Information Structure in Polish and Bulgarian: Accent Types and Peak Alignment in Broad and Narrow Focus.

Chapter · January 2005

CITATIONS
3

READS
56

2 authors, including:

Bistra Andreeva
Universität des Saarlandes
87 PUBLICATIONS **552** CITATIONS

SEE PROFILE

Information Structure in Polish and Bulgarian: Accent Types and Peak Alignment in Broad and Narrow Focus: A Cross-Language Study.

Bistra Andreeva and Dominika Oliver
Institute of Phonetics, Saarland University, Saarbrücken

1. Introduction

1.1. Background

For over thirty years the levels-vs.-configuration debate has been a very controversial issue for research in intonational phonology. Earlier traditions, such as that of the British school (e.g. Crystal, 1969, O'Connor and Arnold, 1973) describe the distinctive units of intonation in a holistic way in terms of complex moves of the contour. These configurations or moves (fall, rise, rise-fall etc.) are associated with an intonational phrase as a whole. The autosegmental-metrical approach to intonation (see Ladd, 1996 for an overview) argues against the configurations as a primitives of the linguistic analysis and analyses the intonation contour as a sequence of phonological level tones such as H(ighs) and L(ows) or a combination of the two, occurring at specific structural positions in the utterance. For example, a rising f_0 movement is, in this view, taken as merely a transition from its beginning point (f_0 minimum value) to its ending point (f_0 maximum value). It is usual to refer to these points as "tonal targets" which can be defined in terms of two dimensions, i.e., "alignment" and "scaling".

Tonal alignment can be defined as the temporal synchronization of tones with some specific segments or prosodic locations (such as syllable onset, syllable offset or rhyme onset) and can be related to phonological and/or phonetic factors. Phonological factors are qualitative and categorical (e.g. align target with syllable X rather than syllable Y) and imply different accent patterns (L*+H vs. L+H*). Phonetic factors are gradient and can often be modelled by means of interacting quantitative parameters (e.g. align target ear-

lier the closer it is to the next target). These factors additionally "fine-tune" the alignment of tonal targets, determining the differing phonetic realisation of the same phonological tones.

A number of recent studies have reported that pragmatic information such as information structure and sentence mode play a crucial role in the timing of tonal alignment (Kohler, 1987, Miševa, 1991, Frota, 2000 among others). In these studies, the tonal targets are claimed to appear at different locations with respect to the segmental references depending on the word's or sentence's pragmatic status (e.g., broad vs. narrow focus, contrastive vs. non-contrastive focus, statement vs. question).

Other studies (Silverman and Pierrehumbert, 1990, Arvaniti et al., 1998, Ladd et al. 2000 among others) have suggested that the specification for the alignment of tonal targets is a function of speech tempo, phonological vowel length, syllabic structure and segmental effects (intrinsic vowel duration, consonant voicing etc.), adjacency to word and intonational boundaries as well as proximity to other tones. The data analyzed in this article reveal that the timing of L and H values have a relatively stable alignment with the onset or offset of the syllable carrying a pitch accent, thus confirming the prediction of the level view (i.e., the existence of well-defined targets as well as "segmental anchors" to which the tones would be aligned).

1.2. Aim

On the basis of experimental data, we investigate whether the information structure affects the choice and realisation of the nuclear pitch accents in Bulgarian and Polish with respect to peak alignment and whether the phenomenon of segmental anchoring can be observed in these languages.

The following three hypotheses regarding the factors affecting the variability vs. stability of nuclear peaks are investigated:

1. Different focus types are associated with specific nuclear pitch accents.
2. The peak of the falling vs. rising pitch accent is consistently anchored to specific points in the segmental structure.
3. The phonetic realisation of phonologically specified accent types is language specific.

2. Material and method

A production experiment was carried out for Bulgarian and Polish. Since we are primarily interested in the contribution of the intonation for signalling focus, the canonical word order was used for the test sentences, i.e., *subject < verb < direct object < indirect object< oblique*[1]. This increases the role of the intonation as an information-structuring factor, allowing us to analyse the realisation of the focus-associated accent patterns in Bulgarian and Polish statements with respect to the f_0 peak alignment, independent from the syntactic structure.

Moreover, we designed the material to make the data easily comparable in both languages (cf. the four test sentences for Bulgarian and Polish below). There are one to four unaccented syllables between the metrically strong syllables with the same maximally sonorant segmental structure (*'ma'*) in order to avoid microprosodic effects.

