Original Paper

Phonetica 2016;73:256–278
DOI: 10.1159/000448044

Received: May 4, 2015
Accepted after revision: June 27, 2016
Published online: February 23, 2017

Local and Global Cues in the Prosodic Realization of Broad and Narrow Focus in Bulgarian

Bistra Andreeva[a] William John Barry[a] Jacques Koreman[b]

[a]Computational Linguistics and Phonetics, Saarland University, Saarbrücken, Germany; [b]Department of Language and Literature, NTNU, Trondheim, Norway

Abstract

In this study, local and global prosodic cues for information structure are examined in the elicited production of six Bulgarian sentences. The sentences were produced in response to different questions, devised to prompt different focus realizations (broad focus and non-contrastive and contrastive narrow focus). Results show that speakers consistently differentiate broad and narrow focus by means of both local and global acoustic cues, by producing different pitch accent types on the nuclear syllable and reducing the 'phonetic strength' of the default pre-nuclear accent in the narrow focus condition. Thus, the difference between the acoustic properties of the nuclear and the pre-nuclear accented syllables is smaller in the broad focus condition and greater in the narrow focus condition. Contrastive and non-contrastive narrow-focus accents are differentiated by *local* cues, i.e., by longer duration when the focus is early in the sentence and by *global* cues, i.e., by enhancing the tonal contrast between the nuclear prominence of CW2 and the pre-nuclear prominence of CW1 when the focus is late in the sentence.

1. Introduction

Most languages employ mechanisms that give prominence to mark the relative informational importance of particular words in a phrase, often combined with word order and special lexical items or syntactic constructions. It is common to distinguish three elements of information structure: 'topic' (the subject matter on which new information is to be offered), 'focus' (the new information offered) and the 'given information' (information given previously or assumed to be known) (e.g., Féry and Krifka, 2008).

These elements can be realized prosodically by means of a 'topic accent', a 'focus accent' or by 'de-accentuation'. At some basic production level, the speaker

KARGER

E-Mail karger@karger.com
www.karger.com/pho

Bistra Andreeva
Computational Linguistics and Phonetics
Saarland University
DE–66123 Saarbrücken (Germany)
E-Mail andreeva@coli.uni-saarland.de

invests more effort in accentuated words than in words that convey given information, with the consequent acoustic effects of greater duration and intensity, higher or changing fundamental frequency (*f0*) and in some way more distinct spectral properties (Bertinetto, 1981; Dauer, 1987; Kochanski et al., 2005; Sluijter and van Heuven, 1996). However, there is evidence that languages differ in the amount each of the acoustic dimensions changes under accentuation (Andreeva et al., 2014; Koreman et al., 2008; Koreman et al., 2009), and there is considerable debate about which properties are used by the listener to identify prominent words or syllables. Pitch (measured as *f0*) is often seen as dominant (Cooper et al., 1985; Eady and Cooper, 1986; Lieberman, 1960), but duration (Beckman, 1986; Fry, 1955), intensity (Beckman, 1986; Kochanski et al., 2005; Turk and Sawusch, 1996) and voice quality (Sluijter and van Heuven, 1996) have also been singled out as important if not dominant determinants of perceived prominence.

The observation that one and the same sentence can be an acceptable answer to different questions as long as it is spoken with different accentuation patterns was made as early as 1880 by Hermann Paul (1937). Depending on the information provided by the pre-context, the focused part of a phrase can be extended over much of the phrase – 'broad focus' – or restricted to one word – 'narrow focus' –, which in turn can be non-contrastive or contrastive (but see discussion below). In (1), (2a) and (2b) below illustrations of the three focus types are given using the sentence 'ПЕТЪР СЕ ПРЕМЕСТИ В СОФИЯ' (*Peter has moved to Sofia*).

1. broad focus:
   ИМА ЛИ НЕЩО НОВО? (*Is there anything new?*)
   [ПЕТЪР СЕ ПРЕМЕСТИ В СОФИЯ]$_{F}$.

2a. non-contrastive narrow focus:
   КЪДЕ СЕ ПРЕМЕСТИ ПЕТЪР? (*Where did Peter move to?*)
   ПЕТЪР СЕ ПРЕМЕСТИ В [СОФИЯ]$_{F}$.

2b. contrastive narrow focus:
   ПЕТЪР В ПЛОВДИВ ЛИ СЕ ПРЕМЕСТИ? (*Did Peter move to Plovdiv?*)
   ПЕТЪР СЕ ПРЕМЕСТИ В [СОФИЯ]$_{F}$.

The syntactic constituent in square brackets is indexed with 'F' to show its function as the focus corresponding to what the preceding question is asking about. While the whole sentence is focused in (1), only part of it is focused in (2a) and (2b). The question in (2a) and (2b) already implies that Peter has moved, so that part of the information in the answer is 'given', and only the place moved to is focused.

Regarding 'narrow focus', there is considerable disagreement in the literature about whether 'contrastive' and 'non-contrastive' focus are two distinct information structure categories. Some authors argue that 'contrast' can be regarded as an independent grammatical category, and that contrastive focus can be categorically discriminated from non-contrastive focus (see Chafe, 1976; Halliday, 1967; Kratzer, 2004; Molnár, 2002; Rochemont, 1986; Selkirk, 2002; Vallduví and Vilkuna, 1998). Others are of the opinion that contrastive focus cannot be regarded as an independent category of information structure (see Bolinger, 1961; Büring, 2007; Féry and Krifka, 2008; Jackendoff, 1972; Rooth, 1992). Clearly, the context may or may not specify a semantic entity to which the focused word is in explicit contrast, providing a textual basis for a distinction. In (2a), there is no explicit set of possible cities (like Burgas, Russe, Sofia, Varna, Veliko Tărnovo), from which *Sofia* is being selected. In (2b), an

explicit alternative city is being introduced: *Plovdiv*. The information in sentence (2b) explicitly contrasts with a specific item (*Plovdiv*) which has already been mentioned in the discourse. However, e.g., Rooth (1992) sees an implicit contrast in any narrow focus; any expression has two semantic representations: the meaning of the expression itself and a set of alternatives. In the case of explicit contrast, the alternative is known, but for Rooth the meaning of the expression does not change if the alternatives are not explicit.

In the study presented here, *non-contrastive* focus will be used to refer to responses to questions *without* an explicitly mentioned alternative to the answer, where the focus refers to the information required by a question. *Contrastive* focus will be used to refer to cases *with* an explicitly mentioned alternative, where the speaker introduces a contrast into the discourse and wants to actively override an element of what (s)he believes to be the addressee's informational state.

Clear *prosodic* evidence for or against a contrastive vs. non-contrastive distinction is not apparent from the literature. Some have claimed that there is no difference (Bolinger, 1961; Cutler, 1977; 't Hart et al., 1990), while others have found evidence and argued that some acoustic features differ between contrastively vs. non-contrastively focused elements (Bartels and Kingston, 1994; Breen et al., 2010; Bruce, 1977; Ito et al., 2004; Katz and Selkirk, 2011; Krahmer and Swerts, 2001). However, the results are not always consistent and sometimes even contradictory. For English, it has been found that a higher peak in the pitch accent is the most reliable perceptual cue for differentiating a contrastive from a non-contrastive narrow focus, e.g., Bartels and Kingston (1994), and Ito et al. (2004) observed that speakers more frequently used a large pitch excursion (L+H*) rather than simply a high peak (H*) to signal contrast. A clear difference was also found by Breen et al. (2010) in a production task in which subjects were explicitly required to differentiate contrastive and non-contrastive focus, albeit with *f0* findings that contradicted the former two studies. The distinction was expressed with higher intensity, longer durations and surprisingly *lower* mean *f0* in contrastive focus. However, an asymmetry was found between the perception and production results. Listeners were unable to use the distinction made in the production consistently to differentiate between contrastive and non-contrastive focus. The authors, therefore, conclude that there are no consistent semantic differences between contrastive and non-contrastive focus types.

Dialectal differences in the realization of contrastive and non-contrastive focus have also been observed. Data reported by Arvaniti and Garding (2007) showed that Californian speakers consistently used H* accents for new information and L+H* for contrastive focus, while Minnesota speakers used the same accent category (L+H*) for both, but still differentiated contrastive from non-contrastive focus by producing a higher peak in the former case. In Russian, the distinction between contrastive vs. non-contrastive focus is made by the accent type (high rise-fall vs. mid-fall), by a longer syllable duration and by prosodic phrasing (Alter et al., 2001; Mehlhorn, 2002; Meyer and Mleinek, 2008; Zybatow and Mehlhorn, 2000). German contrastive focus is characterized by different accent types (L+H*), greater *f0* excursion, later peak-alignment, longer duration of the accented syllable and higher intensity compared to non-contrastive focus (Alter et al., 2001; Baumann et al., 2006, 2007; Braun, 2005, 2015; Féry and Kügler, 2008; Grice et al., 2009; Kügler, 2008).

