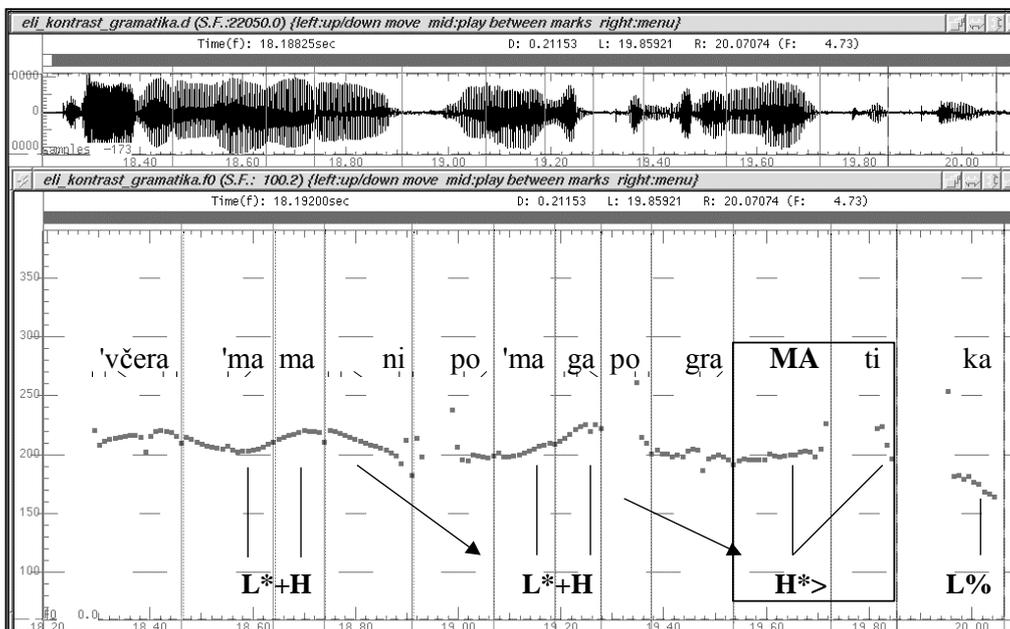


Figure 2: Speaker EK (female); TS 3
Context: *Včera mama po matematika li vi pomaga?*
(‘Was it in maths mum helped you yesterday?’ – No, in grammar.)

(3) non-contrastive broad focus

If the analysis is based solely on the accent pattern, the focus interpretation of the utterance is potentially ambiguous between narrow focus and broad focus readings. For narrow [\pm contrastive] focus the accented syllable of the focussed word carries the focus-associated accent pattern. When complex phrases are focussed, i.e. for broad focus, language-specific focus projection rules define the focus exponent. The pitch accent falls on the stressed syllable of the

⁷ For the distinction between (abstract structural) association and (phonetic) alignment see Ladd (1996, p.55).

focus exponent, which becomes focussed. This means that a single pitch accent signals a complex focus domain in the case of focus projection.

