**Bistra Andreeva and Jacques Koreman**

Institute of Phonetics, Saarland University, Saarbrücken

E-mail: {andreeva, jkoreman}@coli.uni-sb.de

# The Status of Vowel Devoicing in Bulgarian: Phonetic or Phonological?

## ABSTRACT

Vowel devoicing at phrase boundaries is shown to depend on sentence mode (statement, question) and information structure (non-contrastive vs. contrastive and broad vs. narrow focus), which affect the type of nuclear accent and boundary tone. In *li*-questions, considerably less devoicing of the vowels before the phrase boundary is found than in other conditions, because the rising boundary contour prevents devoicing. In all other conditions, mode and information structure only affect the realisation of the vowel in the two syllables before the phrase boundary if the nuclear accent is realised on the last content word of the phrase, just before the boundary. In these positions, the differences in observed vowel devoicing are related to the intonational properties of the utterance, namely to the accent type. In particular, the peak scaling and alignment of the accent explains the presence or absence of voicing, or better the amount of devoicing, since it is clearly a gradient, phonetic phenomenon. In general, the lower and earlier the peak, the greater the tendency towards devoicing.

## 1. INTRODUCTION

In the phonetic typology of Slavic languages Bulgarian is said to be a language with accommodative pronunciation, which is characterised by a large number of contextually conditioned allophones of vowels and consonants, motivated either by the direct segmental context or by their position in the syllable/word or by their distance from the accented syllable. Bulgarian has strong qualitative vowel reduction, depending on accent (Level 2 in the first pre-tonic and in all post-tonic syllables, Level 1 elsewhere) [1]. To our knowledge, no vowel devoicing has been reported in the literature, except in [2].

Vowel devoicing is a well-known phenomenon in a large number of languages [3] and depends on the quality and duration of the vowel, its segmental context, speech rate, level of stress, pitch and morpho-phonological boundaries. It is sometimes also employed for phrase boundary demarcation in addition to boundary tones and F0 resets, final lengthening and voice quality in general [4, 5]. In a recent exploratory analysis of segmental deviation from the canonical phonetic form, intonation-dependent elision and reduction of phrase-final unstressed vowels was observed in a falling (L-L%), but not in a rising (L-H% or H-H%) tonal contour in the spontaneous (Map Task) and read speech recordings of Sofia Bulgarian and of Moscow Russian [6]. In addition, as we reported in [2], phrase endings can also be marked by vowel devoicing under the same condition.

On closer inspection of the Map Task and read speech data, however, we noticed that there is some variation as to the exact positions and tonal context in which vowel devoicing applies. In order to evaluate factors which cause vowel devoicing and to examine them in some more detail under strictly controlled conditions, a production experiment was conducted. Particularly we investigate whether vowel devoicing is a categorical, phonological or a gradient, phonetic process like other segmental reduction phenomena. In this paper we summarise and extend our earlier analyses [2]. The following three hypotheses regarding the factors regulating the distribution of devoiced vowels are investigated:

- Vowel devoicing depends on mode and information structure (focus and contrast).
- Vowel devoicing is more likely if the nuclear accent occurs earlier in the phrase.
- Vowel devoicing depends on accent type.

## 2. MATERIAL AND METHOD

The subjects for the first experiment were five persons from Sofia (4 female and 1 male, aged 25-45 years). The following 4 test sentences (with the material relevant to phrasal demarcation underlined) where the information structure is controlled for the analysis of the relation between focus structure and intonation were recorded several times in random order in a sound-treated studio. These sentences are a subset of a larger data set.

1. *'včera      'mama   ni   po'maga  po  gra'matika*.
   yesterday  mama   us   helped    in   grammar
   'Yesterday mum helped us in grammar.'

2. *'včera      'mareto   po'maga   po  gra'matika?*
   yesterday  Mareto   helped    in   grammar
   'Yesterday Mareto helped in grammar?'

3a. *'včera      'mama   li   vi    po'maga    po  gra'matika?*
   yesterday  mama   Q[1]  you  helped       in   grammar
   'Is it mum who helped you in grammar yesterday?'

3b. *'včera      'mama   vi    po'maga  po  gra'matika   li?*
   yesterday  mama   you   helped    in   grammar     Q
   'Is it grammar that mum helped you with yesterday?'
   'Yesterday mum helped you in grammar, didn't she?'

The statements in the material (sentence 1) were embedded in dialogue exchanges as replies to *wh*-queries uttered by the instructor and were produced three times with broad focus or with contrastive[2] and non-contrastive narrow focus on the first or last content word (*'mama'* or *'gramatika'*).

|  | broad | narrow non-final | narrow final |
|---|---|---|---|
| statements [-contrast] | x | x | x |
| statements [+contrast] | ■ | x | x |
| checks [-contrast] | ■ | x | x |
| *li*-questions [-contrast] | $x^2$ | x | $x^2$ |

**Table 1.**   Realised focus conditions for four sentence modes (black areas indicate missing focus positions)

---

[1]   Q = interrogative particle '*li*'

[2]   It may be more appropriate to call these foci *corrective* instead of *contrastive* (cf. [7]).

