Early vs. late focus: pitch-peak alignment in two dialects of Serbian and Croatian

Rajka Smiljanić

The overall goal of this study was to investigate the extent to which the existence of a lexical tonal contrast affects variation in the acoustic properties that cue the contrast. More specifically, the lexical, pragmatic and prosodic effects on $F_0$ peak alignment in two dialects of Serbian and Croatian with distinct phonological properties were examined: the Belgrade dialect with a lexical contrast between ‘rising’ and ‘falling’ accents and the Zagreb dialect without such lexical pitch-accent contrast. In the dialect with a lexical pitch-accent contrast, the effect of narrow focus on peak alignment is to enhance the contrast. In the dialect without such contrast, the effect of narrow focus is to enhance the prominence of the focused word/syllable by locating the salient pitch peak more strictly on the accented syllable. It was argued that in Zagreb, alignment differences are due to pragmatically defined accent categories. Tonal crowding in the phrase-final positions additionally affects tonal alignment regardless of the presence or absence of the lexical pitch-accent contrast. These results suggest that the variation in the acoustic properties that cue the contrast is restricted by the existence of a phonemic contrast. Furthermore, the phonetic implementation of $F_0$ targets is determined by an interaction of lexical, pragmatic and prosodic factors. Combined, all of these results have implications for the distinction between tonal alignment and association.

1. Introduction

Tune-text association has been one of the central questions in Intonational Phonology (Pierrehumbert 1980, Horne (ed.) 2000). Categorical tonal distinctions defined in the phonology of a language provide tonal targets in a tune that associate with segmental anchors. Work by Bruce (1977) on Swedish and by Pierrehumbert (1980) on English intonation has shown early on
that tonal alignment with the segmental string is precisely controlled. However, alignment of tonal targets varies across and within languages. Alignment differences can arise from factors such as: categorical tone contrasts, pragmatic context (e.g., focus), tonal context (crowding), and position in prosodic phrase. In Swedish, tonal alignment is used to distinguish between two lexically contrastive accents: Accent I and Accent II (HL* vs. H*L, Bruce 1990). In English, tonal alignment makes a similar distinction between contrastive pragmatically-defined pitch-accents (L*+H vs. L+H*). Pragmatic information such as narrow focus can be expressed through changes in pitch contours and changes in alignment. Narrow focus can be expressed through introduction of an additional sentential H tone as in Swedish (Bruce 1977), or by pitch-peak retraction as in Spanish and Greek (Face 2002, Botinis 1998), or through pitch-peak protraction as in European Portuguese and Palermo Italian (Frota 2000, Grice 1995). For instance, in Spanish, a prenuclear pitch-accent has a ‘late’ peak, aligned on the postonic syllable, in a broad focus sentence and an ‘early’ peak, aligned within the accented syllable, when the same word is narrowly focused (Face, 2002). Some of these pragmatically induced alignment changes have been analyzed as consisting of categorical distinctions between accent types: H+L* vs. H*+L for European Portuguese (Frota 2000) and Palermo Italian (Grice 1995), L*+H vs. L+H* for Spanish (Face 2002), etc. In these analyses, the difference in the tonal alignment in narrow focus is placed in the realm of phonology through the inventory of contrastive pitch-accents. However, in other analyses, the same phenomena of timing differences in the expression of pragmatic information are explained through gradient alignment effects rather than categorical distinctions (Nibert 2000, Hualde 2002a, Ladd 1996: 96–98).

Alignment patterns are further finely tuned by proximity to other tones. Prieto et al. (1995) have shown that in Spanish pitch-peak location crucially depends, among other things, on the adjacency to intonational boundaries and the proximity to other tones, which causes earlier alignment. Silverman and Pierrehumbert (1990) show that in English pitch peaks on words in nuclear position overall have earlier alignment than in prenuclear position, presumably due to the upcoming boundary tones in the final position. Furthermore, they showed that pitch-peaks are aligned later in vowels which are elongated due to a slowing down of speech. The opposite is true when vowels are lengthened due to final lengthening effects (e.g. in phrase final positions), i.e., prosodic lengthening induces earlier peak alignment. All of these studies suggest that tonal alignment is sensitive to a combination of the factors mentioned above and that several mechanisms, such as those un-
derlying pragmatic focus and prosodic position, define tonal realization. In the recent years, other factors that additionally affect peak alignment, such as speaking rate (Xu 1998, Ladd et. al. 1999) and phonological vowel length (Ladd et. al. 2000) have been identified in numerous, mostly acoustic, studies of various languages.

In this paper, I explore the interplay of lexical, pragmatic and prosodic factors and how some of these potentially conflicting demands on F0 alignment are resolved, through a comparative study of two dialects of Serbian and Croatian (S/C): Belgrade Serbian and Zagreb Croatian (henceforth Belgrade and Zagreb respectively). Three factors and their interactions that might determine pitch peak alignment in these two dialects are considered: 1) lexical i.e., presence or absence of a lexical pitch-accent; 2) pragmatic, i.e., broad vs. narrow focus in utterance-initial position; and 3) prosodic, i.e., adjacent vs. non-adjacent to an intonation boundary/boundary tones (early vs. late focus). The two dialects investigated here have distinct phonological properties: Belgrade has a lexical contrast between ‘rising’ and ‘falling’ accents with ‘late’ and ‘early’ pitch-peak alignment respectively while Zagreb does not have such a lexical pitch-accent contrast (Lehiste and Ivić 1986, Skarić 1991, Smiljanić and Hualde 2000, Smiljanić 2004).