Somewhat surprisingly, there is also disagreement about the *reliability* of the broad vs. narrow focus distinction. Of course, acceptance of the same utterance following both

a pre-context cueing narrow focus and one cueing broad focus can only occur when the narrow focus is on the final lexical item. Given this condition, equal acceptance has been shown in several perceptual studies (Birch amd Clifton, 1995; Gussenhoven, 1983; Welby, 2003), while others claim that their subjects have consistently been able to make a distinction (Rump and Collier, 1996; Breen et al., 2010). Longer durations in narrow focused words have been found in production studies for Dutch (Hanssen and Cooper, 2008), German (Baumann et al., 2006, 2007), English (Breen et al., 2010; Eady and Cooper, 1986; Sityaev and House, 2003; Xu and Xu, 2005) and Serbo-Croatian (Smiljanić, 2004). Differences in tonal structure have also been reported: higher *f0* peaks in German and English, steeper falls in Dutch and expanded pitch range in Serbo-Croatian for the narrow focus condition and greater intensity for narrow focus in English (see literature cited above). Baumann et al. (2007) found additionally that vowel articulation plays a role in marking different focus structures: Vowels had more peripheral first and second formant values in narrow than in broad focus conditions (see also Cho, 2005 for American English). Smiljanić (2004) investigated the peak alignment patterns in two Serbo-Croatian dialects, one with (Belgrade) and one without (Zagreb) a lexical peak alignment contrast. In Zagreb, the peaks are aligned earlier in narrow focus, while in Belgrade only the peaks in the rising lexical accents are aligned earlier. The peaks in the falling lexical accents are unchanged or are aligned slightly later. In Swedish, narrow focus is marked with an additional H tone (Bruce, 1977). Some languages use different pitch accent types to express the distinction between narrow and broad focus, e.g., mostly H*/H*^ vs. !H*/H+!H* for German (Baumann et al., 2007), H*+L vs. H+L* for European Portuguese (Frota, 2000) and for Palermo Italian (Grice, 1995), L+H* vs. H+L* for Neapolitan Italian (D'Imperio, 1997).

In the research that focuses on the relation between prosody and information structure, it has been shown that focus interpretation is highly context dependent. Bruce (1977) claimed that the domain within which focus is expressed acoustically is larger than just the focused constituent and can affect the prosodic-acoustic realization of the whole sentence. Not only does a focused constituent itself undergo changes in *f0*, duration and intensity, there is also greater post-focus compression of *f0* and intensity, relative to broad focus sentences (Xu and Xu, 2005, for English; Vainio and Järvikivi, 2007, for Finnish; Hanssen et al., 2008, for Dutch, among others). For Polish, Hamlaoui et al. (2015) found the enhancement of *f0* and intensity to be more important in contrastive focus contexts, and the reduction of the given words to be more important in the non-contrastive focus conditions. Romanian speakers signal narrow focus by expanded pitch range on the focused word. However, they do not lengthen the segmental durations of the word under narrow focus compared to broad focus, but they make the narrow focused word relatively longer by compressing the durations of the adjacent words (Manolescu et al., 2009). Speakers seem to employ cues for both the nuclear accented word and the words in the pre-nuclear interval. The importance of such relative prominence patterns is also demonstrated by Breen et al. (2010) for English. They show that for the disambiguation between broad and narrow focus, the distribution of *f0*, word duration and intensity across the sentence is crucial. Calhoun (2006) reports that the performance of a focus prediction model, built using the NXT Switchboard corpus, is significantly improved when information about the acoustics of adjacent words is included in the model.

In summary, it can be said that in order to signal the information structure of a sentence, speakers use: (a) different pitch accent types and/or (b) the same accent type

with different strength in one or more of the acoustic properties, using language- or speaker-specific combinations of *f0*, duration and intensity and/ or (c) they suppress or enhance the prominence of surrounding words.

In this study, we will refer to *local* cues as the phonetic-phonological properties of the nuclear accented syllables (see (a) and (b) above), and to *global* cues as reflecting broader phonetic patterns in the intervals before and after the nuclear accented syllable, which in some cases vary independently of the tonal accent (see (c) above).

In this paper, the prosodic exponents of broad focus and of non-contrastive and contrastive narrow focus are examined in the Sofia variety of Contemporary Standard Bulgarian. The paper is organized as follows: in section 2, we give a brief overview of Bulgarian research with respect to information structure and intonation and we pose the research questions; section 3 outlines the methodology of the production experiment. In section 4 we show and analyze the results and finally we offer a concluding discussion in section 5.

2. Background

Our study is concerned with Contemporary Standard Bulgarian, which is spoken by approximately 8 million people and belongs to the Southern branch of the Slavic language family. Bulgarian is a language with variable (free) stress, where stress can fall on any syllable and can sometimes vary with different grammatical forms of the same word. Stress placement serves as a feature to distinguish otherwise phonemically identical words (e.g., /ˈpara/ ([ˈparə])‘steam’ vs. /paˈra/ ([pəˈra]) ‘coin’). Most Bulgarian scholars have pointed to the occurrence of secondary stress, especially in longer polysyllabic words, compound nouns and adjectives, etc. (Bojadžiev et al., 1999; Stojkov, 1966; Tilkov and Bojadžiev, 1990). In terms of rhythm, Bulgarian occupies an intermediate position on the stress-to syllable-timed continuum and is characterized as being of a mixed type (Barry et al., 2003; Dimitrova, 1998). Bulgarian is a stress-accent language where pitch variation is used for varying functions, such as disambiguation of different syntactic structures, signaling the difference between statements and questions, and between different types of questions, indicating the emotional state and attitudes of the speaker, highlighting important elements of the spoken message and regulating conversational interaction.

Although some basic ideas of the Prague School theory of functional sentence perspective (Mathesius, 1907, 1939, 1947; Trávniček, 1962) have been proposed for the analysis of Bulgarian as early as Szober (1933/1979), they did not receive systematic application until the work of Ivančev (1957/1978) and subsequently in Avgustinova (1997), Dyer (1992), Georgieva (1974), Leafgren (2002), Penčev (1980) and Rudin (1985). Important factors in the realization of the information structure in Bulgarian utterances are:

- word order, remarkably flexible and discourse conditioned, as in all Slavic languages;
- morphological category of definiteness, unusual in the Slavic language family;
- clitic replication of nominal material specific to Bulgarian;
- intonation, fairly malleable, as in languages like English and unlike, e.g., Czech which predominantly uses word order variation to mark information status.

At various times in the past, the intonation of Bulgarian has been the object of scientific interest. Generally Bulgarian researchers start out from the assumption of a complex concept of intonation. Intonation is understood as a complex interaction of fundamental frequency, intensity, duration, pauses, etc. and studied according to its grammatical meanings: question, statement, exclamation, command. Such a phonetic approach is inherent in the Russian intonation school and its followers (see e.g., Miševa, 1991; Nikolaeva, 1977; Tilkov, 1981).

A phonological representation of Bulgarian intonation in relation to the formal-syntactic and semantic aspects of the utterance was developed by Penčev (1980). He dedicated much of his work to the information structure of the Bulgarian sentence. His point of departure is that the

information structure of an utterance is determined by its syntax. Following the American structuralist approach, Penčev describes the intonation with the help of pitch levels. In his approach to the relationship between information structure and intonation, he defines five pitch levels. He introduces the pitch movement as follows: the unstressed final syllables of an utterance usually end at the lowest level 1, level 2 corresponds to the non-focused elements, level 3 to the non-final focus, level 3 or 4 to the final focus and level 4 or 5 to the emphatic focus. Penčev postulates ten basic intonation contours for Bulgarian sentences, six neutral (depending on the focus position) and four emphatic (regardless of the focus position). He distinguishes non-contrastive and contrastive accentuation, and defines the contrastive accent as emphatic accentuation, which he sees in syntactic-semantic terms as the negation of the information provided in the previous context (Y, but not X). The contrast can, but need not necessarily be signaled by a strengthening of the accented syllable, prepared by a crescendo over the preceding syllable sequence. The strengthening of the accented syllable is either the product of an increased *f0* on the syllable itself or a decreased *f0* on the preceding content word, resulting, in either case, in a greater interval between the focused and given parts of the utterance. It is only in these cases that de-accentuation occurs, which Penčev calls 'de-rhematisation'.

Miševa (1991) and Miševa and Nikov (1998) experimentally investigated the regularities of *f0* changes expressing phonetic prominence in terms of the traditional theme-rheme partitioning of the sentence. They conclude that the linguistically relevant phonetic characteristic of the given material (theme) is simply the absence of accentual prominence, i.e. de-accentuation. New material (rheme) shows the same intonational pattern in narrow and broad focus, but the accentual contrast between the prominent and the surrounding syllables is greater in narrow than in broad focus.

Within an autosegmental-metrical framework, Andreeva (2007) investigates a number of phonetic and phonological aspects of Contemporary Standard Bulgarian as spoken in Sofia, based on a set of data which comprises both quasi-spontaneous speech (map-task recordings) and strictly controlled read material. An inventory of pitch accents (L*, L*+H, L+H*, (!)H*, H+!H*/H+L*), phrase accents (L- and H-) and boundary tones (L% and H%) is derived from the combined analysis of both corpora and defined with respect to the various communicative functions we investigate. The analysis of the data from the read corpus did not always show a regular relationship between the focus structure and the intonational features of an utterance. In the broad focus condition, speakers use H+!H*/H+L*[1] (a high tone in the pre-tonic syllable followed by a downstepped/low target point in the accented syllable) or (!)H* (an early or middle, mostly downstepped, peak within the accented syllable). The nuclear pitch accent H* is used for the majority of the non-contrastive narrow focus cases. When the last lexical word is focused, the H* accent is ambiguous because the sentence can also be interpreted as a broad focus sentence. Apart from a greater pitch range, speakers occasionally use an L+H* accent to signal narrow focus and resolve the potential ambiguity, with a low tone at the beginning or shortly before the onset of the accented syllable and a high tone at the end of the accented or in the following syllable. In the contrastive focus condition, speakers show a clear preference for L+H* pitch accent.

Contrary to the findings in Miševa (1991) and Miševa and Nikov (1998) that 'the F0 pattern corresponding to the theme is of monotonic nature without any accentual contrasts' (Miševa and Nikov, 1998:282), Andreeva et al. (2001), Andreeva (2007) and Avgustinova and Andreeva (1999) report that the underlying (phonological) pitch accent pattern for the given material is L*+H. Differences in the particular phonetic realizations depend on the position of the given material in the intonation phrase. When it is pre-nuclear, the underlying L*+H pattern is realized phonetically as a gliding (slow) *f0* rise from a low target within the accented syllable up to the next syllable (if there is enough syllabic material), otherwise only within the syllable itself. When it is post-nuclear, the underlying pattern is not realized phonetically.