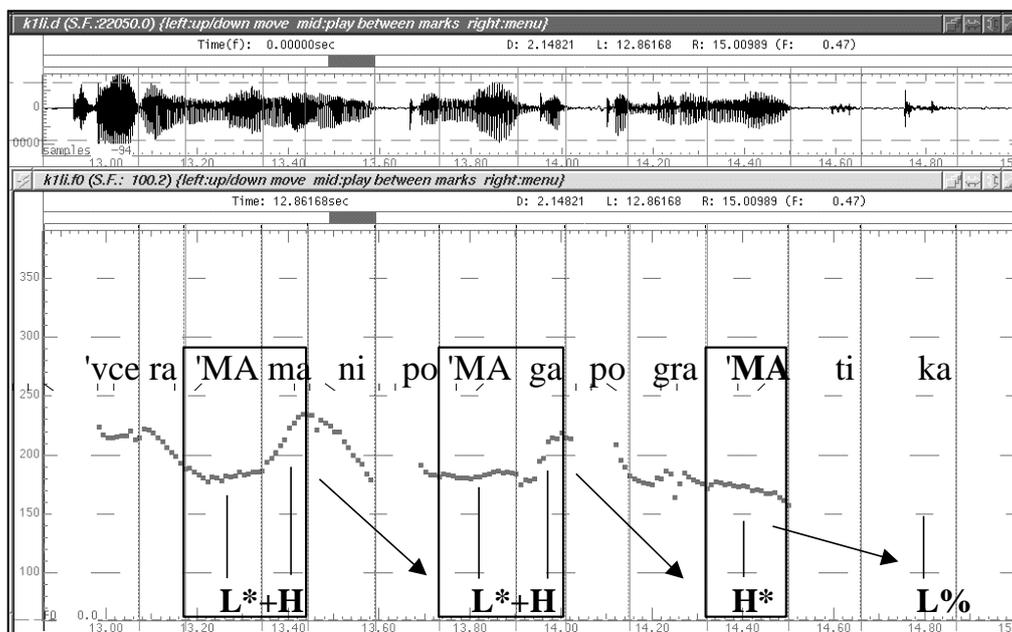


Figure 3: Speaker KP (female); TS 3
Context: *Razkaži za včera.* ('Tell me about yesterday.')

Analysis of the corpus shows that focus projection is avoided for complex focus domains. Instead, the subjects divide the complex focus domain into several accent domains, each of which get a pitch accent. There are 3 accent domains in Figs.3 and 4. Speaker KP (cf. Fig.3) realises the first and second accents as L*+H (low target on the first ma in mama and ma in pomaga and high target on the post-accentual syllable). Speaker EK (cf. Fig.4) realises the first accent differently, namely as H* (local peak at the beginning of the accented syllable), the second as L*+H, like the first speaker. In the case of several accent domains, the rightmost pitch accent is the most prominent of the entire intonation phrase (final accent strengthening) and is realised as a H* by both speakers.

(4) contrastive broad focus

The realisation of the pitch accent patterns in the accent domains is the same as for non-contrastive accents, except that the nuclear pitch accent is realised with emphasis (see section 3.4). In this case the nucleus is higher in pitch than the pre-nuclear pitch accents (cf. Fig.5).

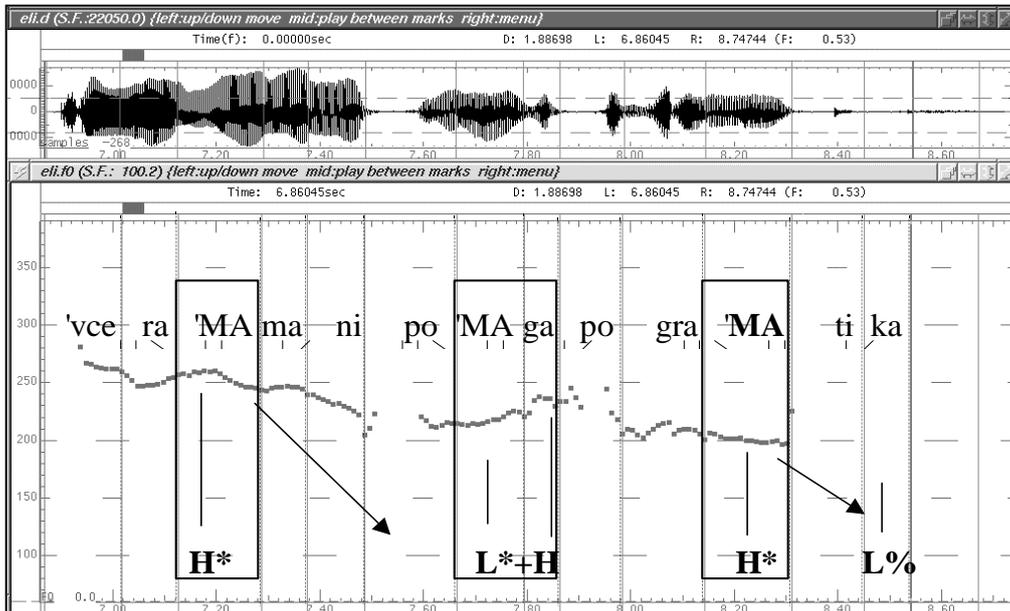


Figure 4: Speaker EK (female) TS 3
Context: *Razkaži za včera.* ('Tell me about yesterday.')