The hypothesis for the present study is that the peak alignment patterns in these two dialects will vary as a function of the presence vs. absence of a lexical pitch-accent contrast (similar to different coarticulation effects in languages with different phonemic inventories, Manuel 1990, Cohn 1993). Zagreb without the lexical pitch contrast will freely use phonetic space through alignment adjustment to express narrow focus. In Zagreb, then, peak alignment patterns would be pragmatically rather than lexically driven as in other intonation languages. However, the expression of narrow focus will be different from Belgrade with the lexical contrast. Belgrade will exhibit less freedom in utilizing the phonetic space since the location of peaks crucially differentiates categorical distinctions at the lexical level. The maintenance of the lexical pitch-accent contrast will, therefore, limit the amount of alignment change available for expressing pragmatic narrow focus. Furthermore, the prosodic effect (proximity to an intonational phrase boundary/boundary tones) is expected to exert influence on F0 peak alignment in both dialects. It is expected that the upcoming boundary will cause earlier alignment for both accent types (Silverman and Pierrehumbert 1990). This effect might potentially jeopardize the lexical contrast in Belgrade where the ‘falling’ peaks are already aligned ‘early’, i.e., with the word-initial tonic syllable. ‘Falling’ peaks might, therefore, have less room for retraction than the word-initial
‘rising’ accents which are aligned with the posttonic syllable. In this situation, the peaks for the two accentual categories might overlap to some extent (see below for the description of the experimental materials).

Through this investigation several questions are addressed: First, what role do lexical, pragmatic and prosodic factors play in shaping pitch contours in S/C? Second, how does the presence vs. absence of a lexical pitch contrast limit pragmatic and prosodic effects on pitch peak alignment? Third, are there differences between pragmatic and prosodic effects on \( F_0 \) peak alignment and, if so, how are these potentially conflicting effects resolved? Finally, how are phonological tonal contrasts implemented phonetically? The results show that pitch peak alignment in Belgrade and Zagreb is greatly sensitive to a combination of lexical, pragmatic and prosodic factors. As hypothesized, there is a difference between the two dialects in patterns of pitch-peak adjustments that corresponds to the presence vs. absence of a lexical pitch-accent contrast. In Zagreb, all peaks are retracted in narrow focus while in Belgrade, the ‘falling’ peaks are retracted and ‘rising’ peaks are unchanged or slightly protracted. Thus, the contrast between the two lexical accent categories is enlarged in Belgrade through the asymmetric manipulation of the ‘rising’ and ‘falling’ peaks rather than just maintained through the possible uniform retraction of both ‘falling’ and ‘rising’ peaks. Also as hypothesized, the upcoming prosodic boundary/boundary tones cause earlier alignment of peaks in utterance-final position when compared with the utterance-initial position in both dialects. However, it was not found that the lexical contrast in narrow-focused words in Belgrade is ever endangered. Finally, the results show that phonetic alignment is not a direct mapping of phonological categories, supporting a crucial distinction between tonal association and alignment (Ladd 1996).

2. Serbian and Croatian

S/C is a stress-language with a lexical pitch alignment contrast only in words with initial stress (Smiljanić and Hualde 2000). This contrast is manifest in ‘late’ vs. ‘early’ alignment of accent peaks. Traditionally these accent types are designated as ‘rising’ and ‘falling’ respectively (Lehiste and Ivić 1963, 1986, Godjevac 1999, 2000, among others). In the ‘rising’ accents \( F_0 \) rises through most of the accented syllable and the peak is aligned with the posttonic while in the ‘falling’ accents \( F_0 \) rises and falls within the accented syllable with the peak being aligned with the accented vowel. In addition, there
is a lexical contrast in vowel length in S/C (Lehiste and Ivić 1986, among others). According to Lehiste and Ivić (1986) similar F0 patterns hold for both long and short series. This description corresponds to most varieties of standard S/C and is reflected in the speech of Belgrade speakers (one of the dialects under investigation here). Although standard Croatian is said to have a contrast in accent type and vowel length similar to those of Standard Serbian, Magner and Matejka (1971) show that some of the accentual distinctions are not found in Zagreb. Furthermore, Zagreb Kajkavian, a non-standard dialect of the area, has been described as lacking both the length and the pitch accent contrasts (Magner 1966). Although the existence of these contrasts is taught in the schools, it is at least doubtful that some Zagreb speakers with the Kajkavian background make either of the two contrasts, in any style (their local dialect or the more formal standard dialect). Therefore, it is expected that some of the Zagreb speakers will exhibit different prosodic characteristics even in their standard speech which is used in the Experiment in this paper. The validity of this hypothesis was established in a pilot study by Smiljanić and Hualde (2000) who show that some Zagreb speakers lack the pitch-accent contrast when speaking the standard variety. For a more detailed description of the dialectal situation see Magner (1966), Inkelas and Zec (1988) and Smiljanić (2004) among others. It is important for this paper to emphasize that standard Croatian and standard Serbian are mutually comprehensible to a very high degree. Impressionistically they do not differ from each other more than, say, British RP and General American English. In the experiments conducted in this investigation, all words are segmentally identical in both varieties. The only expected differences are prosodic.

3. Experiment

Six subjects were recorded: three Belgrade speakers (two female and one male), two Zagreb speakers (one male and one female), and one Karlovac male speaker (some 30 kilometers south of Zagreb). Both Zagreb speakers and the Karlovac speaker are of the same Kajkavian language background and are expected to exhibit the same behavior with regard to the pitch-accent contrast, i.e., they would transfer the prosodic patterns of their local dialect, in which pitch-accent contrast is absent, into their standard dialect productions, which is the style expected in the formal reading situations. All speakers produced the same disyllabic target words in carrier sentences. The target words have either a ‘falling’ (F) or a ‘rising’ (R) accent on the first
syllable. All target syllables have the vowel /a/ and only long vowels were chosen, since accent distinctions are most clearly seen with long vowels (for the behavior of short vowels, see Smiljanić 2004). Mostly sonorant sounds were chosen to avoid segmental pitch perturbations. Example sentences are given in (1) where target words are bolded and the accented syllable within the target word is capitalized:

(1)(R)  a. **Mara** je jela bananu.  
Mara 3rd person Sg-to be PAST-eat ACC-banana.  
‘Mara ate a banana.’

b. **Nada** je vidjela **Mara**.  
Nada 3rd person Sg-to be PAST-sea ACC-Mara.  
‘Nada saw Mara.’