All the studies on intonation and information structure mentioned in this section restricted their analysis to fundamental frequency. Although very important – and some may say of

[1] Because the pitch accent is placed in the utterance-final position and is followed by the low boundary tones L-L%, it is impossible to decide which accent can be regarded as the underlying pattern.

prime importance – *f0* is only one of several acoustic properties associated with the relative prominence of words.

In this study, the question we address – mainly with reference to three of the four acoustic parameters mentioned in the introduction[2] – is whether Contemporary Standard Bulgarian distinguishes prosodically between different types of focus: (a) non-contrastive and contrastive narrow focus and (b) broad and narrow focus. Bearing in mind the literature on the prosodic realization of information structure in different focus conditions and the fact that the focus domain may be larger than the focused constituent and can affect the prosodic-acoustic realization of the whole sentence, the following hypotheses are investigated:

- Speakers produce narrow focus with more acoustic prominence than broad focus; this is reflected in the local cues.
- The prominence of the pre-nuclear and post-nuclear interval is more suppressed in narrow than in broad focus; this is reflected in the global cues.
- Speakers produce contrastive narrow focus with more acoustic prominence than non-contrastive narrow focus; this is reflected in the local cues.
- The prominence of given information is more suppressed in narrow contrastive than in narrow non-contrastive focus; this is reflected in the global cues.

3. Material and Methods

The Bulgarian data that were used in this study were taken from an existing speech corpus consisting of read speech for several languages (Andreeva et al., 2014). The stimulus material consisted of sentences, reflecting the canonical word order *subject* < *verb* < *direct object* < *indirect object* < *oblique*. This increases the role of prosody as an information-structuring factor, allowing us to focus on the acoustic correlates of different focus types. The sentences which were read aloud by our Bulgarian informants are:

1. ДИМО ДАНЕВ ГЛЕДА ДВЕ ДЕЦА.
   Dimo Danev gleda *dve* detsa.
   Dimo Danev looks after two children.
2. БАТЕ СТЕФАН ВЗЕ СЕДЕМ КНИГИ.
   Bate Stefan vze sedem knigi.
   The elder brother Stefan has taken seven books.
3. ИГРАХ НА ДАМА БЕЗ КАКА ТИ.
   Igrax na dama bez kaka ti.
   I played draughts without your older sister.
4. БАТЕ МАНИ ПИ ТЪМНА БИРА.
   Bate Mani pi tămna bira.
   The elder brother was drinking dark beer.
5. ДИМ ДАНЕВ ПЯ ТРИ ПЪТИ.
   Dim Danev pja tri păti.
   Dim Danev has sung three times.
6. КАКА НИНА ТЪРСИ ЧЕРЕН ХЛЯБ.
   Kaka Nina tărsi čeren xljab.
   The elder sister Nina is looking for dark bread.

The stimulus material was designed to investigate the acoustic realization of broad and two different types of narrow focus. Within narrow-focus realizations, we distinguished between non-contrastive and contrastive narrow focus. Appropriate questions were devised for each sentence to elicit the different focus responses (see table 1).

[2] Vowel quality differences (non-reduced vs. reduced vowels) are a feature of +/– stress at word level and are not relevant in the differentiation of +/– accents. As reported briefly below, spectral analysis (F1–F3) showed no significant differences as a function of accentuation.

Table. 1. Questions used for eliciting utterances with broad and (non-contrastive and contrastive) narrow focus

Question	Focus condition
Какво става? (*What happens?*)	broad focus
Кой търси черен хляб? (*Who is looking for dark bread?*)	non-contr. early focus on CW1
Какво търси кака Нина? (*What is elder sister Nina looking for?*)	non-contr. late focus on CW2
Кака Лина ли търси черен хляб? (*Is it elder sister Lina who is looking for dark bread?*)	contr. early focus on CW1
Кака Нина бял хляб ли търси? (*Is it white bread that elder sister Nina is looking for?*)	contr. late focus on CW2

There were two critical words (CWs) in the sentence which could be realized with prominence, one early (CW1) and one late in the sentence (CW2). These words are underlined in the list of sentences above. The five focus conditions (differing in focus type and position) elicited by different questions are: (a) a broad-focus response, (b) a response with a non-contrastive narrow focus on the early and (c) on the late CW and (d) a contrastive focus on the early and (e) on the late CW.

The sentences were read aloud from a PowerPoint presentation in response to pre-recorded questions. The sentences and the questions eliciting different focus responses were pseudo-randomized and offered in six blocks, resulting in six repetitions of each sentence for each focus condition. If the speakers were disfluent or failed to produce the correct words, they were asked to produce the utterance again. The participants in the experiment were six regionally homogeneous speakers of Contemporary Standard Bulgarian as spoken in Sofia (3 female, 3 male). All of them were born in Sofia. At the time of the recordings, they were aged between 20 and 50 years and were students or university staff. They were paid for their participation.

The stimulus material thus consists of 6 speakers × 6 sentences × 5 focus conditions × 6 repetitions = 1,080 sentences. In this paper, we present analysis results for local measurements in the CW for all sentence repetitions. We also present a more detailed analysis of the global prosodic patterns in the entire sentence for the first three of the six available repetitions.

3.1. Recordings and Processing

All recordings were made in a sound-treated studio using an AKG C420IIIPP headset and a Tascam DA-P1 DAT-recorder, and transferred digitally via the optical channel to a PC using the Kay Elemetrics MultiSpeech speech signal processing program.

Segmentation, labeling with SAMPA and further processing were done using the Kiel XASSP speech signal analysis package. Six labeling assistants, students majoring in phonetics at the Saarbrücken Institute of Phonetics who had undergone extensive training in the labeling tasks, were each allocated one of the six repetitions (a, b, c, d, e or f) of each sentence to maximize the labeling consistency across the focus conditions. They segmented and labeled all focus conditions in all repetitions of the specific sentence at the phonemic, syllable and word level. Any segmentation problems were resolved in regular discussions with the first author at group level. In addition, 120 realizations of similar material were labeled by *all six* of the labeling assistants and analyzed for labeler agreement with regard to label selection and boundary placement. The average agreement between labelers at phoneme level (label selection) reached 95.83%. The boundary placement lay within 10 milliseconds in 84.38% of all cases, showing high consistency between the labelers.

Table 2. Local and global acoustic features used for the analysis

Local measurements	Global measurements
Duration	
syllable duration	sb_dur – duration of the pre-nuclear interval
vowel duration	se_dur – duration of the post-nuclear interval
	tempo pre-nuclear – sb_dur divided by the number of syllables
	tempo post-nuclear – se_dur divided by the number of syllables
Fundamental Frequency	
vowel *f0* mean – mean *f0* across the syllable nucleus	H – *f0* maximum
	L – the minimum *f0* value preceding the peak
relative peak alignment to syllable onset	Lpost – the minimum *f0* value following the peak
f0 change – over 3 syllables (accented, preceding and following)	excursion L-H – *f0* excursion between L and H (in semitones)
	excursion H-Lpost – *f0* excursion between H and LPost (in semitones)
Energy	
vowel intensity – mean intensity of the stressed vowel	sb_int – intensity of the pre-nuclear interval
	se_int – intensity of the post-nuclear interval
vowel SpecBalance – spectral balance in the stressed vowel	
Formants	
F1, F2, F3	

In addition to segmental labeling, the pitch accents were also labeled by the first author, with the peak alignment of the L(ow) and H(igh) targets explicitly specified[3]. The positions of the *f0* maxima and minima were double-checked by an automatic procedure for which the Praat pitch tracker was used.

Local and global acoustic measures were calculated using Praat scripts and operationalized as described in the two following subsections. Table 2 summarizes all local and global acoustic features used in the statistical analyses.

3.1.1. Local Measurements

Local measurements of duration, *f0* and intensity were made in the CWs in all the sentences read aloud by the informants.

(a) Duration. The duration of the CWs, the stressed syllable of the CWs and the vowel in the stressed syllable was measured. Since all analyses were compared for different focus realizations within the same sentence, it was possible to normalize all durational measurements as a percentage of the mean duration of the corresponding unit in the sentence.

(b) Fundamental Frequency. f0 was calculated as the mean fundamental frequency [Hz] across the syllable nucleus of the lexically stressed syllable of the CW. These values were also normalized by expressing them as percentages of the mean overall *f0* of the sentence. We consider this to be a valid measure because, although the sentences include numerous consonants with consequent micro-prosodic perturbations in the vowels following them, the balanced design of the corpus assures comparable effects across focus conditions. Differences between individual lexical items in this regard are incorporated into the statistical model as a random effect.

[3] The first author was aware of the condition of the production while labeling.