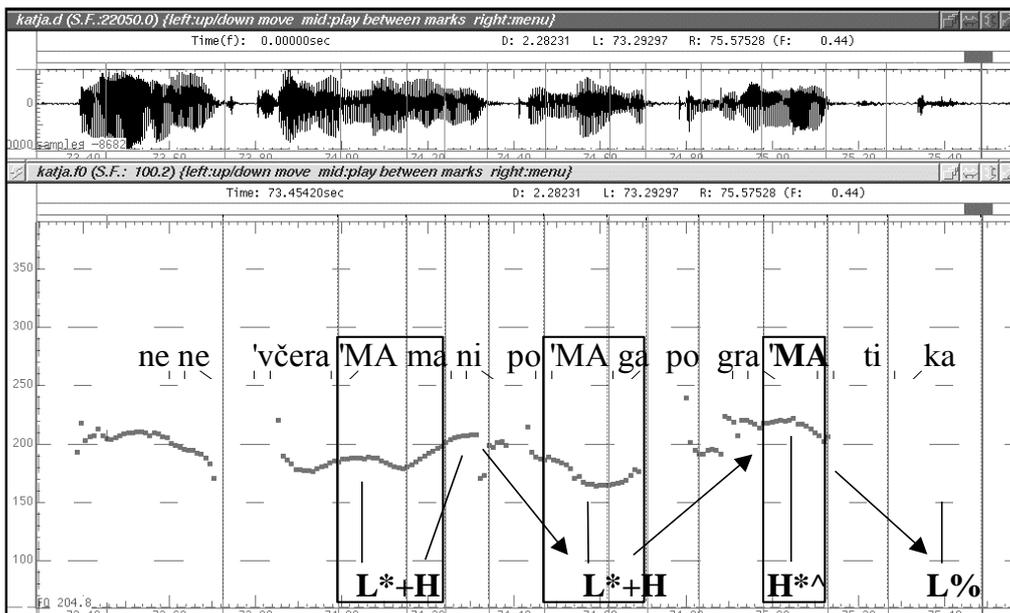


Figure 5: Speaker KP (female); TS 3
Context: *Včera baba prikazka li vi čete?* ('Was it a fairy tale granny read you yesterday?' – No, it was mum and she helped us in grammar.)

(5) non-contrastive double focus

A wh-query about the initial and final words in the utterance was used to elicit non-contrastive double focus. In Fig. 6 the realisation of the focus-associated accents is illustrated. The two foci are realised with two different pitch accent patterns. The first one is realised as L*+H (low target on accented syllable *ma* in *mama* and high target on the post-accentual syllable) and the second one as H* (high syllable *ma* in *gramatika*).

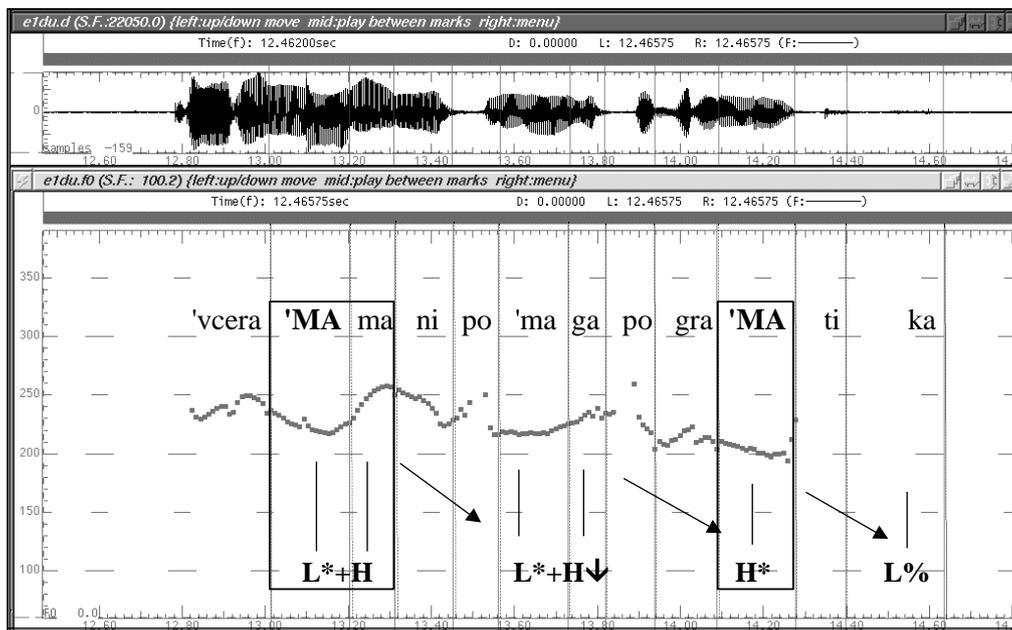


Figure 6: Speaker EK (female); TS 3
Context: *Koj po kakvo vi pomaga včera?* ('Who helped you in what yesterday?')