(F)  c. **MLAda** je jela bananu.  
Bride 3rd person Sg-to be PAST-eat ACC-banana.  
‘The bride ate a banana.’

d. **Mama** je vidjela **MLAdu**.  
Mom 3rd person Sg-to be PAST-sea ACC-bride.  
‘Mom saw the bride.’

Target words were placed in either sentence- initial or final position. The test sentences were read in three blocks: first, with the broad focus reading over the entire test sentence; second, with the narrow focus reading on the sentence-initial target word, and third, with the narrow focus reading on the sentence-final target word. The broad focus reading over the entire test sentence was elicited by the triggering question ‘What happened yesterday?’ where the answer contains all new information. The narrow focus reading for the target word in the initial position was prompted by a question such as ‘Did the bridegroom eat a banana?’ for the answer sentence, ‘The bride ate a banana’ which requires contrastive focus on the target word, in this example ‘bride’. The narrow focus reading for the same target word in the final position was elicited with ‘Did mom see the bridegroom?’ for the answer sentence ‘Mom saw the bride’. The prompting questions and answers were written on index cards. The experimental text was written in the Roman alphabet, with which all speakers were familiar and comfortable. Accent types are not indicated in the orthography. The speakers were instructed to read out loud only the answers to the prompting questions.

In all carrier sentences, the number of syllables in sentences was kept at 8. The number of syllables between targeted accent and the preceding and
following accents was kept at two so that accent clash would be avoided. Arvanití et al. (2000) show for Greek that three syllables are needed between accents for a canonical alignment of accents. It is expected, however, that the two-syllable inter-accent interval will allow for accentual gestures to be executed close to their full form. Furthermore, the syllable interval between accents is kept constant across accent types and pragmatic conditions. If there is any tonal crowding present it is of the same amount for both ‘rising’ and ‘falling’ accents and in both broad and narrow focus conditions.

The sentences within each block were randomized and repeated (5 times in the initial position and 10 times in the final position). 120 sentences per speaker were obtained in this way (30 broad focus initial + 30 narrow focus initial + 60 narrow focus final). The total number of target words obtained was 720. Speakers were recorded under quiet conditions using the Kay Computerized Speech Laboratory (CSL) 4300B at the sampling rate of 16 kHz. The sentences were analyzed using PRAAT software for speech analysis (Boersma 1996).

3.1. Measurements

The sentences were transcribed and segmented manually by a combination of listening and by inspection of the spectrogram, \( F_0 \) track and sound wave. Measurements were made of the accented vowel duration and the \( F_0 \) maximum location with respect to the end of the accented syllable. \( F_0 \) peaks were fairly easy to determine. In cases of high plateau (rather than a clear peak) the beginning of the plateau (elbow) was chosen as the pitch peak. Some of the sentences were discarded due to either bad recording, or the wrong accent type put on the word, or the wrong syllable accented, or other factors, such as extensive coarticulation, that obscured segment boundaries (as determined by the experimenter). This yielded a total of 696 sentences used in the analyses.

The choice to measure the peak alignment data with respect to the \( CV \)-boundary was made because it best captures the phonological ‘targets’ for the ‘falling’ and ‘rising’ accents, as previously described, i.e., the peaks for the ‘rising’ accents target the posttonic syllable and the peaks for the ‘falling’ accents target the accented syllable in Belgrade. The end of the accented vowel/syllable, therefore, seems to be a landmark relevant for peak alignment. The end of the stressed vowel is in most cases the syllable boundary as well, i.e. most stressed syllables are open syllables (with the exception of
the word *marva* ‘cattle’ which has a word-initial closed stressed syllable). Where necessary, peak alignment data relative to CV-boundary are included as well.

4. Results I: expression of focus in the initial position

\(F_0\) peak alignment data from the utterance-initial position are given in scatter plots in Figure 1 separately for each speaker. The left three panels show peak alignment for the Belgrade speakers and the right three panels for the Zagreb speakers. In all scatter plots, the X-axis shows peak alignment with respect to the end of the accented vowel/syllable (0 on X-axis). Peak alignment is plotted against vowel duration (given on Y-axis). It can be observed that patterns of peak alignment differ in two dialects. In Belgrade, the largest distinction seems to be between circles and squares, i.e., ‘rising’ peaks cluster together and are later than the ‘falling’ peaks. The biggest difference in Zagreb, on the other hand, is between empty and filled symbols, i.e., between broad and narrow focus. The ‘falling’ and ‘rising’ peaks within a pragmatic condition (e.g. filled circles vs. filled squares) are greatly overlapped in Zagreb. Peak alignment data were submitted to a two-way ANOVA for each speaker separately. The analyses involved two fixed factors: accent type (‘rising’ vs. ‘falling’) and pragmatics (broad vs. narrow focus). Since numerous ANOVAs were performed in all statistical analyses alpha level was adjusted to .001 to ensure against committing a type I error. For all three Belgrade speakers there was a significant effect of accent on peak alignment: B1: \(F(1,55) = 981.858, p < .0001\); B2: \(F(1,56) = 306.409, p < .0001\); B3: \(F(1,54) = 289.507, p < .0001\). For B2 and B3 there was a significant effect of pragmatics: B1: \(F(1,55) = 4.642, p = .036\) ns; B2: \(F(1,56) = 70.577, p < .0001\); B3: \(F(1,54) = 60.050, p < .0001\). Additionally, for all three speakers there was a significant two-way interaction: B1: \(F(1,55) = 23.292, p < .0001\); B2: \(F(1,56) = 30.534, p < .0001\); B3: \(F(1,54) = 21.592, p < .0001\). Pairwise comparisons reveal that accent is a significant factor for all three speakers for broad focus condition: B1: \(t(13) = 13.892, p < .0001\); B2: \(t(14) = 7.746, p < .0001\); B3: \(t(12) = 7.083, p < .0001\), and for narrow focus condition: B1: \(t(14) = 23.892, p < .0001\); B2: \(t(14) = 19.660, p < .0001\); B3: \(t(14) = 7.083, p < .0001\). In other words, in both pragmatic conditions, ‘falling’ accents are aligned differently than ‘rising’ accents (they are earlier as seen in Figure 1). Furthermore, pairwise comparisons of accent type across two pragmatic conditions show a significant effect of focus for ‘falling’ accents: B1: \(t(13) =
Early vs. late focus: Serbian and Croatian

6.640, p < .0001; B2: t(14) = 9.725, p < .0001; B3: t(12) = 6.965, p < .0001, but not for ‘rising’ accents: B1: t(14) = 2.790, p = .014 ns; B2: t(14) = 2.693, p = .018 ns; B3: t(14) = 3.455, p = .004 ns.