Besides statements, the material contains both confirmation-seeking yes-no questions (checks), marked only by intonation (sentence 2), and information-seeking yes-no questions with the interrogative particle *'li'* [4] (sentences 3a and b) produced with narrow focus on the same content words as in the statements. Sentence 2 was produced three times as a reaction to a described situational context, which was constructed to induce focus on the first or the last content word in the sentence. Sentences 3a and 3b were read four times in a random list. The realised focus conditions across sentence mode are given in Table 1.

The recordings were digitised at a sampling frequency of 16 kHz and with an amplitude resolution of 12 bits, using the Advanced Speech Signal Processing Tool (xassp) [8]. All data were manually labelled on the basis of the synchronised microphone signal and spectrogram. Since we want to investigate devoicing as a cue for phrase demarcation, the material *'-tika(li)'* after the last lexically stressed vowel in the utterance was labelled in great detail (Fig.1), in particular the vowels, which were divided into devoiced and modally voiced portions. We used a slightly modified SAMPA-notation for Bulgarian by adding the following symbols: <I-h:> and <@-h:> for devoiced (portions of the) vowels, <t> and <k> for closure, <t-r> and <k-r> for release, and <t-h> and <k-h> for aspiration of the voiceless stops. The aspiration phase after the release of the syllable-initial plosive was analysed as part of the following vowel. There are two reasons for this: first, although the aspiration phase is usually considered to be part of the plosive, it cannot be distinguished from a devoiced realisation of the vowel; second, we observed that the aspiration phase is longer in weakening context such as unaccented syllables, so that it is more appropriate to analyse it as part of the vowel than as part of the consonant (the latter would imply that the consonant is strengthened in unaccented positions). For each of the vowels, the duration of the devoiced signal was divided by the total vowel duration *(relative devoiced vowel portions)*.
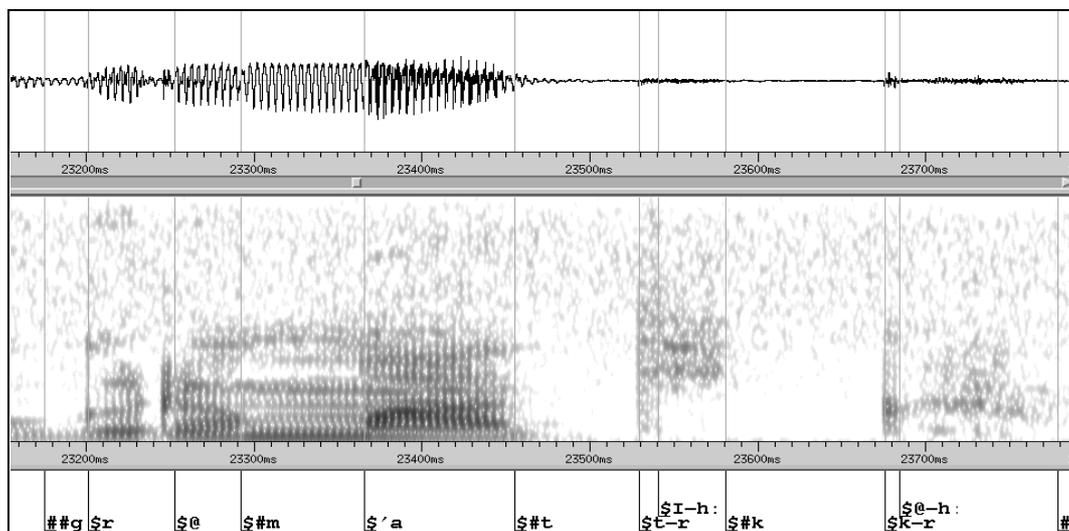


**Figure 1:** Example of the segmental labelling. The sample material is *'gramatika'*. The labels for the modally-voiced vowels and for the consonants are basically phonemic. For the devoiced (portions of the) vowels there are some phonetic labels, too (<I-h:>, <@-h:>). Special labels are used to refer to sub-phonemic events like closure (<t>, <k>) and release (<t-r>, <k-r>) of voiceless stops.

In addition to the segmental labelling the pitch accents, phrase accents and boundary tones were also labelled, using ToBI [9], with the peak alignment of the H(igh) tones explicitly specified.

---

[4]   Typically *'li'* belongs to the focal segment. When it occurs in sentence-final position an ambiguous focus interpretation arises (broad or narrow in final position).

# 3. RESULTS

To analyse the effects on phonatory demarcation, we carried out multivariate analyses of variance, with Scheffé post-hoc tests if appropriate. Significant effects at the 5%-level are reported for the relative devoiced portions of [i] and [a] in the word *'gramatika'*. We shall also report the percentage of *complete* devoicings, which we did not analyse statistically[5].

## 3.1. The effect of sentence mode, focus position and boundary tone

To see if there are any phrase-final devoicing effects of mode, focus and focus position, we first compare two sentence modes: statements (with non-contrastive vs. contrastive focus) and questions (checks and *'li'*-questions). They vary with respect to the position of the focus, which either comes on the first or on the last content word. In [10] the focus-associated pitch movements in statements and questions were analysed as follows: Statements and checks with identical morpho-syntactic structure differ only in the nuclear pitch accent and were analysed as L*+H for the checks, as H* with late peak alignment for the contrastive statements and as H* with early peak alignment for non-contrastive statements with broad and narrow focus. The focus-associated accent in *'li'*-questions was analysed as L*+H. In both sentence modes the boundary tones are referred to as L-L%.