3.2 Focus-Associated Accent Patterns in Yes/No Questions

Like a number of other Eastern European languages, the tone sequence L*+H L% is the basic yes/no question intonation in Bulgarian, i.e. a low nuclear syllable followed by a pitch peak and a final low. However the precise association of the H pitch peak with the segmental string varies in specifiable ways, both between yes/no-question types and between prosodic contexts.

In li-questions the H tone is aligned closer to the prosodic-word boundary if there is enough segmental material after the focussed word, i.e. on the question particle (cf. Fig. 7).

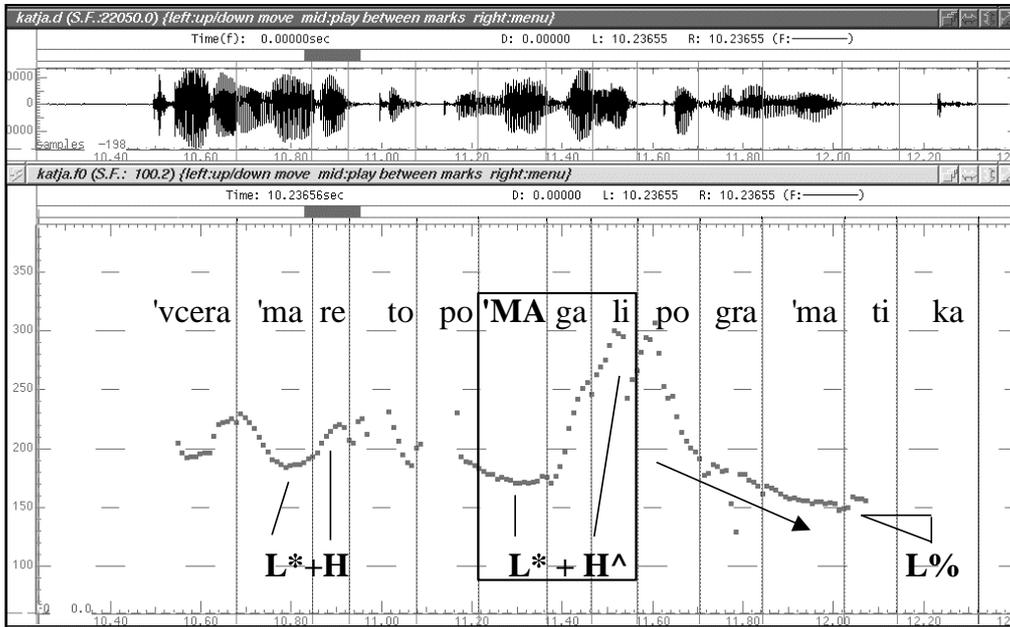


Figure 7: Speaker EK (female); TS 6

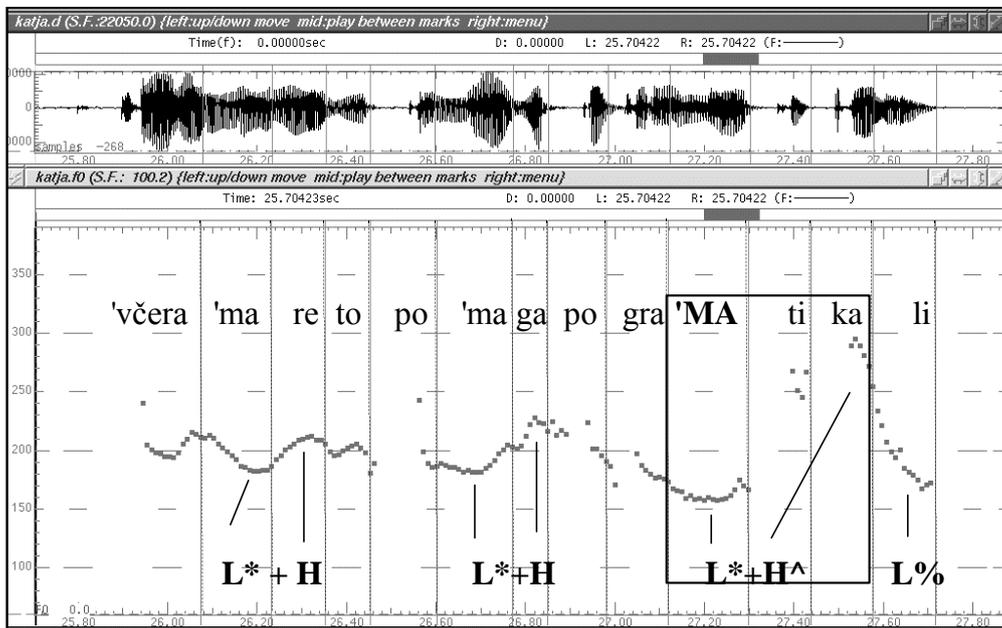


Figure 8: Speaker EK (female); TS 7

Otherwise the trailing tone of a bitonal accent is placed on the penultimate syllable of the prosodic word and the low boundary tone on the last syllable 'li' (cf. fig. 8).