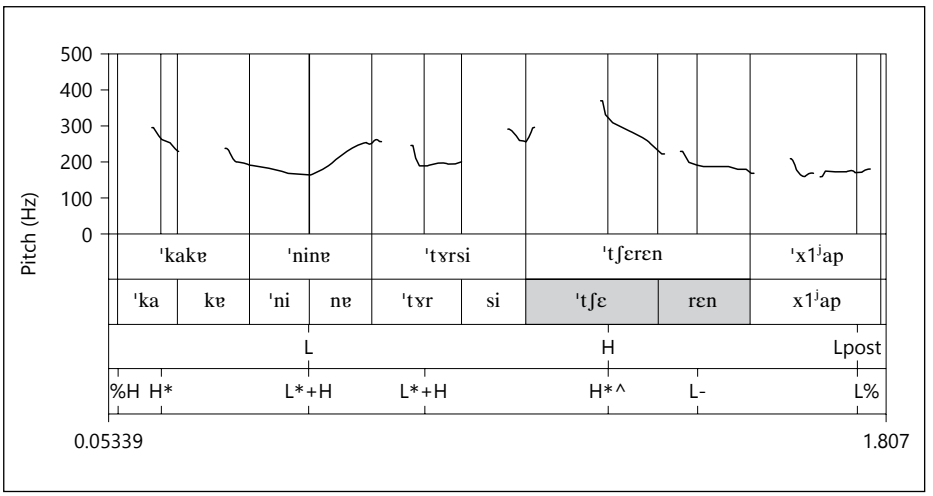

Fig. 1. Labeling example (sentence 6, non-contrastive narrow focus on CW2, highlighted in grey).

In addition, a local '*f0* change' measure was derived, namely the sum of the (absolute) differences between the average *f0* of the focus syllable and the average *f0* of the syllables preceding and following it. This measure of *f0*, computed over three syllables, was chosen to investigate how *f0* changes in the vicinity of the accented syllable might contribute to the differentiation of focus conditions.

As a measure of the peak alignment, the absolute temporal distance from the *f0* peak to syllable onset and rhyme onset were calculated. To compensate for the varying segmental durations on peak alignment, these absolute measures were converted to relative measures, taken as a proportion of syllable and rhyme durations.

(c) Energy. Energy was measured in two ways. First, as the mean intensity (dB) of the stressed vowel in the CW. These intensity values were normalized by subtracting the mean of the sentence intensity. Second, energy was measured as the spectral balance in the stressed vowel. This was computed as the difference in energy between the 70–1,000 Hz and the 1,200–5,000 Hz frequency bands.

(d) Formants. The frequency values for formants 1–3 were measured in the middle of the syllabic nucleus in the lexically stressed syllable of CWs, calculating the LPC coefficients with the algorithm by Burg in Praat using a window size of 0.025 sec and a maximum number of 5 formants.

3.1.2. Global Measurements

Global measurements of duration, speech rate, *f0* and energy were made for the first three sentence repetitions by each speaker in each condition (focus type × sentence).

(a) Tempo. Tempo was determined for a number of different units. First, the duration of each sentence, the interval from the beginning of the sentence up to the onset of the focused syllable (abbreviated as sb) and the interval starting at the end of focused syllable up to the end of the sentence (abbreviated as se) were measured. Since the number of syllables in the sentences varies, the tempo of the whole sentence, and of sb and se was computed by dividing their duration by the number of syllables in the corresponding interval.

(b) Fundamental Frequency. In addition to mean *f0* and peak alignment (section 3.1.1), the minimum *f0* value preceding (L) and following the peak (Lpost) was measured, and the pitch excursion between the preceding *f0* minimum and the peak (L-H) and between the peak and the following *f0* minimum (H-Lpost) was computed (fig. 1).

The measurements were corrected for individual *f0* differences by converting them to semitones using the following formula:

39.863 × log10 (maximum/minimum)

Mean *f0* values for sb and se were also computed and normalized by converting them to percentages of the sentence mean.

(c) Energy. The intensity of sb and se were also measured and normalized using the same procedure as for the stressed vowels.

3.2. Statistical Analysis

We used R (R Core Team, 2012) and lme4 (Bates et al., 2015) to perform a linear mixed effects analysis of the relationship between the dependent variables and the fixed factor focus condition. As random effects, we had random intercepts and slopes for speaker and item. Non-significant terms were removed. Visual inspection of residual plots did not show any obvious violation of the normality assumption or of homoscedasticity. We estimated p-values with the Satterthwaite approximation, with the help of lmerTest (Kuznetsova et al., 2014).

4. Results

In Bulgarian broad focus sentences, content words are expected to be accented. In our data, CW2 is the last content word in sentences 3 and 5, while it is followed by another content word in the remaining sentences. To determine whether the acoustic realization of sentences in which only the object (CW2) carries a narrow focus differs systematically from those in which the entire event is focused (broad focus), we analyze sentences 3 and 5 separately. 144 data points were analyzed for the local features (6 speakers × 2 sentences × 2 focus conditions × 6 repetitions) and 72 for the global features (6 speakers × 2 sentences × 2 focus conditions × 3 repetitions). To investigate whether speakers prosodically differentiate non-contrastive and contrastive narrow focus, we analyze all sentences, excluding the broad focus conditions. Sentences containing an early focus (on CW1) are analyzed separately from those containing a late focus (on CW2). Per CW 432 data points were analyzed for the local features (6 speakers × 6 sentences × 2 focus conditions × 6 repetitions) and 216 data points for the global features (6 speakers × 6 sentences × 2 focus conditions × 3 repetitions).

4.1. Broad versus Narrow Focus

The results from the statistical analysis are summarized in table 3. The reference level in the models is the broad focus condition. Significance at the level $p < 0.05$ is reported.

4.1.1. Local Cues

We first investigate the local acoustic correlates in terms of duration, *f0* and intensity in the nuclear syllable of CW2 (table 3a). The analysis reveals that in the narrow focus condition, the syllable duration is significantly longer and the accentual peak is aligned significantly later in relation to the syllable onset. Furthermore, speakers systematically indicate narrow focus with a higher mean *f0* and a greater amount of *f0* change over three syllables (from the syllable preceding to the syllable following the nuclear accented syllable). We find significantly higher values for vowel intensity and spectral balance in the narrow focus condition (fig. 2). The focus condition shows no significant effect on the first three formants, confirming observations that vowel quality differences are a function of word stress.

Table 3. Summary of the linear mixed-effects models for each of the predicted parameters in broad vs. narrow focus

Reference level	Parameter	Estimate	Std. error	t value	p value	Sig.
Broad focus condition	*a) local acoustic correlates in CW2 (nuclearly accented)*					
	syllable duration	10.866	1.520	7.150	0.001	***
	peak alignment	71.060	6.070	11.707	0.001	***
	vowel *f0* mean	35.794	3.010	11.890	0.001	***
	vowel *f0* diff	9.486	2.612	3.374	0.001	***
	vowel intensity	4.617	0.312	14.778	0.001	***
	vowel SpecBalance	2.993	0.549	5.452	0.001	***
	b) global acoustic correlates in pre-nuclear and post-nuclear interval					
	excursion LH	3.681	0.566	6.501	0.001	***
	excursion HLpost	2.039	0.488	4.180	0.001	***
	tempo pre-nuclear (sb)	–12.830	2.927	–4.384	0.001	***
	intensity pre-nuclear (sb)	–1.479	0.238	6.203	0.001	***
	intensity post-nuclear (se)	1.535	0.343	–4.481	0.001	***
	c) global acoustic correlates in CW1 (pre-nuclearly accented)					
	vowel duration	–4.071	2.009	–2.027	0.05	*
	syllable duration	–3.987	0.916	–4.355	0.001	***
	word duration	–5.353	0.720	–7.436	0.001	***
	vowel *f0* mean	–2.688	0.911	–2.952	0.003	**
	vowel *f0* change	–6.962	0.741	–9.394	0.001	***
	vowel intensity	–1.132	0.136	–8.304	0.001	***

The results from the statistical analysis provide evidence that speakers systematically differentiate between narrow and broad focus by producing the critical word in the narrow focus condition with more acoustic prominence. However, the observed phonetic differences in *f0*, duration, intensity and peak alignment might in principle reflect the use of different pitch accent types: H+!H*/H+L* for broad and (L+)H* for narrow focus (see Andreeva, 2007). We shall discuss this in section 4.3.

4.1.2. Global Measurements

Considering non-local (global) effects, we also investigated the realization of the pre-nuclear accented CW1 for broad vs. narrow focus differences (table 3b, c). Although CW1 is not expected to be de-accented in narrow focus, and the pre-nuclear pitch accents realized in broad and narrow focus conditions are expected to be identical (L*+H) (compare also Andreeva et al., 2001; Andreeva, 2007; Avgustinova and Andreeva, 1999), we observe a difference in *duration*, *pitch* and *intensity* in the pre-nuclear accented CW1, with a longer duration (for vowel, syllable and word), higher mean *f0* in the accented vowel, greater amount of *f0* change and higher intensity in broad than in narrow focus. Additionally, speakers produce narrow focus with significantly larger pitch excursion (for both L-H and H-Lpost), lower tempo and intensity in the pre-nuclear interval (sb) and higher intensity in the post-nuclear interval (se). This indicates that speakers suppressed the prominence of the pre-nuclear