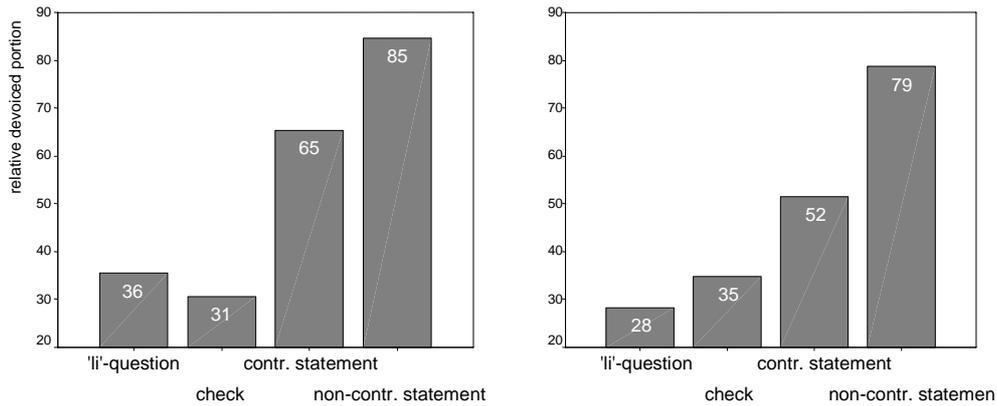### 3.1.1. Comparison of statements and questions with focus in final position

Although *'li'*-questions are analysed in [10] as having the same nuclear pitch movement as checks (L*+H L-L%), they are not directly comparable. First, when the focus is on the last content word in the utterance, this word is followed by the particle *'li'*. Therefore, the /a/ context in *'li'*-questions is different from that in checks (as well as in all other conditions), since it is followed by /l/ instead of a phrase boundary. This fact is expected to reduce the tendency towards devoicing. Furthermore, a comparison of the penultimate and antepenultimate syllables in *'li'*-questions with the final /ka/ and the penultimate /ti/ in the other conditions is not warranted.

When the focus is on the last content word (final), no complete vowel devoicings occur for checks, while in contrastive and non-contrastive statements 33% and 53% of the /i/'s and 20% and 47% of the /a/'s, respectively, are completely devoiced. On average, /i/ has significantly shorter relative devoiced portions in checks (31%) than in contrastive and non-contrastive statements (65% and 85%, respectively). The relative devoiced portion for /a/ is significantly shorter in checks and contrastive statements (35% and 52% respectively) than in non-contrastive statements (79%). The vowels in non-contrastive vs. contrastive statements never contain modal voicing, while this is only true for /a/ in checks. A large portion of /i/ is modally voiced in checks (69%). In general, a stronger tendency towards vowel devoicing is found in non-contrastive than in contrastive statements, while they are weakest in checks.

We observe that neither /i/ nor /a/ is ever completely devoiced in *'li'*-questions with narrow focus on the last content word (final), although we do sometimes find a short relative devoiced portion (36% for /i/, 28% for /a/). Besides checks with final focus positions, the *'li'*-questions in this condition were the only utterances with modal voicing in /i/ and /a/ (53% and 72%, respectively). Figure 2 shows the distribution of devoicing rate as a function of the sentence mode and focus type in final position.

---

[5]  Wilcoxon tests on the averaged data per subject are possible, but we decided against this because of the small dataset (5 observations per cell).

**Figure 2:** Mean relative devoiced portion for /i/ (left graph) and /a/ (right graphs) for 'li'-questions, checks, contrastive and non-contrastive statements in the final condition.

### 3.1.2. Comparison of statements and questions with focus in non-final position

When the focus is on the first content word (non-final), vowel devoicing in the *'li'*-questions (77% for /i/ and 43% for /a/) is significantly different from that in the checks and contrastive and non-contrastive statements, where a tendency is found towards complete devoicing of both vowels (cf. Fig. 3).
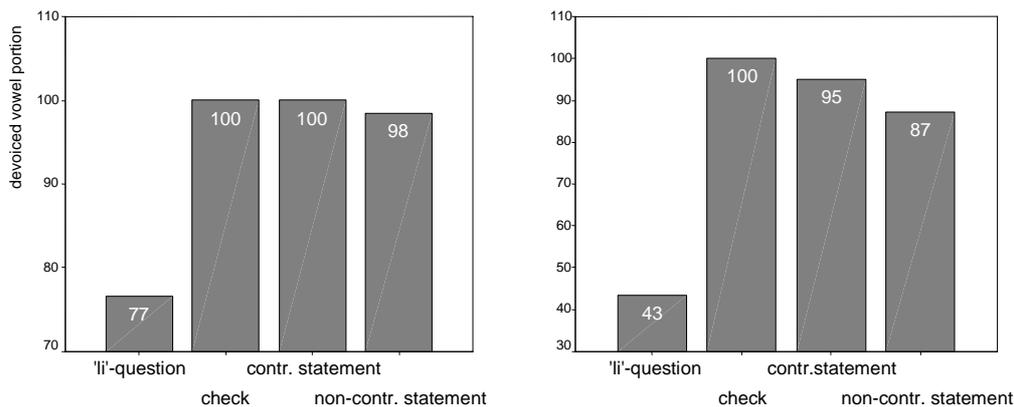


**Figure 3:** Mean relative devoiced portion for /i/ (left graph) and /a/ (right graphs) for 'li'-questions, checks, contrastive and non-contrastive statements in the non-final condition.