In Yes/No questions marked only by intonation, the pitch accent association is realised differently. The start of the rise on 'MA (cf. Figs. 9 and 10) indicates that both the low and high targets of the pitch accent are associated with the accented syllable; the high target is also associated with the following syllable. In other words, the bitonal L^*+H accent is dumped on the accented syllable and, if there is enough segmental material, its trailing tone spreads onto the next syllable.

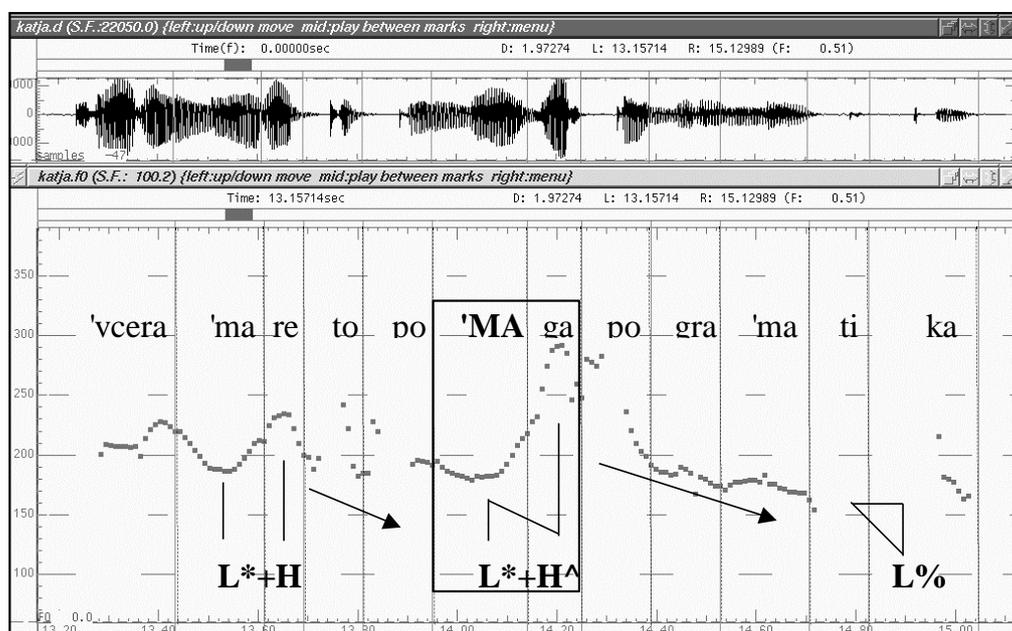


Figure 9: Speaker EK (female); TS 8

The diacritic '^' marks the greater pitch interval between the low and high tones. This modification is called "raised peak" (Ladd 1983).

3.3 Link-Associated Accent Patterns

Constituents can build an accent domain, independent of whether they belong to the focus or to the background. In our database we found two possible link-associated accent patterns, depending on whether the accent is followed by another link accent or by a focus accent.