Speech material for Bulgarian:

1. '*včera* '*mama* '*maza* '*masata*.
   yesterday mama painted the table
   'Yesterday mum painted the table.'

[1] Because of difficulty of constructing the stimuli for Polish (fixed stress on the penult) the word order in test sentence 3 and 4 is not the canonical one.

2. '*včera* '*mama* *po*'*maga* *na* '*Mareto*.
   yesterday mama helped to Mareto
   'Yesterday mum helped Mareto.'
3. '*včera* '*mama* *ni* *po*'*maga* *po* *gra*'*matika.*
   yesterday mama us helped in grammar
   'Yesterday mum helped us in grammar.'
4. '*utre* '*mama* *šte* *ni* *po*'*maga* *po* *mate*'*matika*.
   tomorrow mama will us help in mathematics
   'Tomorrow mum will help us in mathematics.'

Speech material for Polish:

1. '*mama* '*ma* *te*'*maty.*
   mother has topics.
   'Mother has topics.'
2. '*mama* *wy*'*maga* *te*'*matu.*
   mother requires topic.
   'Mother requires a topic.'
3. *a*'*mator* *nas* *na*'*mawiał* *do* *te*'*matu.*
   amateur us persuaded to topic.
   'An amateur persuaded us to accept the topic.'
4. a'*mator* *nam* *niedo*'*magał* *przy* *Kaza*'*matach.*
   amateur us felt unwell in Kazamaty
   'An amateur felt unwell in Kazamaty.'

The subjects for the production experiment were two tertiary-level educated female speakers of Sofia Bulgarian and two of standard Polish. They produced the sentences six times in random order at normal and fast speech rate in a sound-treated studio at the Institute of Phonetics (Saarland University). No explicit instructions regarding accentuation were given to the subjects. In order to elicit broad, narrow non-contrastive and narrow contrastive focus, the test sentences were embedded in dialogue exchanges as replies to *wh*-queries uttered by the instructor and directed towards the first, second or last content word (cf. Table 1).

	broad	narrow initial	narrow medial	narrow final
statements [-contrast]	x	x		x
statements [+contrast]		x	x	x

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

In total there were 288 utterances per speaker for Polish and Bulgarian. 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). All target words in the data were manually labelled on the basis of the synchronised microphone signal and spectrogram using a slightly modified SAMPA transcription.

In addition to the segmental labelling the pitch accents, phrase accents and boundary tones were also labelled[2], using ToBI (Beckman & Ayers, 1994), with the peak alignment of the L(ow) and H(igh) targets explicitly specified. The positions of the f_0 maxima and minima were double-checked by an automatic procedure for which the Praat pitch tracker was used.

The peak delay was calculated a) as the absolute distance in time from the f_0 peak to syllable onset, syllable offset and rhyme onset. and b) as the proportion of the rise/fall duration relative to the syllable or rhyme duration.

3. Results

3.1. Focus-driven peak alignment

The framework adopted in the present study is Pierrehumbert's autosegmental-metrical model of intonational phonology (Pierrehumbert, 1980). The phonological correlate of focus is a pitch accent which is realised on one of the prominent syllables.

[2] For Polish speaker KA only half the data in each condition was analysed.

Before we account for the results, we want to explicitly emphasise the different strategies used by the Bulgarian and the Polish subjects in the realisation of the sentences under different focus conditions. The number of pitch accent types used in the different test conditions is summarised in Table 2 for Bulgarian and in Table 3 for Polish.

Speaker	Focus	Accent type							
		H+!H*		!H*		H*		L+H*	
		norm	fast	norm	fast	norm	fast	norm	fast
BA	broad	0	0	18	0	7	20	0	4
	non-contr.	0	0	1	0	42	40	7	9
	contrastive	0	0	0	0	3	31	69	42
EK	broad	24	24	0	0	0	0	0	0
	non-contr.	4	8	0	0	3	12	32	29
	contrastive	0	0	0	0	2	0	67	69

Table 2: Accent types used by the two Bulgarian subjects in the different focus conditions.

In the Bulgarian data we observe four different accent types, namely L+H* with late peak alignment, H* with early peak alignment, !H* with early peak alignment and H+!H*/L*. The boundary tones in the test sentences are realised as L-L%.

In the case of broad focus in both normal and fast speech speaker EK uses a pitch accent which can be analysed as either H+!H* (a downstepped high target preceded by a high tone) or H+L* (a low target preceded by a high tone). There are also 12 realisations of this pitch accent type when narrow non-contrastive focus is on the last content word. Because of the sentence-final position and the upcoming low boundary tones L-L% it is impossible to distinguish between the two accent types or mark the position of the peak in the signal (cf. Fig.1). For this reason we shall exclude the broad focus data for this speaker.