Incomplete but strong devoicing is found for /i/ in non-contrastive statements, of which 98% are completely devoiced, and for /a/ in non-contrastive and contrastive statements, with 87% and 95% devoicing, respectively. Therefore the results in section 3.1.1 cannot be entirely explained by sentence mode per se. In the case of checks, contrastive and non-contrastive statements it is the distance of the nuclear accent from the intonation phrase boundary which also affects the degree of vowel devoicing: the greater the distance from the phrase boundary, the greater the relative devoiced portion of the vowels in the last two syllables of the phrase. In the case of *'li'*-questions on the other hand the boundary tones were realised as L-H% instead of L-L% (cf. Fig 5 and Fig. 6). This may be because we are dealing with read speech, or because of a list effect (since only the *'li'*-questions were read in a list).
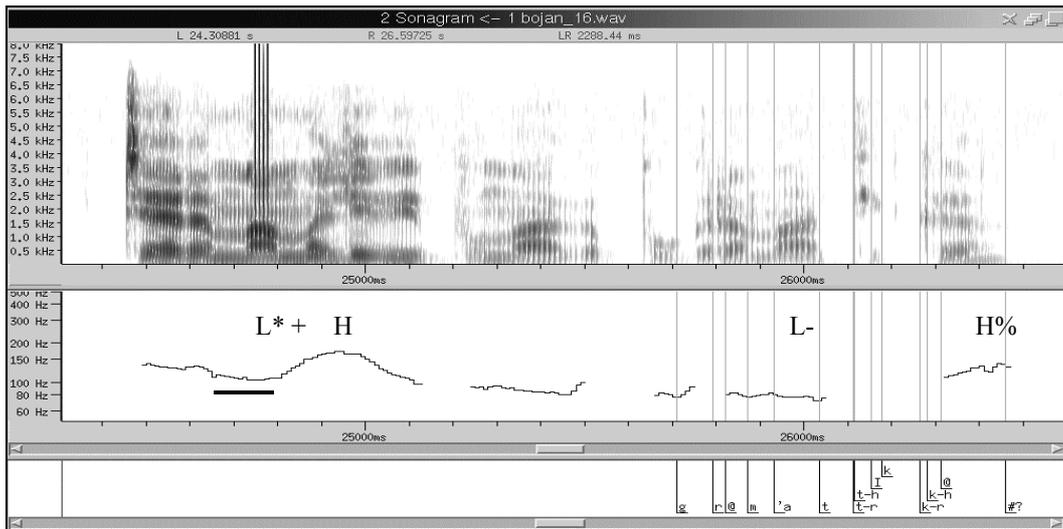
**Figure 5:** Example of the realisation of L-H% boundary tones in *'li'*-questions (non-final focus).
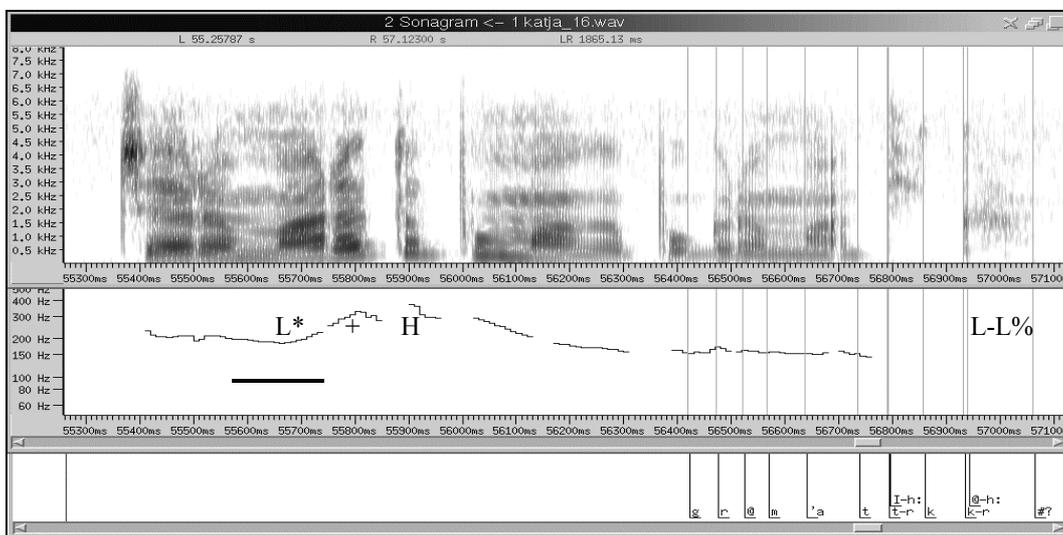