Figure 1: $F_0$ peak alignment in broad (filled symbols) and narrow (empty symbols) focus conditions in utterance-initial position. The panels on the left show the results for the Belgrade speakers and on the right for the Zagreb speakers. ‘Rising’ accents are shown as squares and ‘falling’ accents as circles. X-axis shows peak alignment with respect to the end of the accented vowel/syllable (0). Y-axis shows vowel duration.
This shows that the ‘falling’ accents change their alignment with respect to the end of the accented vowel significantly (empty circles vs. filled circles), i.e., they are aligned earlier with respect to the end of the accented vowel/syllable when compared to ‘falling’ peaks in broad focus. The ‘rising’ accents remain fairly stable in their alignment with respect to the end of the accented vowel (empty squares vs. filled squares).

For none of Zagreb speakers was there a significant effect of accent on $F_0$ alignment: Z1: $F(1,55) = 4.965$, $p = .031$, ns; Z2: $F(1,48) = 1.630$, $p = .208$ ns.; Z3: $F(1,54) = 1.132$, $p = .293$ ns. For all three speakers there was a significant main effect of pragmatics on $F_0$ peak alignment: Z1: $F(1,55) = 99.594$, $p < .0001$; Z2: $F(1, 48) = 46.878$, $p < .0001$; Z3: $F(1,54) = 431.079$, $p < .0001$. There were no significant two-way interactions: Z1: $F(1,55) = 1.658$, $p = .204$ ns; Z2: $F(1,48) = .116$, $p = .735$ ns; Z3: $F(1,54) = .742$, $p = .039$ ns. These results show that, in Zagreb, $F_0$ alignment is not different for ‘rising’ and ‘falling’ accents (circles and squares within a pragmatic condition are largely overlapped in Figure 1, right panels). $F_0$ peaks do, however, change their alignment in narrow focus, i.e., they are aligned significantly earlier (empty symbols are earlier than filled symbols in Figure 1).

4.1. Discussion

The results of the broad focus condition confirm earlier experimental findings that Belgrade speakers differentiate between ‘early’ and ‘late’ peak alignment corresponding to the lexical pitch accent categories. (Lehiste and Ivić 1986, Smiljanić and Hualde 2000). However, it can be seen from the plots in Figure 1 that the ‘falling’ peaks are not necessarily on the tonic syllable as traditionally described. Speakers B2 and B3 align ‘falling’ peaks in broad focus on both tonic and posttonic syllables (filled circles). Despite this ‘non-canonical’ alignment of ‘falling’ peaks they are distinguished from the ‘rising’ peaks which are placed significantly later onto the posttonic syllable. This ‘non-canonical’ alignment is not due to tonal crowding, since $F_0$ peaks are aligned ‘later’ than their traditional descriptions suggest, i.e., closer to the following tonal target. Such alignment patterns could possibly be due to hypo-articulation and a less-precise coordination of segmental and tonal targets characteristic of more informal speaking styles to which broad focus sentences might belong (Lindblom 1990, de Jong 2000). The inherent variability observed in larger distribution of the ‘falling’ accents data can partly be explained by the variable syllable structure of the words involved. One
of the words, *marva* ‘cattle’, has a closed accented syllable and one of the words, *mlada* ‘bride’, has an onset cluster. Both of these factors could affect peak alignment and could account for the variability in the results.  

The results of the broad focus condition further confirm earlier experimental findings that the pitch accent contrast is absent for some Zagreb speakers (Smiljanić and Hualde 2000). For Zagreb speakers, $F_0$ peaks in broad focus can be characterized as ‘late’ peaks with somewhat variable alignment for different Zagreb speakers (filled symbols, right panels in Figure 1). This peak alignment does not correspond clearly with the late/’rising’ peak alignment in Belgrade for all speakers. The variability in peak alignment observed among Zagreb speakers could partly be accounted for by the fact that one speaker (Z2) is from Karlovac. It could be argued that this speaker does not clearly belong to the Zagreb dialect but to a distinct dialectal group. However, the fact that this speaker lacks the pitch accent contrast makes him more similar to Zagreb speakers than to Belgrade speakers. In any case, the variability between/among dialects suggests that the prenuclear ‘rising’ tonal categories are phonetically implemented in different ways. The variability in the alignment of ‘rising’ peaks cannot be accounted for by the varying structure of the test words since all ‘rising’ words have the same CVCV structure. It is more likely that there are dialect differences in the instantiation of this accentual category (similar to Atterer and Ladd, 2004). Finally, differences in peak alignment among the speakers within a dialect (even excluding Z2) suggest that the alignment of tonal targets in this case is not as invariable as that of tonal targets in Greek pre-nuclear accents (Arvaniti et al. 1998).