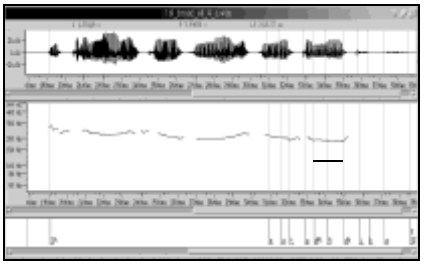

Figure 1: Realisation of H+!H*/L* in final position (Bulgarian speaker EK)

In fast speech the same H* accent type is used by speaker BA in the majority of the narrow non-contrastive and broad foci. This focus-associated accent H* is manifested as a small rise (from the middle of the speaker's voice range) on the onset of the accented syllable, where the H target is a local peak, aligned around the beginning of the syllable rhyme. The tonal movement from the high target to the low boundary target is not phonologically specified. It is realised as a linear interpolation, i.e. a transition between tonal targets. The way the H* is realised is different in the final position (cf. Fig. 2) to that in the non-final position (cf. Fig. 3) and depends on how long the stretch is between the accent and the boundary tone. In the non-final position the fall to the low phrase accent (L-), associated with the metrically strong syllable in the foot following the accented syllable is usually more gradual, while in the final position it is steeper, since L- must be realised on the same syllable.

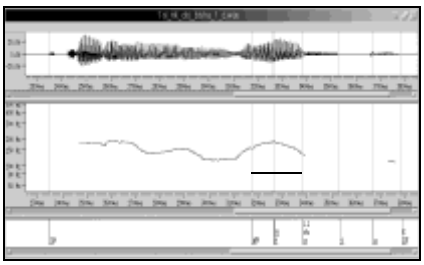

Figure 2: Realisation of H* in final position

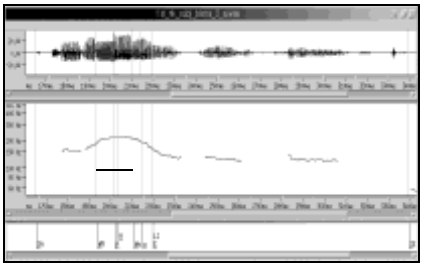

Figure 3: Realisation of H* in non-final position

In the case of focus on the last content word in the utterance there is an ambiguity between the broad focus and narrow non-contrastive focus. This ambiguity is resolved by subject BA in the frequency domain. The pitch range of the narrow non-contrastive focus-associated H* is significantly higher than that used in the broad focus.

The H* accent is also used by speaker BA in 31 cases in fast speech rate in narrow contrastive focus condition. The speaker disambiguates between the narrow contrastive and non-contrastive focus in initial position by using surprisingly significantly higher f_0 peak values for non-contrastive than for contrastive focus. In final position the same tendency is observed but it is not significant. The non-intuitive distribution of peak heights is counteracted by a slightly later peak alignment for the contrastive foci. This tendency did not reach significance level.

Speaker BA mostly realises broad focus in normal speech with an early downstepped !H* peak (18 occurrences) on the last content word in the utterance. The difference between the downstepped accent types (!H*) and the same pitch accents without the down-step (H*) is on the one hand in the height and on the other hand in the alignment of the peak. In the downstepped accent the peak is distinctly lower than that of a preceding H-tone and is aligned in the beginning of or just before the syllable onset (cf. Fig 4).

In the realization of the pitch accent in the contrastive focus this speaker prefers L+H* with phonologically specified late peak alignment. Phonetically, the bitonal L+H* is manifested as a high peak preceded by a gradual rise from a valley in the lowest part of the pitch range. The L is aligned at the very beginning or slightly before the onset of the accented syllable, and the H at the end of the accented or in the first post-accentual vowel (cf. Fig. 5).