**Figure 6:** Example of the realisation of L-L% boundary tones in checks (non-final focus).

The finding that the rising boundary contour L-H% prevents devoicing because it requires an increased sub-glottal pressure and a different glottal adjustment compared to the final falling contour [11], fits with the observation in [6] that L-H% also prevents vowel loss and reduction. The fact that we find more devoicing in *'li'*-questions with the non-final accent (cf. Fig. 3) than with a final accent (cf. Fig. 2) is that the syllables under investigation (*'tika'*) are closer to the boundary in the former – since the question particle *'li'* occurs after the nuclear accent early in the phrase in the non-final condition. Additionally, the fact that /a/ occurs in a strong foot position (*'gramatika li'*) and is followed by a sonorant consonant /l/ further reduces devoicing.

## 3.2. The effect of accent type and position

As expected, differences are only found when there is a focus on the last content word in the phrase and the boundary tone is realised as L-L%. This shows that the vowel realisation in the ultimate and penultimate (post-nuclear) syllables is not dependent on mode per se, but is a consequence of the realisation of the boundary tone as well as the nuclear pitch accent and its position in the utterance. Table 2 summarises the number of pitch accent types used in the different test conditions (numbers given across final and non-final positions). In checks, the pitch accent is *always* realised as L*+H.

| Focus x contrast | Accent type | | | | |
|---|---|---|---|---|---|
| | H*(early) | !H*(early) | H*(late) | !H*(late) | L*+H |
| Broad | 3 | 12 | 0 | 0 | 0 |
| Narrow non-contrastive | 17 | 1 | 12 | 0 | 0 |
| Narrow contrastive | 2 | 0 | 22 | 8 | 0 |
| check | 0 | 0 | 0 | 0 | 27 |

**Table 2.** Realised accent types for the different test conditions.

Broad focus in non-contrastive statements is *always* realised with an early (mostly also downstepped) peak on the last content word in the utterance. The pitch accent signals a complex focus domain in the case of focus projection.

This type of accent was also used in the majority of the narrow non-contrastive foci. When focus is on the last content word the early accent causes ambiguity, since it can be interpreted as a broad focus. Besides an increase in pitch range or emphasis, speakers sometimes use a late peak in an attempt to resolve the ambiguity: we counted 11 early and 4 late peaks in non-contrastive statements with narrow focus in final position.

In contrastive statements, the subjects differ in their realisation of the pitch accent. Like non-contrastive statements, these also have falling nuclear pitch accents, but there is a stronger tendency towards realising late peaks in contrastive statements (2 early and 30 late peaks in final and non-final position).

To test the effect of the particular pitch accent, we shall compare the accents listed in Table 2 in final and non-final position, namely L*+H, with late peak alignment in the post-accentual syllable(s), (!)H*, with late peak alignment at the end of the accented syllable, and (!)H*, with early peak alignment (early in the accented or in the pre-accented syllable). The boundary tones in the test sentences are realised as L-L%. In the non-final conditions there was no significant effect on the devoiced portions in /i/ and /a/: Both vowels showed a strong tendency to devoice, caused by the decrease of sub-glottal pressure and the resulting low volume velocity airflow, which in turns inhibits voicing [3, p. 100]. Therefore we shall report the results for focus *in final position only* (cf. Fig. 7-11).
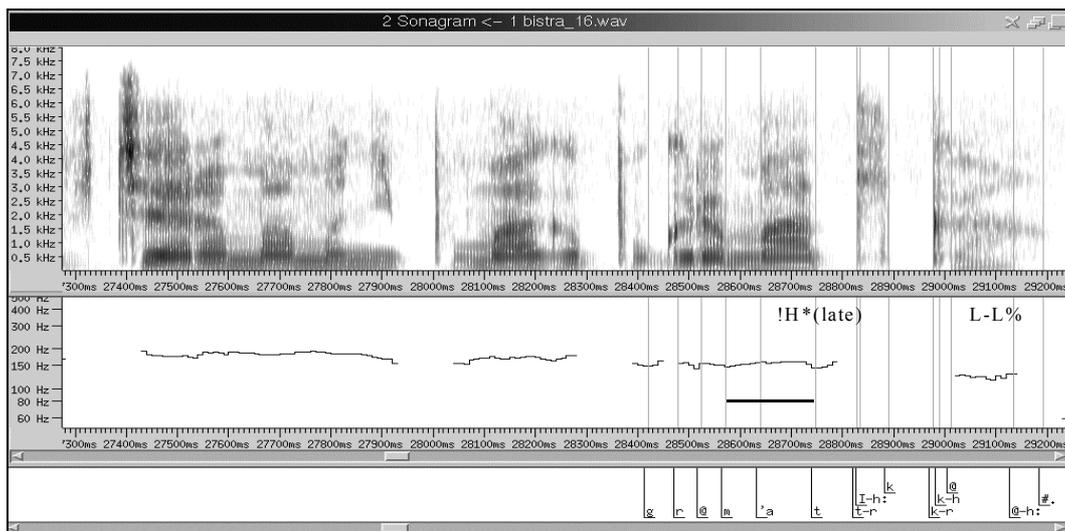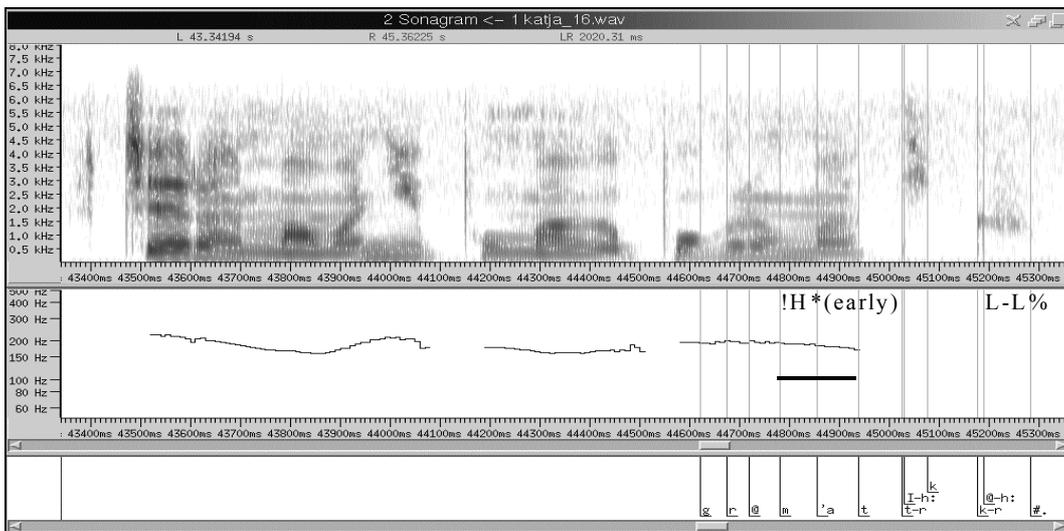