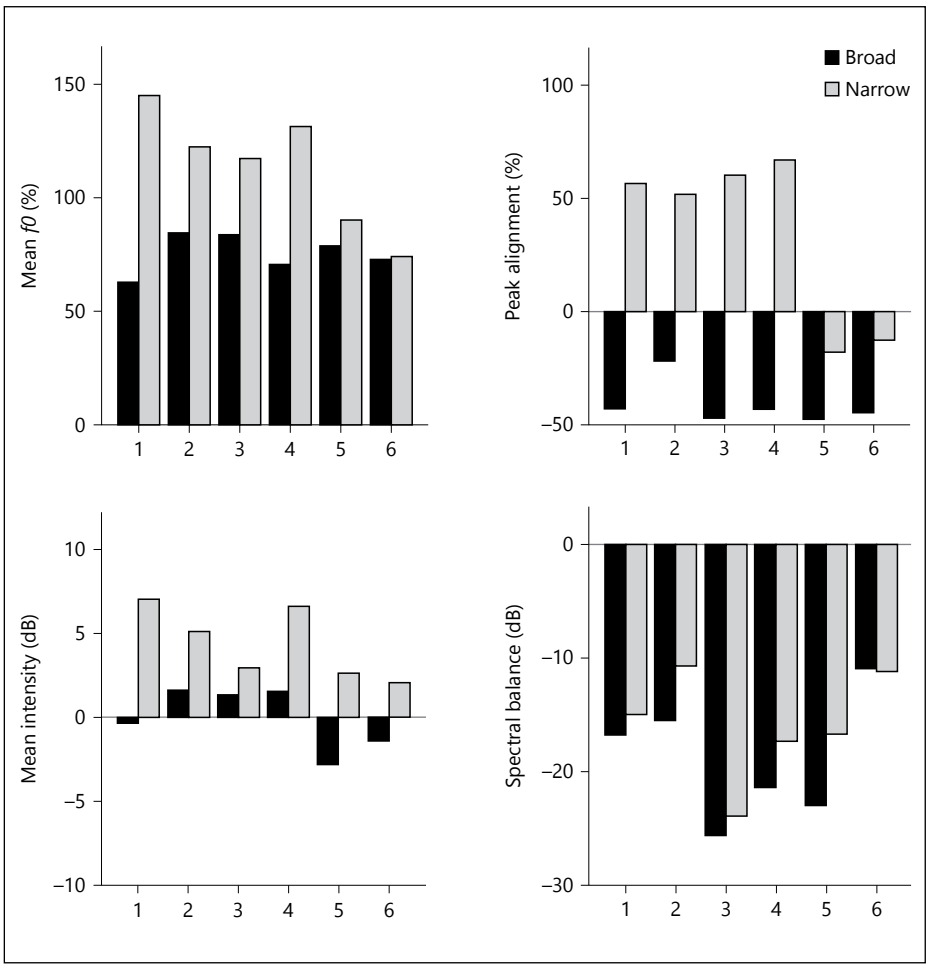

Fig. 2. Normalized values for mean fundamental frequency, peak alignment, mean intensity and spectral balance, broken down by speaker.

interval (including the pre-nuclear accented word) in the narrow focus condition, with the result that the utterances realized under narrow focus exhibit an increase in prominence-lending cues from pre-nuclear to nuclear accents (fig. 3). In other words, the phonetic effects triggered by focus should also be seen in relation to pre-nuclear accents. This finding is in line with those obtained by Calhoun (2006), Breen et al. (2010) and Katz and Selkirk (2011), among others. The fact that the intensity in the post-nuclear interval is higher in narrow focus than in broad focus seems to contradict the hypothesis that the prominence of post-nuclear material is more suppressed in narrow than in broad focus and that this is reflected in the global acoustic cues. This outcome is explained by the presence of final vowel devoicing in Bulgarian reported by Andreeva and Koreman (2008). The authors found that towards the end

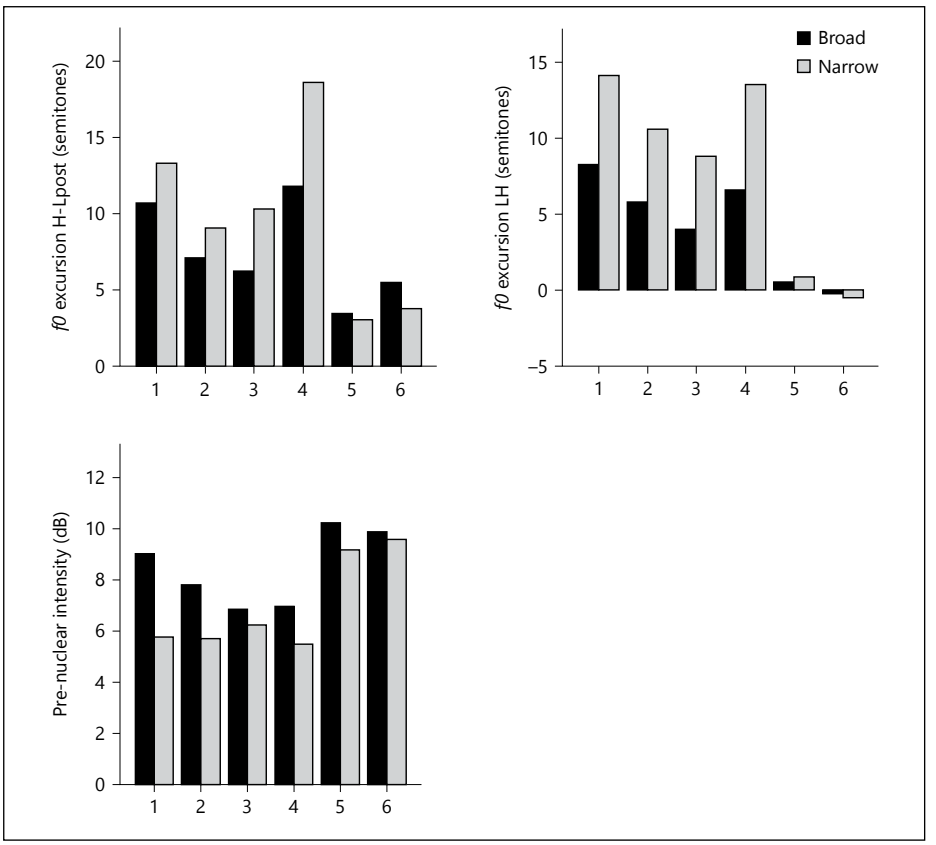

Fig. 3. Normalized values for *f0* excursion H-Lpost and L-H in semitones, and intensity in the pre-nuclear interval, broken down by speaker.

of the utterance, the peak scaling and alignment exert a strong influence on the vowel realization, since earlier and lower peaks tend to allow for more extensive vowel devoicing than later and higher peaks.

4.2. Contrastive versus Non-Contrastive Focus

The results from the statistical analysis are summarized in table 4. The reference level in the models is the narrow contrastive focus condition. Significance at the level $p < 0.05$ is reported.

4.2.1. Local Measurements

The statistical analysis indicates that in terms of local acoustic parameters in the nuclear-accented syllable, speakers do not systematically differentiate between contrastive and non-contrastive focus condition in the late sentence position. When the focus is realized on CW1, speakers do not use *f0* and intensity to discriminate between non-contrastive and contrastive focus, but they produce systematically longer vowel, syllable and word durations in the contrastive focus condition (table 4a).

Table 4. Summary of the linear mixed-effects models for each of the predicted parameters for contrastive vs. non-contrastive focus

Reference level	Parameter	Estimate	Std. error	t value	p value	Sig.
Narrow contrastive focus condition	*a) local acoustic correlates in CW1 (nuclearly accented)*					
	vowel duration	–4.613	1.303	–3.539	0.001	***
	syllable duration	–5.113	1.812	–2.822	0.005	**
	word duration	–4.769	1.527	–3.122	0.003	**
	b) global acoustic correlates in CW1 (pre-nuclearly accented)					
	vowel *f0* diff	1.342	0.649	2.070	0.04	*

4.2.2. Global Measurements

No systematic differences were found between the global measurements for contrastive vs. non-contrastive focus when the focus is realized *early* in the sentence, i.e., on CW1. The statistical analysis indicates that when the focus is realized *late* in the sentence, i.e., on CW2, the *f0* change on CW1 is greater in the non-contrastive focus condition (table 4b). While CW2 in the two focus conditions is realized identically in both focus conditions (see section 4.2.1), reduction of the prosodic prominence of CW1 in contrastive utterances makes CW2 more prominent. These results suggest that a clear difference between the two focus conditions emerges, if the phonetic realization of the pre-nuclear accent is taken into account.

4.3. Speaker Variation in the Phonological Choice of the Pitch Accent Type

The statistical analyses of the local prosodic cues in section 4.1.1 and section 4.2.1, reveal that the speakers differentiate systematically between the different focus conditions. They use longer duration and higher pitch and intensity in the narrow-focused word compared to the corresponding word in the broad focus condition, and longer duration in the contrastive focus condition compared to the non-contrastive when the focused word is early in the sentence. But it is important to ascertain whether these differences are the result of gradient prosodic processes or the selection of categorically different pitch accent types for focus marking. This section will, therefore, examine the pitch accent types used across all conditions.

In our data, narrow focus is mostly realized with H* pitch accents. When the focus is realized on CW1, the H target is reached close to the end of the accented syllable in 57%, in the middle of the accented syllable in 23% and close to the beginning of the accented syllable in 20% of the cases. When the focus is realized on CW2, the H target is reached close to the beginning of the accented syllable in 75%, in the middle of the accented syllable in 21% and close to the end of the accented syllable in 4% of the cases (fig. 4). Speakers also use L*+(!)H, L+H*, H+!H*, H+L* and in very few cases L* (4 times) and !H* (3 times). Note that the use of H+!H*, H+L*, L*+H and L* in the narrow focus condition is not described by Andreeva et al. (2001) and Andreeva (2007); their use here requires further investigation of possible pragmatic meanings.