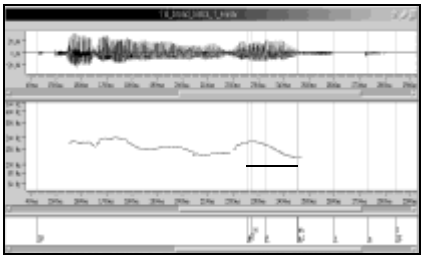

Figure 4: Realisation of !H* in final position (Sp. BA)

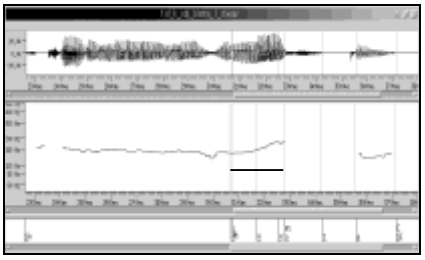

Figure 5: Realisation of L+H* in final position (Sp. BA)

In contrast to speaker BA, speaker EK shows a strong tendency towards realising L+H* in both narrow non-contrastive and contrastive focus condition. This subject disambiguates between the two focus conditions not by peak alignment, but by a higher peak in narrow contrastive than in narrow non-contrastive focus condition. However, the main effect of focus type on peak height is not significant.

With respect to the acoustic properties of H* and L+H*, there are contradictory views in the intonational research about whether these accents are categorically different or that they are just two extremes of a simple accent type. Contrary to claims by Pierrehumbert (1980) and Pierrehumbert and Hirschberg (1990) that only L+H* can be preceded by a low target, Ladd and Schepman (2003) provide statistical evidence that this is also true for H*. A related issue is whether these two accents are associated with different meanings. With regard to our Bulgarian data we can argue that the domain of interpretation of H* and L+H* overlap. Both accent types can signal either new information or a presence of contrast.

In the Polish data we observe four different pitch accent types: !H+L*, H+L*, H*+L, L+H*. The first three are phonetically realised as a fall with an early peak aligned at different positions with respect to the accented syllable. The fourth one (L+H*) represents

a rising movement with a late peak. The two Polish subjects differ in their choice of pitch accent type across focus conditions.

Speaker	Focus	Accent type							
		!H+L*		H+L*		H*+L		L+H*	
		norm	fast	norm	fast	norm	fast	norm	fast
WM	broad	14	14	10	10	0	0	0	0
	non-contr.	6	1	0	0	27	50	15	14
	contrastive	0	0	0	0	47	46	25	26
KA	broad	0	0	0	0	12	12	0	0
	non-contr.	0	0	0	0	12	12	12	12
	contrastive	0	0	0	0	0	0	36	36

Table 3: Accent types used by the two Polish subjects in the different focus conditions.

For example, in the broad focus condition, at both speech rates, only speaker WM uses !H+L* and H+L* accents. These accent types are realised as a fall from a high target in the preceding syllable to a low target situated just after the rhyme onset. The difference between the two pitch accents is that the peak in the downstepped one is perceived lower in comparison to the preceding high target in the utterance (cf. Fig 6 and Fig. 7).

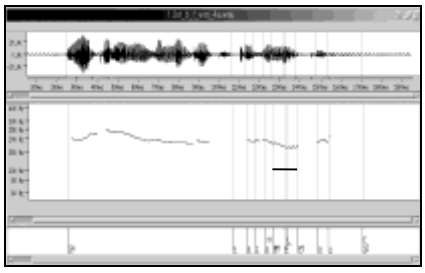

Figure 6: Realisation of !H+L* (Polish Speaker WM)

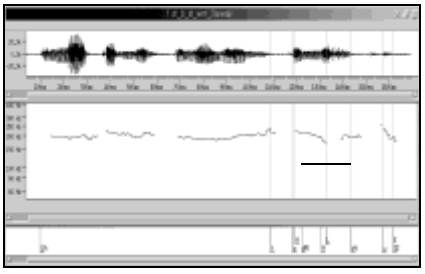

Figure 7: Realisation of H+L* (Polish speaker WM)

In contrast to speaker WM, speaker KA uses H*+L in broad focus condition. This accent type is also used by both speakers in narrow non-contrastive focus condition (there are also seven realisations of !H+L* by speaker WM). In comparison to H+L* the high

target of phonological category of H*+L is aligned later, namely just after the rhyme onset (cf. Fig 8 and 9).