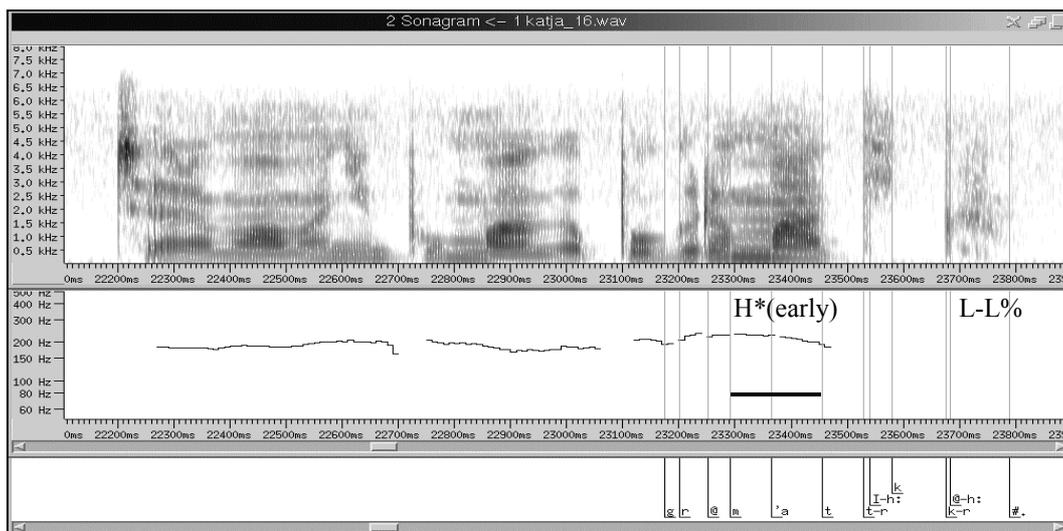**Figure 7:** Example of the realisation of !H*(late) in final position (narrow contrastive focus).
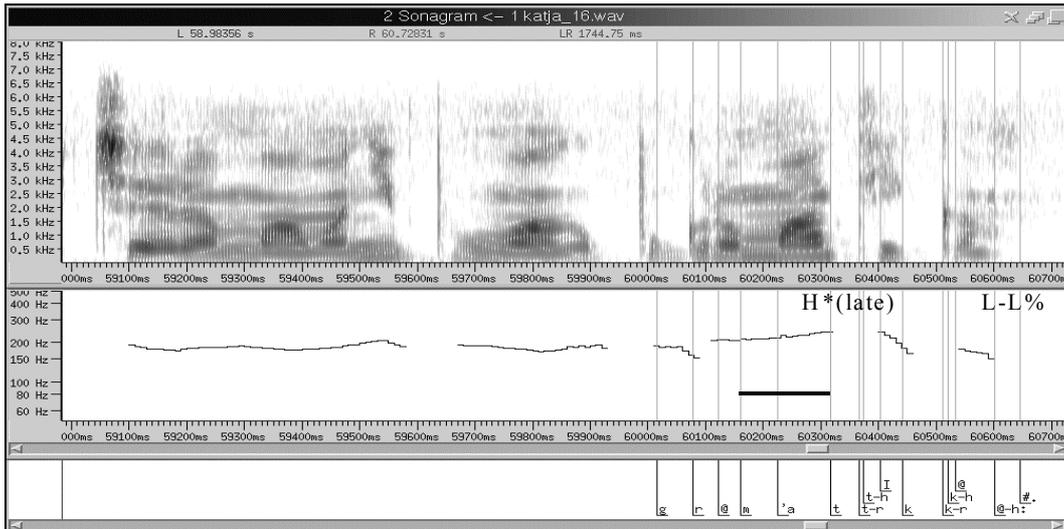
**Figure 8:** Example of the realisation of H*(late) in final position (narrow contrastive focus).
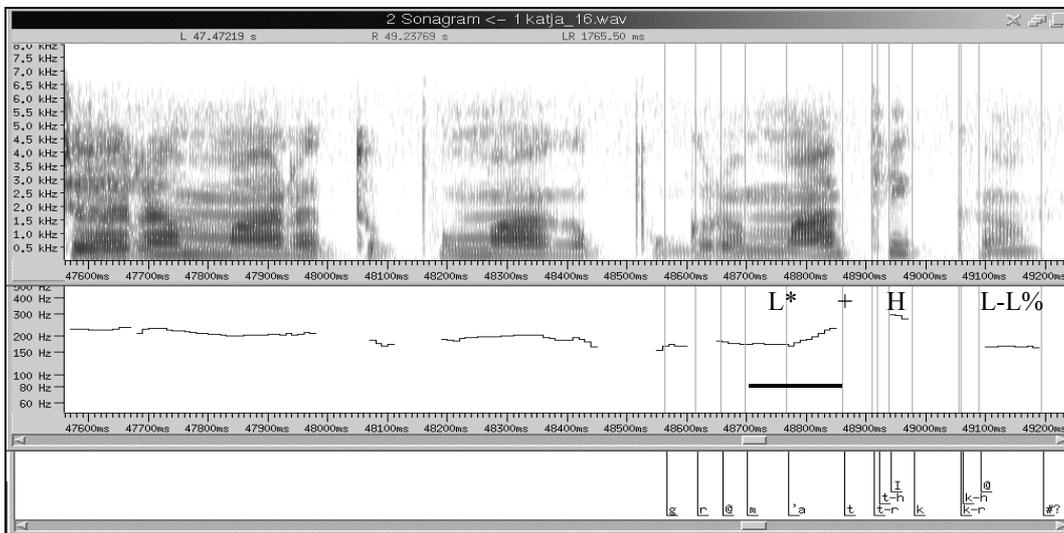