In the broad focus condition, we consistently find a H+!H* or H+L* accent with early peak alignment on the pre-tonic syllable. With respect to the narrow focus condition, speakers vary as to their preferred choice of phonologically specified accent types

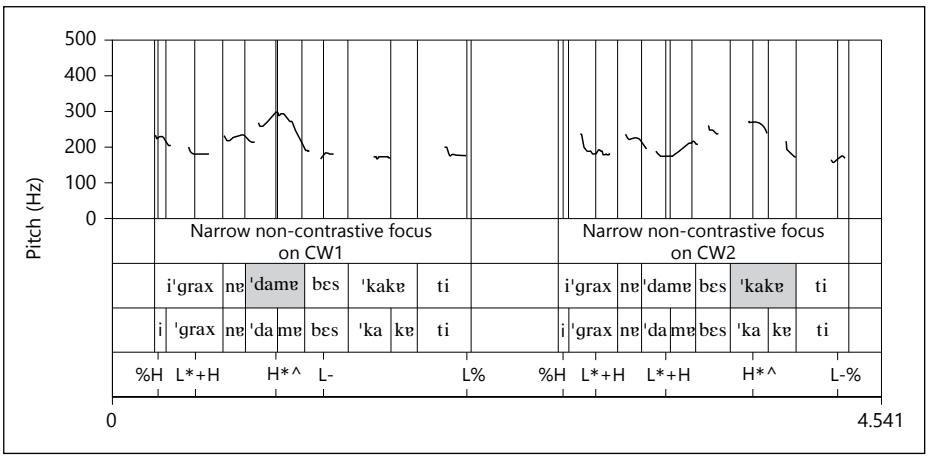

Fig. 4. Test sentence produced by female speaker 1 in non-contrastive narrow focus on CW1 (left) and on CW2 (right). Both CWs are highlighted in grey.

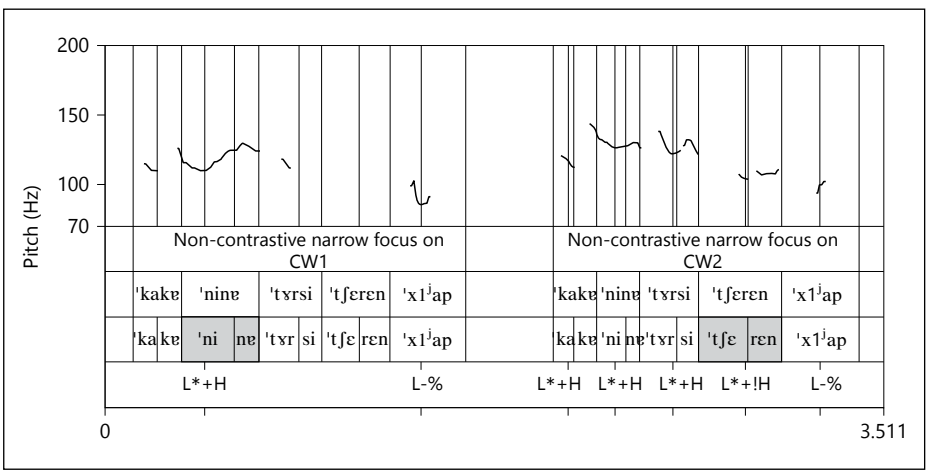

Fig. 5. Test sentence produced by male speaker 5 in non-contrastive narrow focus on CW1 (left) and on CW2 (right). Both CWs are highlighted in grey.

and their phonetic realization. They seem to be insensitive to the focus type and make the same phonological choice for contrastive and non-contrastive focus. SP6 exclusively uses H+!H* or H+L*, regardless of the position within the sentence (CW1 or CW2). SP5 has a strong preference for (downstepped) rising nuclear accents (L*+(!) H), again regardless of the position within the sentence (fig. 5). SP4 uses L+H* on CW1 and H* on CW2. Speakers 1, 2 and 3 use H*, regardless of the position within the sentence. The percentage of the pitch accent types used in the different focus conditions is summarized in table 5.

Table 5. Pitch accent types distribution per focus condition

Focus type	Accent type					
	H+!H*	H+L*	H*	L+(!)H*	L*+(!)H	L*
CW2 broad	62.5	37.5	–	–	–	–
CW2 non-contrastive	4.6	15.7	63.9	0.5	13.4	1.9
CW2 contrastive	7.9	14.4	59.7	2.8	13.4	1.9
CW1 non-contrastive	13.4	–	53.7	15.3	17.6	-
CW1 contrastive	17.2	0.5	50.7	14.4	17.2	-

Additionally, we examined the realizations of CW1 when the focus is late in the sentence, i.e., on CW2. It turns out that in all three focus conditions (contrastive, non-contrastive and broad), the stressed syllable of CW1 bears a pitch accent L*+H, where the L tone is aligned within the accented syllable and the trailing H tone is aligned in the first post-tonic syllable or slightly after it. This accent type is described as the *default* pitch accent in pre-nuclear position (Andreeva, 2007), but L* and L+H* were also observed. When the nuclear accent is placed on CW1, i.e., when narrow focus occurs early in a sentence, all following words are de-accented.

The analyses reported in this section support the assumption made in section 4.1.1, to some extend namely that the differences in the analyzed local acoustic cues in the broad and narrow focus condition reflect the use of different pitch accent types. In the case of narrow contrastive and non-contrastive focus, we found that even when speakers differ in their choice of accent types, the number of different accent types in the two focus conditions is the same. We therefore feel justified in claiming that, in the contrastive and non-contrastive conditions, we observe the *phonetic variation* of phonological categories. Speakers boost the accent type with longer durations, to signal the contrast.

5. Discussion and Conclusions

In this study, we explored the prosodic realizations of information structure categories in Bulgarian declarative sentences, drawing particular attention to the specific contribution of global and local cues. The main question we asked is whether Contemporary Standard Bulgarian distinguishes prosodically between: (a) broad and narrow focus, and (b) non-contrastive and contrastive narrow focus.

With regard to the difference between *broad* focus and *narrow* focus on CW2, it was found that narrow-focused syllables in CW2 were consistently realized with a longer duration, later peak alignment (but still early in the syllable), greater mean *f0* and greater pitch change (including the previous and next syllable), higher intensity and spectral balance than syllables with broad focus. This finding is not surprising, since all the subjects but one used different pitch accent types to signal narrow vs. broad focus: predominantly H* vs. H+!H*/H+L*. These results generally support our first hypothesis that speakers produce narrow focus with more acoustic prominence than broad focus, and that this is reflected in the local cues. To be clear, the local cues for this comparison turn out to be the different pitch accent types and their corresponding acoustic parameters. Our second hypothesis was that the prominence of the pre-nuclear (sb) and post-nuclear (se) intervals (i.e., non-local properties) is more

suppressed in narrow focus than in broad focus, because that is where the 'given' information is provided and this should be reflected in the strength of the global cues. In agreement with the results from previous research (Andreeva et al., 2001; Andreeva, 2007; Avgustinova and Andreeva, 1999), no phonological de-accentuation of the pre-nuclear CW1 was found when the narrow focus was on CW2. Broad and narrow focus conditions were not distinguished by the accent type realized on CW1. The default pre-nuclear L*+H was used for both. However, the non-focused, pre-nuclear interval in the narrow focus condition differs from the focused, pre-nuclear interval in the broad focus condition in terms of global acoustic properties. In responses with broad focus, the pitch excursion between the pre-nuclear interval (L-H) and the focused syllable is lower, which reflects the use of different pitch accent types, and the pre-nuclear interval has a lower speaking rate and higher intensity than responses with narrow focus on CW2. Equivalent differences are also found in the pre-nuclear CW1 accented vowel itself. Speakers consistently produce this vowel in the broad focus condition with a higher *f0* mean value and higher *f0* change, longer duration and higher intensity. In other words, speakers seem to employ both local cues for the nuclear accented CW2 and dispersed (global) cues in the pre-nuclear accented interval, including CW1 in a way that boosts the prosodic prominence on the later narrow focus nuclear accented syllable. The pitch excursion between the accented syllable and the post-nuclear interval (H-Lpost) is smaller in the broad focus condition. The interval following the nuclear accent has a higher intensity in responses with broad focus than in responses with narrow focus on CW2. This finding does not support our hypothesis, but it is consistent with the observed relationship between post-nuclear vowel devoicing and pitch accent type in Bulgarian (Andreeva and Koreman, 2008). Towards the end of the utterance, earlier and lower peaks tend to allow for a stronger vowel devoicing since the transglottal pressure has dropped so low by the time the coda is reached that vocal fold vibration can no longer be sustained.

With regard to our third hypothesis that speakers produce *contrastive* narrow focus with more acoustic prominence than *non-contrastive* narrow focus in terms of local acoustic parameters, we observed that only when focus occurs *early* in the utterance, i.e., on CW1, was contrastive focus marked more prominently than non-contrastive focus, namely by longer vowel, syllable and word durations. Since our analysis in section 4.3 showed that speakers use similar proportions of different accent types in both focus conditions, this indicates phonetic variation of phonological categories being used to reinforce the contrastive condition. This is not true for focus on CW2, where no such differences were observed. This result partially confirms our third hypothesis. Our results also partially support our last hypothesis, namely that the prominence of given information is suppressed more in narrow contrastive than in narrow non-contrastive focus, and that this is reflected in the global cues. When CW1 is focused, all content words in the post-nuclear interval are de-accented and no systematic differences are found between the global measurements for contrastive vs. non-contrastive focus. But when the focus is realized *late* in the sentence, i.e., on CW2, *f0* change, reflecting the realization of the pre-nuclear accent L*+H on CW1, is greater in the non-contrastive focus condition. We can conclude that the difference between contrastive and non-contrastive focus condition is realized by local and global cues in a mutually compensatory way, as a function of focus position: by longer duration of the focus word when focus is early in the sentence, or by a decrease of the prominence in the pre-nuclear interval when focus is late in the sentence. Our results replicate our previous findings and relativizes

Penčev's (1980) general claim about the necessary de-accentuation of the given material in Bulgarian utterances with contrastive focus. The greater duration found in our data and the enhanced tonal contrast between the nuclear prominence of CW2 and the pre-nuclear prominence of CW1 may be seen as a covariant of the greater pitch range for contrastive-focused syllables reported by Penčev. Our results are, however, not consistent with the findings for another Slavonic language, Russian (Mehlhorn, 2002; Meyer and Mleinek, 2008), where the contrastive-focused elements are higher and have a different accent type (L+H* compared to (!)H* for the non-contrastive focus).