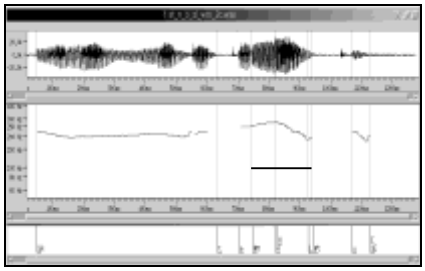

Figure 8: Realisation of H*+L (Polish speaker WM)

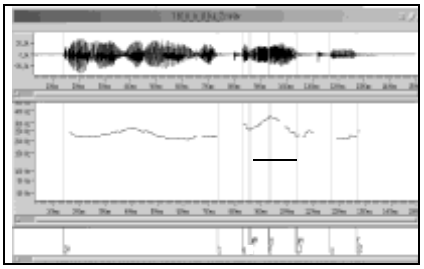

Figure 9: Realisation of H*+L (Polish speaker KA)

When narrow non-contrastive focus is on the final content word in an utterance the speakers have to disambiguate between this condition and a broad focus where the focus exponent is in the same position. While speaker WM achieves this by using two different accent types ((!)H+L* vs. H*+L, for broad and narrow non-contrastive respectively), speaker KA uses the frequency domain. We found peak f_0 values for narrow non-contrastive to be significantly higher than the broad focus f_0 values for this speaker. When narrow non-contrastive focus is in sentence-initial position (subject in focus) both speakers use L+H*. The low target of the L+H* accent is situated just before or at the beginning of the accented syllable. The high target occurs at the end of the accented syllable or at the beginning of the next syllable (cf. Fig 10 and 11).

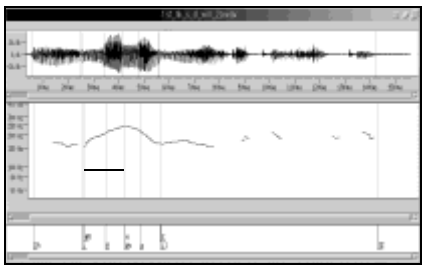

Figure 10: Realisation of L+H* (Polish speaker WM)

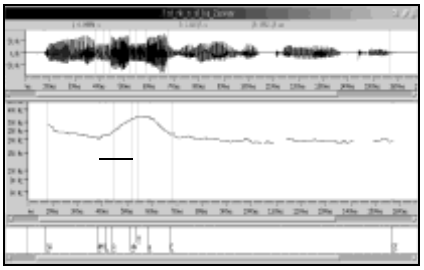

Figure 11: Realisation of L+H* (Polish speaker KA)

In narrow contrastive focus speaker WM uses both L+H* (51 occurrences) as well as H*+L accents (93 occurrences). In the same condition speaker KA uses just L+H* accent type. Because both speakers use L+H* on the sentence-initial word in narrow non-contrastive as well as in contrastive conditions they need to disambiguate between them. Speaker WM achieves this in the time domain by varying f_0 peak alignment, reaching the peak significantly later in the contrastive condition. Speaker KA on the other hand disambiguates these cases in the frequency domain by using significantly higher f_0 values in narrow contrastive focus.

Additional disambiguation is used by speaker WM in the case of focus on sentence-final items in narrow non-contrastive and contrastive focus. This speaker uses an H*+L in both cases and significantly shifts the f_0 peak to later in the syllable in the contrastive condition.

3.2. Phonetically driven peak alignment

According to the hypothesis in the Introduction, the peak of the falling vs. rising pitch accent is consistently anchored to specific points in the segmental structure and is language specific. If so, the peak alignment, measured as an actual proportion of the syllable/rhyme length (relative alignment), should not be affected by increasing speech rate the resulting shorter duration of the accented word. On the other hand, the absolute distance in time from the syllable/rhyme onset/offset should differ significantly with changing speech rate. To analyse the effects of speech rate on peak alignment we carried out multivariate analyses of variance.

As expected the statistical analysis of the data for both languages shows that at a 5% significance level speech rate influences the absolute and not the relative peak alignment measure. However, both languages behave differently under time pressure. As shown in Figure 12 the direction of f_0 peak shift diverges in Bulgarian and Polish with respect to the type of pitch accent. In rising pitch accents with increasing speech rate the peak is placed earlier in Bul-

garian and later in Polish. On the other hand, under the same condition (fast speech) in falling accents Bulgarian speakers reach the high target later and Polish speakers earlier.

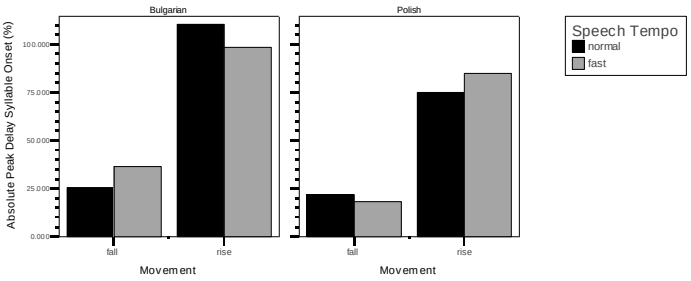

Figure 12: Speech rate and accent type interaction in Bulgarian (left panel) and Polish (right panel).