The results for narrow focus show that $F_0$ peak alignment is changed in both dialects. However, the alignment change is different in the two dialects. The alignment adjustment crucially depends on the ‘availability’ of the phonetic space to realize the change in the alignment. For Belgrade speakers the phonetic space in broad focus is already carved in two parts for the lexical distinction between ‘rising’ and ‘falling’ accents. Such division of the phonetic space limits the available alignment change for Belgrade speakers. In narrow focus the lexical distinction is maintained and, despite the alignment changes that peaks undergo, they do not cross over these spaces (‘rising’ peaks are not realized on the tonic syllable although potentially the retraction of both ‘falling’ and ‘rising’ peaks could be a signal of narrow focus). In Zagreb, there is no such division of phonetic space in broad focus. The availability of phonetic space for ‘early’ alignment is utilized completely for the expression of narrow focus. For all Zagreb speakers all $F_0$ peaks in
narrow focus are aligned with the tonic syllable (where the lexical category of ‘falling’ accents is realized for Belgrade).

An important acoustic correlate of narrow focus, which enhances the salience of the accented syllable in both dialects equally, is vowel lengthening. For all speakers white symbols are above the filled symbols in Figure 1. It is important to look closer at vowel duration data in order to establish whether vowel lengthening is the cause of the change in the peak alignment rather than the independent peak retraction. It is possible that the peaks have the same alignment with respect to the beginning of the accented vowel in both pragmatic conditions. The observed change in the alignment could then entirely be due to the lengthening of vowels in narrow focus. Two-way ANOVAs for each Belgrade speaker separately were performed for the vowel duration data. The analyses involved two fixed factors: accent type (‘rising’ vs. ‘falling’) and pragmatics (broad vs. narrow focus). For all three Belgrade speakers there is a significant main effect of pragmatics on vowel duration: B1: F(1,55) = 75.463, p < .0001; B2: F(1,56) = 85.597, p < .0001; B3: F(1,54) = 14.057, p < .0001. For B2 and B3 there is a significant effect of accent on vowel duration: B1: F(1,55) = 7.629, p = .008 ns; B2: F(1,56) = 12.540, p = .001; B3: F(1,54) = 15.620, p < .0001. No two-way interactions were significant: B1: F(1,55) = 1.345, p < .251 ns; B2: F(1,56) = 6.288, p < .015 ns; B3: F(1,54) = .010, p = .919 ns. Vowels are lengthened in narrow focus on average 35 ms for B1, 37 ms for B2, and 18 ms for B3. Vowel length also differs between ‘rising’ and ‘falling’ accents. ANOVAs with only pragmatics as a fixed factor were performed on the vowel duration data for Zagreb speakers (accent is no longer considered to be a factor). For all three speakers there is a significant main effect of pragmatics on vowel duration: Z1: F(1,55) = 43.259, p < .0001; Z2: F(1,48) = 15.744, p < .0001; Z3: F(1,54) = 256.396, p < .0001. Vowels are lengthened in narrow focus on average 33 ms for Z1, 28 ms for Z2 and 50 ms for Z3.

In order to check whether the difference in the alignment is due to vowel lengthening or to the active peak alignment manipulation, mean peak alignment change (ΔH) between broad and narrow focus was calculated with respect to the beginning of the accented vowel. The results are given in Table 1. The results in Table 1 show that for B2 and B3 ‘falling’ peaks are retracted independently of vowel lengthening (‘falling’ peaks for these speakers were in the ‘non-canonical’ posttonic position in broad focus). This is not the case for B1 whose ‘falling’ peaks were already on the tonic in the broad focus condition. For this speaker vowels are significantly lengthened in narrow focus which indicates that the peak alignment results are not a consequence of
‘misreading’ of the pragmatic context. This suggests that the primary function of peak retraction is enhancing the contrast (already large enough for B1 in broad focus). Additionally, ‘falling’ peaks are treated differently from the ‘rising’ peaks, which are either slightly protracted or not changed (for B3) but are never retracted. This strategy allows the Belgrade speakers to additionally exaggerate the contrast between ‘falling’ and ‘rising’ accents. All three Zagreb speakers retract all the peaks independently of vowel lengthening. These results suggest that pitch manipulation due to narrow focus is somewhat different in the two dialects, i.e., it reflects their different phonological properties. In Belgrade lexical contrast is enlarged through the asymmetrical manipulation of ‘rising’ and ‘falling’ accents (similar to the lexical vowel length contrast enhancement reported by, e.g., de Jong and Zawaydeh 2001, Smiljanić 2004). In Zagreb, on the other hand, the $F_0$ peak alignment change is completely a function of expression of pragmatic narrow focus (confirming Smiljanić and Hualde, 2000). Therefore, peak alignment is not utilized in the same way in expressing pragmatic narrow focus in the two dialects.

Table 1: Mean amount of $F_0$ peak alignment change between broad and narrow focus conditions in the initial position. Negative numbers indicate retraction.

<table>
<thead>
<tr>
<th>ΔH (ms)</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>Z1</th>
<th>Z2</th>
<th>Z3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>−1.36</td>
<td>−80.7</td>
<td>−72.2</td>
<td>−32.6</td>
<td>−40.8</td>
<td>−103.1</td>
</tr>
<tr>
<td>R</td>
<td>13.2</td>
<td>25.1</td>
<td>−4.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Results II: early vs. late focus

Here, the results of the comparison between early and late focus (utterance-initial vs. final position) are given (e.g., Mara in the initial narrow focus condition (3a) vs. Maru in the final narrow focus condition (3b)). The data in Figure 2 are plotted in the same way as in Figure 1. In all plots in Figure 2 the empty symbols are data from the narrow focus initial position condition (the same points were seen in Figure 1). The filled symbols are data from the narrow focus final position. For Zagreb speakers, ‘falling’ and ‘rising’ points are collapsed since it was shown in Results I that accentual distinctions are absent in this dialect. Peak alignment in the final position is additionally modified when compared with the initial position (filled symbols vs. empty symbols). The most striking adjustment is in the alignment of the ‘rising’
accents in Belgrade (filled squares). Recall that for this accent type in both broad and narrow conditions in the initial position peaks are consistently ‘late’ (on the posttonic syllable). In the final position they are clustered fairly tightly around the end of the tonic syllable. For Zagreb speakers Z1 and Z3, F₀ peaks in the final position (filled triangles) are earlier than the peaks in the initial position. For Z2 the peaks in both conditions are largely overlapped.