The findings in our study are in line with the results reported by Baumann et al. (2006, 2007) for German and to *some* extent with the results of Breen et al. (2010) for English. Breen at al. (2010) obtained their results within an analysis framework known in the literature as a *direct-relationship* approach between acoustics and meaning (Cooper et al., 1985; Fry, 1955; Lieberman, 1960; Xu and Xu, 2005). However, they pointed out that these results are also consistent with the assumption of mediating phonological categories (pitch accents), as in the intonational phonology framework which underlies this study. Under this assumption, and on the basis of their results, they suggest principles for mapping acoustics to meaning: Focused constituents receive a pitch accent. In narrow focused SVO sentences with focus on the object, there will be only one accent, namely on the narrow-focused word, that leads to higher values of the acoustic measures on the object compared to constituents in the pre-nuclear interval. In the broad focus condition, each of the focused constituents (subject, verb, object) receives an accent, resulting in more similar acoustic measures on the nuclear and pre-nuclear accented constituents. The different accent patterns in the production of broad and narrow focus would support the successful recognition of the two focus conditions. In our data, the two conditions are not distinguished by accent placement, but rather in the choice of nuclear accent type and in the acoustic values of the same accent type in the pre-nuclear interval. But the difference between the nuclear and pre-nuclear intervals follows the pattern postulated by Breen and colleagues: they decrease in the broad focus condition and increase in the narrow focus condition.

With regard to accent placement, it has to be said that evidence from other languages also suggests that given information does not necessarily have to be de-accented (for a cross-linguistic investigation on de-accenting and re-accenting, see Cruttenden, 2006; for German: Baumann et al., 2007, 2015; and Féry und Kügler, 2008; for English: Calhoun, 2010; for Romanian: Ladd, 2008; and Manolescu et al., 2009; for Spanish: Hualde, 2002; among many others). For a language-independent theory of acoustics-to-focus-meaning mapping, accentuation vs. de-accentuation would be insufficient. It would be necessary, as Baumann et al. (2015) stress, to differentiate prosodic prominence more finely on the basis of different accent positions and accent types to achieve a satisfactory (de-)coding of the information status of a constituent.

To conclude, the all-important function of intonation, namely to transmit the relative weighting of information in speech communication, cannot be captured by a purely phonological description of realized accent types or by analyzing only the phonetic implementation of *f0*, particularly when the analysis is restricted to the nuclear accents. Crucially, the information-structure-related patterns of phonetic prominence which are revealed in this study show a complex interplay between phonological categories and multiple local and global phonetic signal properties.

The present study replicates and extends previous research on Contemporary Standard Bulgarian. It shows that not only *f0,* but also other prosodic cues play an

important role in distinguishing different focus types. Clearly, these cues are not restricted to the syllable which carries the focus; more global characteristics are important cues for signaling the information structure.

This is in itself not a completely new insight, although its central importance to the functioning of prosody in speech communication is rarely highlighted to the extent it is here. But it also offers an explanation for the apparently contradictory results that are found in the literature, some cases of which have found mention in the discussions above. A difference in one study between broad and narrow focus, or between contrastive and non-contrastive focus, found in a 'local' *f0* analysis of the nuclear accents, can no longer be branded as contradictory to another study which found no difference unless the lack of difference extends to all the potentially – local and global – differentiating parameters in the complete intonation phrase. This has certainly not been the case in the majority of earlier studies we have discussed.

One important caveat in this conclusion must be that differences must be shown to be perceptually relevant. The present study has not included perceptual investigation of the parameters analyzed, and while the degrees of difference reported are known to be perceptible in terms of discrimination, there is clearly a need to pursue the question of functional perceptibility within the information structure context.

References

Alter K, Mleinek I, Richter N (2001): Prosodic phrasing and accentuation in Russian; in Zybatow G, Junghanns U, Mehlhorn G, Szucsich L (eds): Current Issues in Formal Slavic Linguistics (= Linguistik International; 5). Frankfurt/Main, Peter Lang, pp 317–328.

Andreeva B, Avgustinova T, Barry WJ (2001): Link-associated and focus-associated accent patterns in Bulgarian; in Zybatow G, Junghanns U, Mehlhorn G, Szucsich L (eds): Current Issues in Formal Slavic Linguistics (= Linguistik International; 5). Frankfurt/Main, Peter Lang, pp 353–364.

Andreeva B (2007): Zur Phonetik und Phonologie der Intonation der Sofioter-Varietät des Bulgarischen, PHONUS 12, Institute of Phonetics, University of the Saarland, Saarbrücken, PhD theses.

Andreeva B, Koreman J (2008): The status of vowel devoicing in Bulgarian: phonetic or phonological; in Zybatow G, Junghanns U, Mehlhorn G, Szucsich L (eds): Formal Description of Slavic Languages: The Fifth Conference. Frankfurt/Main, Peter Lang, pp 81–91.

Andreeva B, Barry W, Koreman J (2014): A Cross-Language Corpus for Studying the Phonetics and Phonology of Prominence. Reykjavik, The 9th Edition of the Language Resources and Evaluation Conference (LREC 2014), pp 26–31.

Arvaniti A, Garding G (2007): Dialectal variation in the rising accents of American English; in Cole J, Hualde JI (eds): Papers in Laboratory Phonology 9. Berlin/New York, Mouton de Gruyter, pp 547–576.

Avgustinova T (1997): Word Order and Clitics in Bulgarian. Saarbrücken Dissertations in Computational Linguistics and Language Technology. Vol 5.

Avgustinova T, Andreeva B (1999): Thematic intonation patterns in Bulgarian Clitic Replication. Proceedings of the 14th ICPhS, San Francisco, pp 1501–1504.

Barry W, Andreeva B, Russo M, Dimitrova S, Kostadinova T (2003): Do rhythm measures tell us anything about language type; in Recasens D, Solé MJ, Romero J (eds): Proceedings of the 15th ICPhS, Barcelona, pp 2693–2696.

Bartels C, Kingston J (1994): Salient pitch cues in the perception of contrastive focus; in Boach P, van der Sandt R (eds): Focus and Natural Language Processing. Proceedings of J Sem Conference on Focus. IBM Working Papers. TR-80, pp 94–106.

Bates D, Mächler M, Bolker B, Walker S (2015): Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.

Baumann S, Grice M, Steindamm S (2006): Prosodic marking of focus domains – categorical or gradient? Proceedings of Speech Prosody, Dresden, Germany, pp 301–304.

Baumann S, Becker J, Grice M, Mücke D (2007): Tonal and articulatory marking of Focus in German; in Tourvain J, Barry WJ (eds): Proceedings of the 15Ith ICPhS, Pirrot GmbH, Dudweiler, pp 1029–1032.

Baumann S, Röhr CT, Grice M (2015): Prosodische (De-)kodierung des informationsstatus im deutschen. Zeitschrift für Sprachwissenschaft 34:1–42.

Beckman ME (1986): Stress and Non-Stress Accent. Netherlands Phonetic Archives, Series No. 7. Foris.
Bertinetto PM (1981): Strutture prosodiche dell'italiano. Firenze, Accademia della Crusca.
Birch S, Clifton C Jr (1995): Focus, accent, and argument structure: effects on language comprehension. Lang Speech 38:365–391.
Bojadžiev T, Kutsarov I, Penčev J (1999): Săvremenen bălgarski ezik. IK Petăr Beron, Sofija.
Bolinger D (1961): Contrastive accent and contrastive stress. Language 37:83–96.
Braun B (2005): Production and perception of thematic contrast in German. Oxford, Peter Lang.
Breen M, Fedorenko E, Wagner M, Gibson E (2010): Acoustic correlates of information structure. Lang Cogn Process 25:1044–1098.
Braun B (2015): What causes the activation of contrastive alternatives, the size of focus domain or pitch accent type? Proceedings of the 18th ICPhS. Glasgow, UK.
Bruce G (1977): Swedish word accents in sentence perspective. CWK Gleerup, Lund.
Büring D (2007): Intonation, semantics and information structure; in Ramchand G, Reiss C (eds): The Oxford Handbook of Linguistic Interfaces. Oxford, Oxford University Press.
Calhoun S (2006): Information Structure and the Prosodic Structure of English: A Probabilistic Relationship. PhD Thesis, University of Edinburgh.
Calhoun S (2010): How does informativeness affect prosodic prominence? Lang Cogn Process 25:1099–1140.
Chafe W (1976): Givenness, contrastiveness, definiteness, subjects, topics and points of view; in Li CN (ed): Subject and Topic. Academic Press, pp 27–55.
Cho T (2005): Prosodic strengthening and featural enhancement: evidence from acoustic and articulatory realizations of /ɑ, i/ in English. J Acoust Soc Am 117:3867–3878.
Cooper WE, Eady SJ, Mueller PR (1985): Acoustical aspects of contrastive stress in question-answer contexts. J Acoust Soc Am 77:2142–2156.
Cruttenden A (2006): The de-accenting of given information: a cognitive Universal; in Bernini G, Schwartz ML (eds): Pragmatic Organization of Discourse in the Language of Europe. Mouton de Gruyter, The Hague, pp 311–355.
Cutler A (1977): The Context-Independence of 'Intonational Meaning'. Chicago Linguistic Society (CLS 13), pp 104–115.
Dauer R (1987): Phonetic and phonological components of language rhythm. Proceedings of the 11th ICPhS. Tallinn, Estonian Academy of Sciences, vol 5, pp 447–450.
Dimitrova S (1998): Bulgarian speech rhythm: stress-timed or syllable-timed? J Int Phonet Assoc 27:27–33.
D'Imperio M (1997): Narrow focus and focal accent in the Neapolitan variety of Italian; in Botinis A, Kouroupetroglou G, Carayiannis G (eds): Intonation: Theory, Models and Applications. Proceedings of an ESCA Workshop, Athens, Greece, pp 87–90.
Dyer D (1992): Word Order in the Bulgarian Simple Sentence: A Study in Grammar, Semantics and Pragmatics. Amsterdam, Rodopi.
Eady SJ, Cooper WE (1986): Speech intonation and focus location in matched statements and questions. J Acoust Soc Am 80:402–415.
Féry C, Krifka M (2008): Information structure – notional distinctions, ways of expression; in Sterkenburg PV (ed): Unity and Diversity of Languages. Amsterdam, John Benjamins, pp 123–136.
Féry C, Kügler F (2008): Pitch accent scaling on given, new and focused constituents in German. J Phon 36:680–703.
Frota S (2000): Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation. New York, Garland Publishing.
Fry DB (1955): Duration and intensity as physical correlates of linguistic stress. J Acoust Soc Am 27:765–768.
Georgieva E (1974): Slovored na prostoto izrečenie v bălgraskija knižoven ezik. Sofia, BAN.
Grice M (1995): The intonation of Interrogation in Palermo Italian; implications for intonation theory. Tübingen, Niemeyer.
Grice M, Baumann S, Jagdfeld N (2009): Tonal association and derived nuclear accents: the case of downstepping contours in German. Lingua 119:881–905.
Gussenhoven C (1983): Testing the reality of focus domains. Lang Speech 26:61–80.
Halliday M (1967): Intonation and grammar in British English. The Hague, Mouton.
Hamlaoui F, Żygis M, Engelmann J, Wagner M (2015): Acoustic correlates of focus marking in Polish. Proceedings of the 18th ICPhS.
Hanssen J, Peters J, Gussenhoven C (2008): Prosodic effects of focus in Dutch declaratives. Proc Speech Prosody 2008:609–612.
Hualde JI (2002): Intonation in romance: introduction to the special issue. Probus 14:1–7.
Ito K, Speer SR, Beckman ME (2004): Informational status and pitch accent distribution in spontaneous dialogues in English. Proceedings of the International Conference on Spoken Language Processing, Nara, Japan, pp 279–282.
Ivančhev S (1957/1978): Nabljudenija vărxu upotrebata na chlena v bălgarskija ezik; in Pašov P (ed): Pomagalo po Bălgarska Morfologija. Imena. Sofia, Nauka i Izkustvo, pp 186–211.
Jackendoff R (1972): Semantic Interpretation in Generative Grammar. Cambridge, MIT Press.
Katz J, Selkirk E (2011): Contrastive focus vs. discourse-new: evidence from phonetic prominence in English. Language 87:771–816.