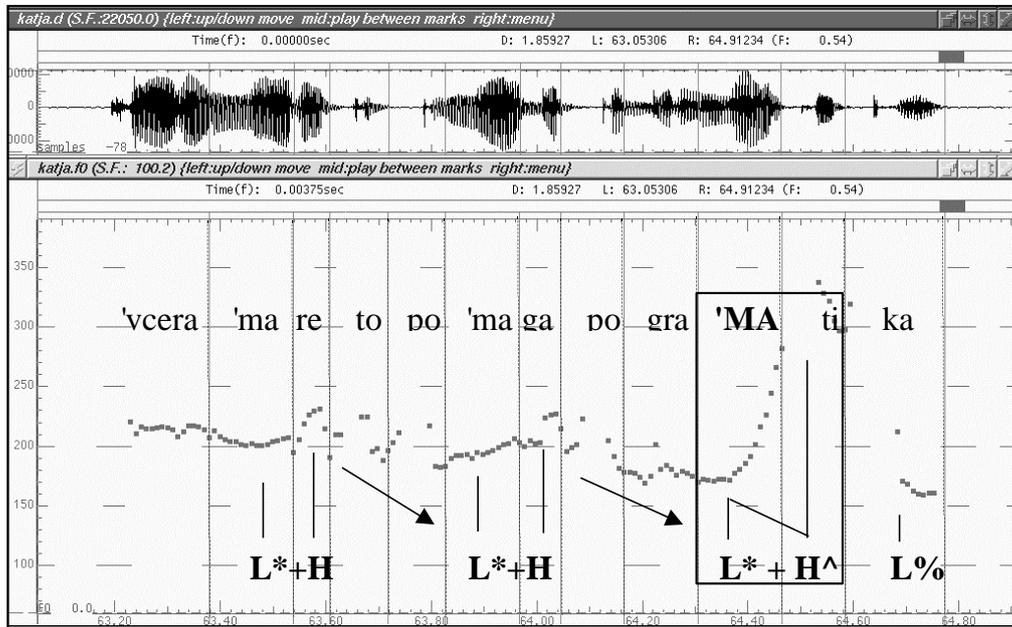


Figure 10: Speaker EK (female); TS 8

The link accents preceding a focus accent show a gliding (slow) rise that starts within the accented syllable and peaks on the following unstressed syllable in both questions and statements (see Figs. 1 and 2 for statements and Figs. 8 and 10 for questions). Preceding other link accents two of the four speakers choose the same tonal pattern (cf. Fig.11, Speaker KP), namely L^*+H , regardless of whether the accent occurs in a question or in a statement. The other two speakers use an H^* (see Fig.1, Speaker EK) in statements and L^*+H in questions (cf. Fig. 8 and 10, Speaker EK). Despite the different realisation of the link before another link by speakers EK and KP, each speaker consistently uses an H^* or an L^*+H for the first link, irrespective of the focus type (non-contrastive broad focus and non-contrastive narrow focus). The ambiguity of the focus type is resolved by context for speaker EK (non-contrastive narrow and broad focus in Figs. 1 and 4, respectively) or by the use of emphasis for speaker KP (non-contrastive narrow and broad focus in Figs. 11 and 12, respectively).

In Bulgarian only the focus, the relevant ground information (link) and the sentence mode(statement/question) are tonally encoded. In Avgustinova and Andreeva (1999) the tail-associated intonation realisations are investigated. After nuclear accents the underlying L^*+H is completely reduced (deleted). In the case of Fig. 5 we observed a double focus information structure with intonationally reduced ($L^*+H\downarrow$) grounding material between the two foci.

3.4 Contrast vs. Emphasis

Figs. 11 and 12 are only distinguished by the higher F_0 maximum in the first figure. The pitch range from the accent peak (H^*) to the low boundary tone ($L\%$) is considerably less for non-contrastive broad focus. Speakers often use an increase of the pitch range, or emphasis, to mark narrow focus. Formally this can be indicated by an additional beat assigned to the metrically most prominent syllable of a narrow focus constituent, reflecting an increase of the F_0 maximum. Therefore, the interaction between the phonologically specified prominence and its tonal realisation leads to the realisation of phonetically different F_0 contours. Emphasis therefore helps to disambiguate the focus structure of the utterance. But the role of the emphasis is twofold. It can also be used to mark particular parts of the utterance in order to express contrasts (cf. Fig. 13 for contrastive narrow focus and Fig. 5 for contrastive broad focus).