**Figure 9:** Example of the realisation of L*+H in final position (check).



**Figure 10:** Example of the realisation of !H*(early) in final position (broad focus).

**Figure 11:** Example of the realisation of H*(early) in final position (narrow non-contrastive focus).

If we reanalyse the utterances with final focus with respect to accent type, we find that the relative devoiced portion of the /i/ and /a/ is dependent on the peak alignment and height of the pitch accent. The relative devoiced portion of /i/ is significantly larger in !H* (early) than in (!)H* (late)  than in L*+H conditions. H* (early without downstep) groups together with H* (late) and !H* (early). Therefore, the earlier and lower the peak, the stronger the tendency to devoice the vowel. For /a/ a significant difference was only found between the (!)H* (early) and the other conditions (L*+H and (!)H* (late)). This is probably because the earlier the peak is aligned, the sooner a low pitch is reached, allowing vowel devoicing.

The pooled results for the data discussed in section 3.2 are shown in Figure 12.
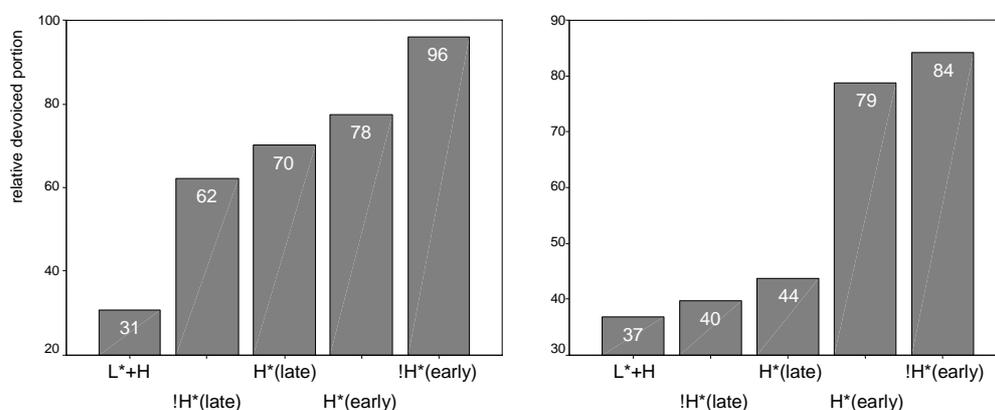


**Figure 12:** Mean relative devoiced portion for /i/ (left graph) and /a/ (right graphs) for all H* (early), H* (late) and L*+H accents on the last content word, excluding '*li*'-questions.