Because the absolute peak alignment differences were found to be significant in both languages, we can shed light on the nature of the anchoring points of the tonal targets in the segmental structure of these languages. In Bulgarian, the anchoring points are the syllable onset, rhyme onset and syllable offset, and in Polish they are syllable onset and offset.

Figure 12 also shows cross-language differences in the placement of the high target point: The Polish speakers align the peaks earlier than the Bulgarian speakers. Because of the different types of falling pitch accents ((!)H+L* and H*+L for Polish vs. (!)H* for Bulgarian) they cannot be directly compared across languages. This is not so in the case of rising pitch accents where all speakers have the same phonologically specified L+H*, but differ on the phonetic level in terms of the peak alignment. On average the Polish speakers place the peak 24% earlier in the syllable than the Bulgarian speakers.

With respect to the position of the focused item in the utterance we find the following tendency: the later the focused item in the utterance the earlier the peak alignment. A possible explanation is the phenomenon of “tonal repulsion”. The proximity of the intonation

phrase boundary tones leads to temporal readjustments of peak location (Silverman and Pierrehumbert 1990).

4. Conclusions

The goal of this study was to investigate the realisation of broad and narrow contrastive vs. non-contrastive focus in terms of accent type and temporal alignment of high tonal targets for different speech rates and position within the utterance.

The following accent types are used by the Bulgarian speakers: H+!H*/L*, (!)H* and L+H*. Polish speakers use (!)H+L*, H*+L and L+H*. For both languages we found different accent types in the same focus condition and the same accent types in different focus conditions, which refutes our first hypothesis. In both languages speakers employ both peak alignment and peak height to obtain a phonological contrast between the different focus conditions. The fact that the absolute measures for peak alignment differ significantly with changing speech rate clearly lends support to the claim that speakers carefully control the peak alignment in an extremely consistent way. The results show that the segmental anchoring points for Bulgarian are syllable onset, rhyme onset and syllable offset and for Polish syllable onset and syllable offset. With this evidence the prediction of our second hypothesis is confirmed. According to our third hypothesis the phonetic realisation of the phonologically specified pitch accents would differ in both languages. The tonal high target of L+H*, the accent type present in both languages, was found to be aligned significantly earlier by Polish than by Bulgarian speakers, but may also be caused by the greater number of post-nuclear syllables in Bulgarian test sentences.

This study by no means exhausts all the factors affecting the phonetics of tonal alignment in Bulgarian and Polish. Further research is needed to determine what these factors are, which of them are language-specific, and which might be considered “universal”.

References

Arvaniti, A., D. Ladd, and I. Mennen (1998). Stability of tonal alignment: The case of Greek prenuclear accents. *Journal of Phonetics 26*, 3–25.

Beckman, M.E. & Ayers, G. (1994). *Guidelines for ToBI Labeling.* Dept.Linguistics, Ohio State University.

Crystal, David (1969). Prosodic Systems and Intonation in English. Cambridge: CUP.

Frota, S. (2000a). Prosody and Focus in European Portuguese. Phonological Phrasing and Intonation (PhD Dissertation, University of Lisbon). New York: Garland.

Kohler, K. (1987). Categorical pitch perception. In *Proc. XIth Int'l Congress of Phonetic Sciences*, vol. 5, Tallin, pp. 331–333.

Ladd, Robert (1996). Intonational Phonology. Cambridge: CUP.

Ladd, D. R., I. Mennen, and A. Schepman (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *J. Acoust. Soc. Am. 107*(5), 2685–2695.

Ladd, D.R., and Schepman, A. (2003) "Sagging transitions" between high accent peaks in English: experimental evidence. *J. Phon.* 31: 81-112.

Miševa, Anastassia (1991). Intonacionna sistema na bâlgarskija ezik, Bâlgarska akademija na naukite, Sofia.

O'Connor, J. D. & Arnold, G.F. (1973). Intonation of colloquial English. Longman.

Pierrehumbert, J. (1980) The phonology and phonetics of English intonation. Ph.D. Thesis, MIT.

Pierrehumbert, J. & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen, J. Morgan, and M.E. Pollack (eds.), *Intentions in Communication.* MIT Press, Cambridge, MA. 271-311.

Silverman, K. and Pierrehumbert, J. (1990) The timing of prenuclear high accents in English. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech.* J. Kingston and M. Beckman (eds.), CUP, Cambridge, 71-106.