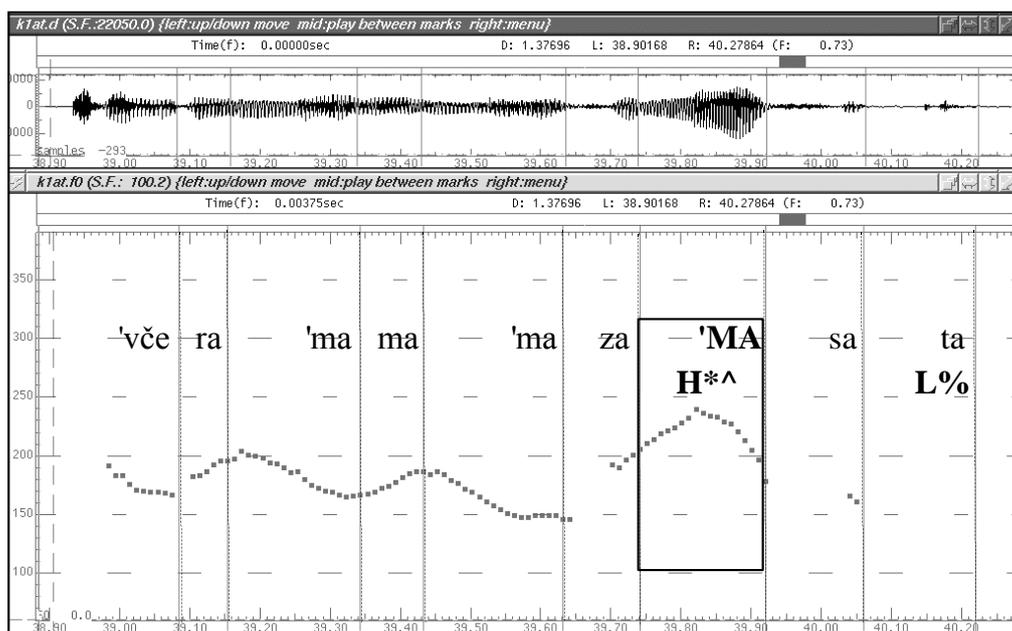


Figure 11: Speaker KP (female); TS 1 (non-contrastive narrow focus)

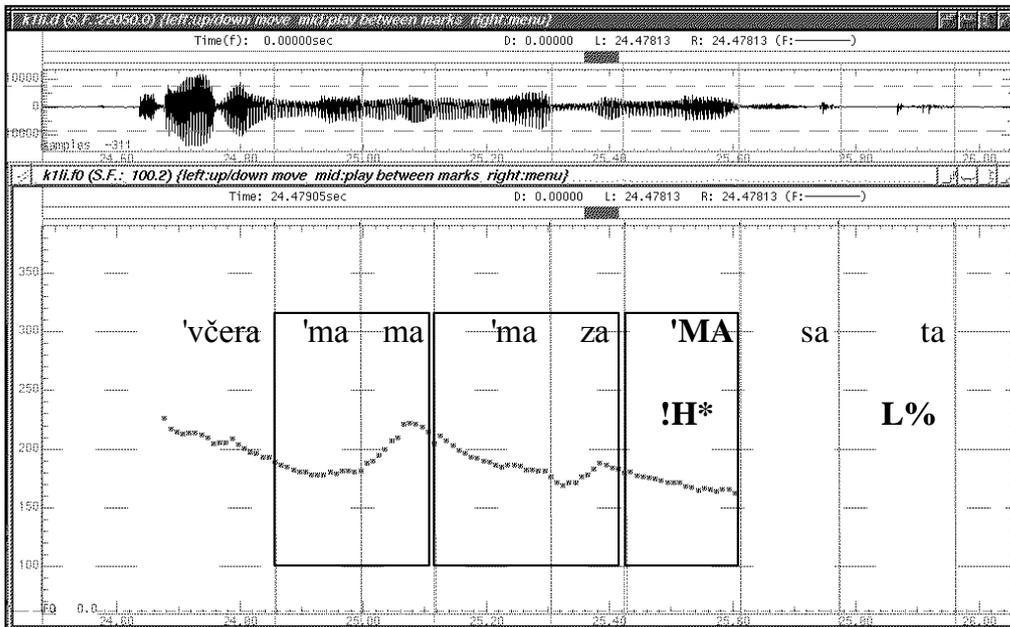


Figure 12: Speaker KP (female); TS 1 (non-contrastive narrow focus)