Kochanski G, Grabe E, Coleman J, Rosner B (2005): Loudness predicts prominence: fundamental frequency lends little. J Acoust Soc Am 118:1038–1054.
Koreman J, Andreeva B, Barry W (2008): Accentuation cues in French and German; in Barbosa PA, Madureira S, Reis C (eds): Proceedings of the 4th International Conferences on Speech Prosody. Campinas, Editora RG/CNPq, pp 613–616.
Koreman J, Andreeva B, Barry WJ, van Dommelen W, Sikveland RO (2009): Cross-language differences in the production of phrasal prominence in Norwegian and German; in Vainio M, Aulanko R, Aaltonen O (eds): Nordic Prosody. Proceedings of the Xth Conference, Helsinki 2008. Frankfurt/Main, Peter Lang, pp 139–150.
Krahmer E, Swerts M (2001): On the alleged existence of contrastive accents. Speech Commun 34:391–405.
Kratzer A (2004): Interpreting focus: presupposed or expressive meanings? Theoretical Linguistics (30(1), special issue on Interpreting Focus, pp 123–136.
Kügler F (2008): The role of duration as a phonetic correlate of focus; in Barbosa PA, Madureira S, Reis C (eds): Proceedings of the Speech Prosody 2008 Conference. Campinas, Editora RG/CNPq, pp 591–594.
Kuznetsova A, Brockhoff PB, Christensen RHB (2014): lmertest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package): R package version 2.0-25.
Ladd RD (2008): Intonational Phonology. Cambridge, Cambridge University Press.
Leafgren J (2002): Degrees of Explicitness: Information Structure and the Packaging of Bulgarian Subjects and Objects. Pragmatics and Beyond New Series 102. Amsterdam, John Benjamins.
Lieberman P (1960): Some acoustic correlates of word stress in American English. J Acoust Soc Am 32:451–454.
Manolescu A, Olson D, Ortega-Llebaria M (2009): Cues to contrastive focus in Romanian; in Vigário M, Frota S, Freitas MJ (eds): Interactions in Phonetics and Phonology. Amsterdam/Philadelphia, John Benjamins.
Mathesius V (1907): Studie k dějinám anglického slovosledu. Věstník České Akademie 16:261–274, 17:195–214, 299–311.
Mathesius V (1939): O tak zvaném aktuálním členění větném. Slovo Slovesnost 5:171–174.
Mathesius V (1947): Podstata aposice a její druhy; in Mathesius V (ed): Čeština a Obecný Jazykozpyt. Praha, pp 302–318.
Mehlhorn G (2002): Kontrastierte Konstituenten im Russischen. Experimentelle Untersuchungen zur Informationsstruktur. Frankfurt/Main, Peter Lang.
Meyer R, Mleinek I (2008): How prosody signals force and focus a study of rise-fall accents in Russian yes-no questions; in Zybatow G, Junghanns U, Mehlhorn G, Szucsich L (eds): Proceedings of the Conference FDSL-5, Leipzig 2003. Frankfurt/Main, Peter Lang.
Miševa A (1991): Intonacionna Sistema na Bălgarskija Ezik. Sofija, Bălgarska Akademija na Naukite.
Miševa A, Nikov M (1998): Intonation in Bulgarian; in Hirst D, Di Cristo A (eds): Intonation Systems: A Survey of 20 Languages. Cambridge, Cambridge University Press, pp 275–287.
Molnár V (2002): Contrast – from a contrastive perspective; in Hasselgård H, et al (eds): Information Structure in a Cross-Linguistic Perspective. Amsterdam/New York, Rodopi, pp 147–161.
Nikolaeva TM (1977): Frazovaja intonacija slavjanskix jazykov. Moscow, Nauka.
Paul H (1937): Prinzipien der Sprachgeschichte. [Reprint of ed 5. Niemeyer, Halle.].
Penčev J (1980): Osnovni intonacionni konturi v bălgarskoto izrečenie. Bălgarska Akademija na Naukite, Sofija.
R Core Team (2012): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/.
Rochemont MS (1986): Focus in Generative Grammar. Amsterdam/Philadelphia, John Benjamins.
Rooth M (1992): A theory of focus interpretation. Nat Lang Seman 1:75–116.
Rudin C (1985): Aspects of Bulgarian syntax: complementizers and wh-constructions. Columbus, Slavica Publishers.
Rump HH, Collier R (1996): Focus conditions and the prominence of pitch accented syllables. Lang Speech 39:1–17.
Selkirk E (2002): Contrastive focus vs. presentational focus: prosodic evidence from right node raising in English; in Bel B, Isabelle M (eds): Proceedings of the 1st International Conference on Speech Prosody. France, Aix-en-Provence, pp 643–646.
Sityaev D, House J (2003): Phonetic and phonological correlates of broad, narrow and contrastive focus in English. Proceedings of the 15th ICPhS, pp 1819–1822.
Szober S (1933): Edin osnoven tip nareždane dumite v bălgarskoto izrečenie; Reprinted in Popov K (1979): Pomagalo na bălgarski sintaksis. Sofia, Nauka i Izkustvo, pp 275–285.
Sluijter AM, van Heuven VJ (1996): Spectral balance as an acoustic correlate of linguistic stress. J Acoust Soc Am 100(4 pt 1):2471–2485.
Smiljanić R (2004): Lexical, Pragmatic, and Positional Effects on Prosody in Two Dialects of Croatian and Serbian: An Acoustic Study. Outstanding Dissertations in Linguistics. New York, Routledge.
Stojkov St (1966): Uvod văv Fonetikata na Bălgarskija Ezik. Sofija, Nauka i Izkustvo.
't Hart J, Collier R, Cohen A (1990): A Perceptual Study of Intonation. Cambridge, Cambridge University Press.
Tilkov D (1981): Intonacijata v Bălgarskija Ezik. Sofija, Narodna Prosveta.
Tilkov D, Bojadžiev T (1990): Bălgarska Fonetika, ed 3. Sofija, Nauka i Izkustvo.
Trávniček F (1962): O tak zvanem aktualnim cleneni vetnem. Slovo Slovesnost 22:163–171.
Turk AE, Sawusch JR (1996): The processing of duration and intensity cues to prominence. J Acoust Soc Am 99:3782–3790.

Vainio M, Järvikivi J (2007): Focus in production: tonal shape, intensity and word order. J Acoust Soc Am 121:EL55–EL61.
Vallduví E, Vilkuna M (1998): On rheme and contrast; in Culicover PW, McNally L (eds): The limits of Syntax (Syntax & Semantics 29). San Diego, Academic Press, pp 79–108.
Welby P (2003): Effects of pitch accent position, type, and status on focus projection. Lang Speech 46(pt 1):53–81.
Xu Y, Xu CX (2005): Phonetic realization of focus in English declarative intonation. J Phon 33:159–197.
Zybatow G, Mehlhorn G (2000): Experimental evidence for focus structure in Russian; in King TH, Sekerina IA (eds): Formal Approaches to Slavic Linguistics: The Philadelphia Meeting 1999. Ann Arbor, pp 414–434.