Figure 2: F₀ peak alignment for the Belgrade speakers (left panels) and the Zagreb speakers (right panels) in narrow focus in initial and final positions. The data are plotted in the same way as in Figure 1.
Two-way ANOVAs with position (initial vs. final) and accent (‘falling’ vs. ‘rising’) as fixed factors show that for all Belgrade speakers there is a main effect of position on peak alignment: B1: F(1,56) = 179.541, p < .0001; B2: F(1,56) = 9.231, p < .001; B3: F(1,56) = 256.289, p < .0001. There is also a significant main effect of accent on peak alignment: B1: F(1,56) = 692.805, p < .0001; B2: F(1,56) = 636.459, p < .0001; B3: F(1,56) = 764.281, p < .0001. Additionally, there is a significant two-way interaction for all three speakers: B1: F(1,56) = 45.922, p < .0001; B2: F(1,56) = 47.137, p < .0001; B3: F(1,56) = 48.684, p < .0001. It was established earlier that ‘falling’ and ‘rising’ accents are significantly different in narrow focus in the initial position for all three speakers. Pairwise comparisons reveal that accent is a significant factor in the final position as well: B1: t(14) = 15.890, p < .0001; B2: t(14) = 12.649 p < .0001; B3: t(14) = 15.305, p < .0001. Furthermore, for all three speakers ‘rising’ accents change their alignment significantly in the final position when compared with the initial position: B1: t(14) = 11.402, p < .0001; B2: t(14) = 11.073, p < .0001; B3: t(14) = 17.038, p < .0001. Finally, the alignment of ‘falling’ accents is significantly different in the final position for B1 and B3 only: B1: t(14) = 6.693, p < .0001; B2: t(14) = 1.895, p = .097 ns; B3: t(14) = 5.319, p < .0001. These results show that in both positions, ‘falling’ and ‘rising’ peaks are distinct with ‘falling’ accents being aligned earlier. Additionally, F₀ peaks are aligned earlier in the final position than in the initial position (filled vs. empty symbols). For two Zagreb speakers ANOVAs with the fixed factor of position (initial vs. final) showed a significant effect of position on peak alignment: Z1: F(1,56) = 107.058, p < .0001; Z2: F(1,55) = 3.773, p = .057 ns; Z3: F(1,51) = 103.669, p < .0001. For Z1 and Z3, F₀ peaks are aligned earlier in final position. For Z2 there is no difference in the alignment between two conditions.

5.1. Discussion

The results show that ‘falling’ and ‘rising’ peaks are distinct for Belgrade speakers in narrow focus final position. The results also show that peaks are additionally modified in the final position in both dialects. The ‘falling’ peaks for Belgrade speakers are just slightly earlier in the vowel (not for B3) but the ‘rising’ peaks are much earlier than in the sentence-initial position. For Zagreb speakers, peaks are earlier in the final position except for Z2.

Vowel lengthening was examined in the final position as well. Two-way ANOVAs with position (initial vs. final) and accent (‘rising’ vs. ‘falling’) as
fixed factors for each Belgrade speaker show that there is a main effect of position on vowel duration: B1: F(1,56) = 62.805, p < .0001; B2: F(1,56) = 16.792, p < .0001; B3: F(1,56) = 30.819, p < .0001. Accented vowels are on average longer in the final position for B1 by 29 ms and for B3 by 28 ms but are shorter for B2 by 19 ms. There is also a main effect of accent for all speakers: B1: F(1,56) = 34.041, p < .0001; B2: F(1,56) = 96.706, p < .0001; B3: F(1,56) = 12.844, p < .001. For all speakers, there is an overall significant durational distinction between ‘rising’ and ‘falling’ accents. Finally, there is a significant two-way interaction for B2: B1: F(1,56) = 1.948, p < .168 ns; B2: F(1,56) = 20.455, p < .0001; B3: F(1,56) = .080, p < .778 ns. Pairwise comparisons for B2 show that in both positions vowels with ‘rising’ accents are different from vowels with ‘falling’ accents: t_{fin}(14) = 4.489, p = .001; t_{fin}(14) = 10.827, p < .0001. It can be seen in Figure 2 that for B2, vowels for ‘rising’ accents are longer than vowels for ‘falling’ accents (squares vs. circles). The effect of position is significant for vowels with ‘falling’ accents: t(14) = 5.934, p < .0001 but not for vowels with ‘rising’ accents: t(14) = .288, p = .778. Vowels for ‘falling’ accents are shortened while there is no major change in the duration of ‘rising’ accents in the final position. ANOVAs with position (initial vs. final) as a fixed factor show that there is a main effect of position on vowel duration for Z3: Z1: F(1,56) = 1.022, p = .316 ns; Z2: F(1, 55) = 8.098, p = .006 ns; Z3: F(1,51) = 120.898, p < .0001. Vowels are lengthened on average 4 ms for Z1, 18 ms for Z2 and 42 ms for Z3. The overall effect of final position on vowel duration (filled symbols vs. empty symbols in Figure 2) seems to be of smaller magnitude than was the effect of focal lengthening in the initial position for all speakers.

The mean change in peak alignment (ΔH) in the final position with respect to the beginning of the accented vowel is given in Table 2. This was done in order to see whether the observed peak alignment change is due to vowel lengthening alone or to an independent retraction:

Table 2: The change in the peak alignment between the initial and final positions with respect to the beginning of the accented vowel.