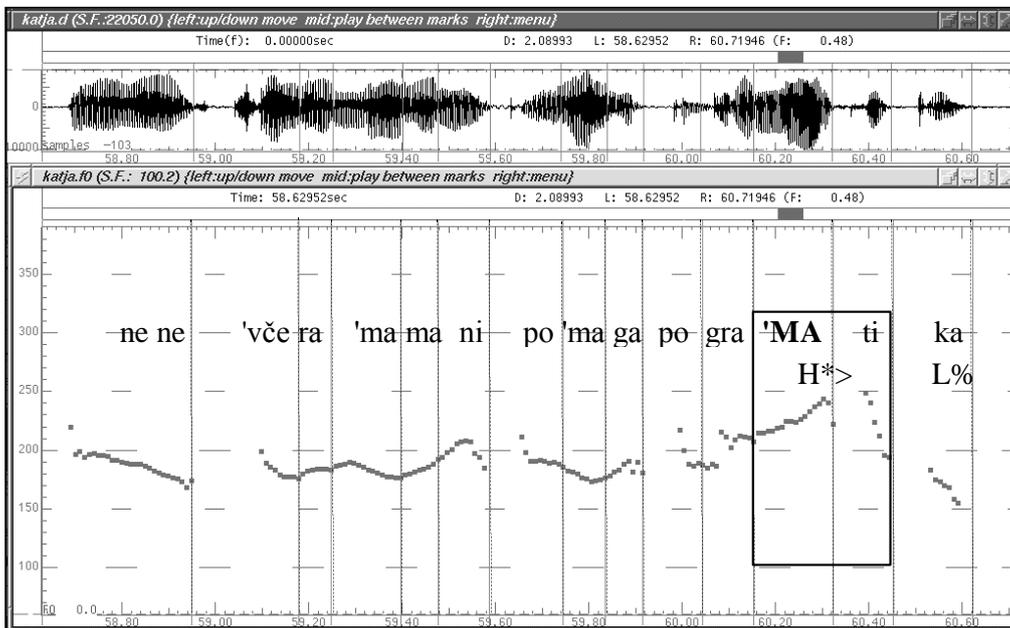


Figure 13: Speaker KP (female); TS 3 (contrastive narrow focus)

4. Conclusions

The canonical configuration in syntactically and prosodically unmarked statements in the investigated data is $H^*/L^*+H L^*+H H^* L\%$. In the opposition narrow vs. broad focus the underlying H^* for the nucleus is realised with an emphasis [+raised peak] in the marked member of the opposition (i.e. narrow focus) if the focus type is not disambiguated by the context. The emphasis is obligatory for the marked member in the opposition contrastive vs. non-contrastive. In the case of contrastive narrow focus, the phonetic realisation of the shape of the underlying H^* is also different, namely $H^*>$ [+raised peak; +delayed peak]. In the case of contrastive broad focus there is H^* with a disambiguating context and $H^*>$ otherwise.

The realisation of Yes/No questions is $L^*+H L^*+H L^*+H^\wedge$. Some speakers produce the accent pattern from one link to another link and from the link to focus differently. In questions a neutralization of this opposition is observed. This neutralization is the subject of an ongoing research.

References

- Tania Avgustinova (1997). *Word Order and Clitics in Bulgarian*. Saarbrücken Dissertations in Computational Linguistics and Language Technology. Volume 5.
- Tania Avgustinova and Bistra Andreeva (1999). *Thematic Intonational Patterns in Bulgarian Clitic Replication*. In: *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco. 1501-1504.
- Carlos Gussenhoven (1984). *On the Grammar and Semantics of Sentence Accents*. Dordrecht: Foris.
- D. Robert Ladd (1983). *Phonological Features of Intonational Peaks*. In: *Language*. Volume 59. 721-59.
- D. Robert Ladd (1996). *Intonational Phonology*. Cambridge University Press.
- Anastasija Miševa (1991). *Intonacionna sistema na bâlgarskija ezik*, Sofija: BAN
- Jordan Penčev (1980) *Osnovni intonacionni konturi v bâlgarskoto izrecenie*, Sofija: BAN
- Janet Pierrehumbert (1980). *The Phonology and Phonetics of English Intonation*. PhD thesis, MIT, published 1988 by IULC.
- Enric Vallduví and Elizabet Engdahl (1995). *Information Packaging and Grammar Architecture*. In: *J. N. Beckman Proceedings of North East Linguistic Society 25*. 1. University of Pennsylvania. 519-533.
- Enric Vallduvi and Elizabet Engdahl (1996). *The Linguistic Realisation of Information Packaging*. In: *Linguistics* 34, S. 459-519.