## 4.   SUMMARY

In analogy to the observations for vowel loss and reductions [6], we never observe complete vowel devoicing with final rising intonation contours ('*li*'-questions), while it is quite frequent in falling intonation contours (statements and checks). This is also reflected in the shorter relative devoiced portions in '*li*'-questions as compared to statements and checks. Conversely modal voicing is only found in '*li*'-questions (with focus in final and non-final positions) as well as checks with focus in final position.

For the final falling intonation contours (L-L%), the vowels just before the phrase boundary are (almost) completely devoiced when the nuclear accent comes early in the intonation phrase. When the nuclear pitch accent is realised on the last content word in the phrase, we found that the peak scaling and alignment exert a strong influence on the vowel realisation, since early and downstepped peaks tend to allow for a stronger vowel devoicing than late peaks.

It was also shown that vowel quality determines the strength of the effect, the more closed vowel /i/ being more sensitive to devoicing than /a/.

## 5. DISCUSSION

The present study, which focuses on Bulgarian, investigates whether vowel devoicing at intonation phrase boundaries is a categorical, functional phenomenon or a gradient, phonetic effect of the implementation of the accent types which signal linguistic functions.

Although vowel devoicing can be phonologically determined [3, 12], our results show that it is not in the conditions investigated in the present paper. The choice of the pitch accent is linguistically determined,  depending on sentence mode and focus. In turn, the pitch accent influences vowel devoicing. Devoicing occurs when the transglottal pressure is so low that vocal fold vibration can no longer be sustained. The fact that accents early in the phrase do not affect vowel devoicing at the phrase boundary shows that this is not a phonological process. When pitch accents occurs on the last content word in the phrase, the amount of vowel devoicing depends on the scaling and alignment of the accent peak clearly showing a gradient rather than a categorical relationship.

The fact that no devoicing occurs at rising (L-H%) intonation phrase boundaries supports a phonetic analysis. At rising phrase boundaries the transglottal pressure is kept high in order to sustain voicing despite worse phonation conditions caused by stiffening of the vocal folds to reach the high $F_0$ target, thus preventing devoicing.  Further support for the phonetic status of vowel devoicing comes from the observation that there is a stronger tendency towards devoicing for /i/ than for /a/. This coincides with the phonetic argumentation for devoicing, since the drop in transglottal pressure in /i/ is enhanced by the narrower vocal tract compared to /a/.

In conclusion, it was shown that vowel devoicing at the intonation phrase boundary is a gradient, phonetic phenomenon which depends on the phonetic realisation (peak scaling and alignment) of phonological categories (pitch accent type), the choice of which depends on sentence mode and focus type.

## REFERENCES

[1]  D. Tilkov, T. Bojadžiev, E. Georgieva, J. Penčev, V. Stankov, *Gramatika na sâvremennija bâlgarski knižoven ezik (vol. 1): Fonetika,* Sofia: Bâlgarska akademija na naukite, 1982.

[2]  B. Andreeva, J. Koreman & W.J. Barry, "Phonatory demarcations of intonation phrases in Bulgarian." *Proc. of the 15th Int. Congress of Phonetic Sciences (ICPhS 2003),* Barcelona, pp. 611-614, 2003.

[3]  M. Gordon, "The phonetics and phonology of non-modal vowels: a cross-linguistic perspective," *Berkeley Linguistics Society*, vol. 24, pp. 93-105, 1998.

[4]  R. Dauer, "The reduction of unstressed high vowels in Modern Greek," *Journal of the International Phonetics Association.* vol. 10, pp. 17-27, 1980.

[5]  C. Smith, "Marking the boundary: utterance-final prosody in French questions and statements." *Proc. of the XIVth Int. Congress of Phonetic Sciences (ICPhS 1999)*, San Francisco, pp. 1181-1184, 1999.

[6]  W. J. Barry and B. Andreeva, "Cross-language similarities and differences in spontaneous speech patterns," *Journal of the International Phonetic Association,* vol. 31/1, pp. 51-66,

2001.

[7]  C. Gussenhoven (forthcoming), "Types of Focus in English." In Daniel Büring, Matthew Gordon & Chungmin Lee (eds.) *Topic and Focus: Intonation and Meaning. Theoretical and Crosslinguistic Perspectives*. Dordrecht: Kluwer.

[8]  http://www.ipds.uni-kiel.de/forschung/xassp.de.html

[9]  M. Beckman and G. A. Elam, "Guidelines for ToBI labeling," version 3. Linguistics Department, Ohio State University, 1997

[10] B. Andreeva, T. Avgustinova and W. J. Barry, "Link-associated and focus-associated accent patterns in Bulgarian." In: , G. Zybatow, U. Junghanns, G. Mehlhorn and L. Szucsish, (eds.) *Current Issues in Formal Slavic Linguistics*, Frankfurt: Peter Lang, pp. 353-364, 2001.

[11] P. Ladefoged, *Three Areas of Experimental Phonetics.*Oxford University Press, Oxford, 1967.

[12] Loren Bliese, „Afar." In Bender, M. Lionel (ed.). *The Non-Semitic Languages of Ethiopia*. East Lansing, MI: African Studies Center, Michigan State University, pp. 232-277, 1976.