<table>
<thead>
<tr>
<th>ΔH (ms)</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>Z1</th>
<th>Z2</th>
<th>Z3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>-9.8</td>
<td>+18.2</td>
<td>-12.32</td>
<td>-41.7</td>
<td>-0.23</td>
<td>-28.6</td>
</tr>
<tr>
<td>R</td>
<td>-68.3</td>
<td>-53.1</td>
<td>-79.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Peak alignment change in the final position has the effect of ‘repelling’ the peaks away from the boundary in both dialects (similar to Silverman and
Early vs. late focus: Serbian and Croatian (Pierrehumbert 1990) rather than exaggerating the phonemic accentual contrast. For most speakers peaks are retracted independently of vowel lengthening. For Z2, however, peaks are not retracted in the final position. It can be noted that for Z2 already in the narrow focus initial position F₀ peaks were placed earlier than for the other two speakers (a consequence of earlier peak alignment in broad focus for this speaker, i.e., peaks are on the tonic in broad focus). For Belgrade speakers, the retraction is large for the ‘rising’ peaks and smaller or absent for the ‘falling’ peaks. Table 3 shows the amount of accentual difference between ‘falling’ and ‘rising’ peaks in the two positions for Belgrade speakers:

Table 3: The mean amount of difference between ‘falling’ and ‘rising’ peak alignment in the initial and final narrow focus conditions.

<table>
<thead>
<tr>
<th>‘falling’ vs. ‘rising’ distinction</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-R/Initial</td>
<td>167.6 ms</td>
<td>178.3 ms</td>
<td>160.4 ms</td>
</tr>
<tr>
<td>F-R/Final</td>
<td>125.3 ms</td>
<td>167.8 ms</td>
<td>112.42 ms</td>
</tr>
</tbody>
</table>

The change in the peak alignment induced by the final position decreases the difference in the peak alignment between ‘rising’ and ‘falling’ accents. The proximity to the IP boundary affects peak alignment in the opposite direction from the pragmatic effect of narrow focus which increases the lexical pitch distinction (cf. intial broad focus vs. intial narrow focus). Furthermore, the effect of the final position goes against the claim that phonetic space is carved into two areas and that peak adjustments never cause ‘falling’ and ‘rising’ peaks to cross over these spaces. The ‘rising’ alignment in the final position is the same as the ‘falling’ alignment in broad focus in the initial position by some Belgrade speakers (see Figure 1). This suggests that in a situation when ‘late’ alignment is unavailable due to tonal crowding and/or unavailability of segmental material, ‘rising’ peaks can ‘encroach’ on the ‘falling’ territory. However, for all Belgrade speakers, the difference in the alignment between ‘falling’ and ‘rising’ in the final position is still significant (as shown in Results II). Asymmetric vowel shorthening/lengthening of ‘rising’ and ‘falling’ vowels could function as a repair mechanism allowing for the distinction between the two accent types to be maintained (see B2 results). It remains to be investigated further whether the changes in the alignment of peaks in the final position, such that the ‘rising’ peaks are ‘early’ compared to their alignment in the initial position, affect listeners’ perception of the these lexical categories.
6. General discussion

The results show that lexical, pragmatic and prosodic factors determine pitch peak alignment in both dialects of S/C. In postulating phonological tonal categories, therefore, all of the discussed factors have to be taken into account. The following analyses are proposed for the two dialects: in Belgrade there is a lexical contrast between L*+H and L+H* for the ‘rising’ and ‘falling’ accents respectively (for additional arguments for the proposed analysis, see Smiljanić 2003, 2004; for an alternative account see Godjevac 2000, 2001 and Godjevac and Arvaniti, 2003). The timing difference between the late-aligned ‘rising’ peaks and the early-aligned ‘falling’ peaks is indicated by the starred notation. In Zagreb, the phonological tonal category is L*+H. The actual alignment data show inter- and intra-speaker variation in preferred peak position. The prenuclear ‘rising’ category, L*+H, is not implemented in the same way in the two dialects. The variation observed shows that these categories can be implemented in different ways on an alignment continuum (similar argument is given in Atterer and Ladd, 2004). The proposed pitch-accent categories, thus, present idealized associations with segmental anchors rather than the actual alignment facts.

In narrow focus, the pattern of F0 peak alignment change in Zagreb could be explained by postulating a pragmatically-defined pitch-accent category L+H* (similar to the analyses for Spanish and Portuguese, Face 2002, Frota 2000, and possibly for Greek, based on data described in Botinis 1998). The Belgrade narrow focus data, on the other hand, do not lend themselves to the same analysis. Despite peak alignment modifications, we wouldn’t want to postulate two new pitch-accent categories (‘earlier’ early peak vs. ‘later’ late peak). It is not clear what the anchors for these tonal targets would be (they would have to refer to the pitch-accents in broad focus and specify the alignment to be ‘earlier’ or ‘later’ in narrow focus although still on the tonic syllable for ‘falling’ and on the posttonic syllable for ‘rising’). It was shown above that the peak alignment is modified in such a way that the pitch-accent contrast is enlarged. Therefore, it is proposed that F0 peak alignment changes in Belgrade are due to phonetic level adjustments in the expression of narrow focus without changing the accent category. This analysis would capture the distinction in the mechanisms underlying narrow focus effects in the two dialects, whereby in Belgrade the lexical contrast is enlarged while in Zagreb the focus mechanism is entirely driven by pragmatic considerations.

We need to, briefly, consider the alignment patterns in the final position as well. For Belgrade speakers, ‘rising’ peaks in the final position are aligned
Early vs. late focus: Serbian and Croatian

513

earlier (on the tonic) when compared with the initial narrow focus condition. By looking at the alignment facts in this condition alone one could argue that in fact H is a starred tone (aligning with the end of the accented vowel). However, through the comparison of all three conditions investigated here we can conclude that ‘rising’ peak alignment in the final position is not a new pitch-accent category (somewhat intermediate to ‘falling’ and ‘rising’ alignment in the initial position). Such alignment is a result of the interaction of pitch accents and boundary tones. Furthermore, the same analysis accounts for the Zagreb data and reflects the fact that the effect of the prosodic position is the same for both dialects. All of these analyses support the difference between tonal alignment and association as given in Ladd (1996). These results show that tonal alignment is controlled by various mechanisms and that in order to arrive at the right phonological analyses of pitch patterns we have to take into account lexical, pragmatic, and prosodic components which can become obvious only through experimental investigation.

Finally, it is important to note that the results in the final position are taken from the narrow focused words, which may be characterized as hyper-articulated or ‘listener oriented’ (Lindblom 1990, de Jong 2000). A different pattern is observed for target words in the final position with broad focus reading (not discussed here). There, lexical contrast seems to be jeopardized due to deaccenting, i.e., there are no discernable peaks in this condition. It seems that the overall contrast, in both broad and narrow focus, is diminished in the final position. This is in agreement with the finality effects discussed by Hock (1999) whereby the pitch contrasts are diminished or obliterated in prepausal contexts. Furthermore, the displacement of H to the left in the final narrow focus could lend support to the claim made by Becker (1979, as cited in Hock 1999) that the leftward accentual shift, which occurred in S/C in the fifteenth century and reorganized the accentual paradigm, originated in such prepausal contexts.

7. Conclusions

This paper has examined lexical, pragmatic and prosodic effects on F0 peak alignment in two dialects of S/C. It finds that phonetic implementation of F0 targets is determined by all of these factors. The effects on peak alignment vary for these two closely related language varieties. We have compared segmentally identical texts produced by S/C speakers with and without a lexical pitch-accent contrast. In the dialect with a lexical pitch-accent con-
contrast, the effect of narrow focus on peak alignment is to enlarge the contrast. In the dialect without such contrast, the effect of narrow focus is to enhance the prominence of the focused word/syllable by locating the salient pitch peak more strictly on the accented syllable similar to Spanish (Face 2002), Italian (Grice 1995) and Portuguese (Frota 2000). It was argued that in Zagreb, alignment differences are due to pragmatically defined accent categories. The effect of the final position in the prosodic phrase (proximity to a boundary and boundary tones) is similar in both dialects. There seems to be a constraint against $F_0$ peak alignment on the posttonic syllable which is the last syllable in the utterance and a host for the realization of boundary tones. Therefore, in both dialects $F_0$ peaks are further away from the boundary in final position. This causes a conflict in the Belgrade dialect between the demands to maintain the lexical pitch-accent contrast in a hyperarticulated narrow focus context and to accommodate the upcoming boundary tones with little segmental material available. However, the contrast is maintained albeit of smaller magnitude.

This investigation by no means exhausts all the factors that determine phonetic implementation of tonal targets in S/C. A number of other effects need to be calculated for deriving the right model of phonetic tonal implementation. Peak alignment might change as a function of vowel duration differences due to changes in speech rate (Xu 1998, Ladd et al. 1999), of phonological vowel length (Xu 1998, Ladd et al. 2000), of segmental syllable composition (Arvaniti et al. 1998, Ladd et al. 2000), etc. Furthermore, it is possible that it is not just the $F_0$ peak that changes its alignment. The preceding L target could change as well (Face 2002). In that case the investigated factors would affect the entire rising gesture rather than just the $F_0$ peaks. Finally, articulatory constraints, such as the maximum speed of pitch change, and perceptual constraints may impose additional constraints that determine tonal alignment (Xu 2002, D’Imperio 2000). Research is needed to determine what these factors are and how they interact, potentially shedding light on the interplay between universal and language-specific constraints.

Notes

* I am grateful to Jennifer Cole, Jose Hualde, Zsuzsanna Fagyal, Carlos Gussenhoven and Ken de Jong for help at various stages of this project. I am also grateful to two anonymous reviewers and Louis Goldstein
for helpful suggestions and comments on an earlier draft. All errors are mine.

1. The label ‘Zagreb’ dialect is used for the speakers of standard Croatian with Kajkavian background as opposed to the speakers of standard Serbian. The speakers included in this group might differ among themselves, in addition to being different from the Belgrade speakers.

2. The data in the initial position are a subset of a larger set collected in another experiment (Smiljanić 2004). The subset used here is limited to the words with the long vowel /a/ to make the comparison with the same words in the final position more felicitous. This is the reason for the asymmetry in the number of repetitions between initial and final positions.

3. I want to thank an anonymous reviewer for bringing up this point.

4. For statistical analyses of peak alignment data only first five repetitions of each word obtained for the final position were included. This was done in order to balance the number of data points per cell for initial and final position.

5. It would be possible to argue that neither tone is starred as was done by Arvaniti et al. (2000) for Greek. However, this analysis would not capture the regularities in tonal alignment observed in these data.

References

Arvaniti, A., D.R. Ladd, and I. Mennen


Atterer, M. and R. Ladd, R.

Becker, L.

Boersma, P.
Botinis, A.  

Bruce, G.  

Cohn, A.  

D’Imperio, M.  

De Jong, K.  
2000 Attention modulation and the formal properties of stress systems. *CLS* 36.

Face, T.  

Frota, S.  

Godjevac, S.  
2001 *Serbo-Croatian ToBI (SC_ToBI)*. http://www.ling.ohio-state.edu/~godjevac/SCToBI5.pdf

Godjevac, S. and A. Arvaniti  
2003 Phonological structure and phonetic realization of Serbian pitch accents. Presented at the *Experimental Approaches to Tone and Intonation* Workshop of the European Science Foundation Network „Tone and Intonation in Europe.“ Oxford.
Grice, M.

Hock, H.H.

Horne, M. (ed.)

Hualde, J. I.

Inkelas, S., and D. Zec

Ladd, R.

Ladd, R., D. Faulkner, H. Faulkner and A. Schepman

Ladd, R., I. Mennen and A. Schepman

Lehiste, I. and P. Ivić

Lindbloom, B. E. F.

Magner, T.

Magner, T., and L. Matejka
Manuel, S.

Nibert, H.

Pierrehumbert, J.

Prieto, P., J. v. Santen and J. Hirschberg

Silverman, K. and J. Pierrehumbert

Skarić, I.

Smiljanić, R.

Smiljanić, R. and J. I. Hualde

Xu, Y.