Studying Neoštokavian Serbocroation prosody
Keijsper, C.E.

Published in:
Studies in Slavic and General Linguistics

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
STUDYING NEOŠTOKAVIAN SERBOCROATIAN PROSODY*

C.E. KEIJSPER

CONTENTS

Introduction

Chapter One: Preliminaries
  1.1 Mode
  1.2 Three types of perceptual contrast
  1.3 The NeoŠtokavian accent shift
  1.4 Prominence, tone, intonation
  1.5 Notational conventions

Chapter Two: Prominence
  2.1 Cues for accent and for type of accent
  2.2 Word accent versus sentence accent
  2.3 Words

Chapter Three: Tone
  3.0 Introduction
  3.1 Linear and logarithmic scales
  3.2 Consonants
  3.3 Vowels versus configurations
  3.4 Long versus short
  3.5 The notion "peak" (vrhunac)
  3.6 Minimal realizations
  3.7 Northeast versus (South)west?

Chapter Four: Intonation
  4.0 Introduction
  4.1 Preliminary remarks
  4.2 The reverse pattern and related contours
  4.3 Type I versus type III
  4.4 Type III versus type IV
  4.5 Type II (and V) versus type IV
  4.6 Final falls
  4.7 Minor continuations?
  4.8 No sentence-final rise
  4.9 Conclusion

References
INTRODUCTION

Much research has already been devoted to Serbocroatlan prosody, but a further army of accentologists, dialectologists, intonologists, phoneticians and semanticists could easily spend a lifetime unravelling its problems. A complicated historical development has given rise to a great diversity of dialects where, in addition to normal sentence intonation and accentuation, up to five word tones can be found in accented syllables (three if one subtracts length oppositions) and where vowel length can be distinctive in both accented and unaccented syllables; further, the meaning of the notion "word" is unclear as a consequence of scope problems created by the existence of both enclitic words which cannot be accented and which must appear in a fixed order, and proclitic words which can be accented, the next word then becoming enclitic.

The so-called "standard language" has: 1. the "word" problem mentioned in the preceding paragraph; 2. distinctive vowel length in accented and postaccentual syllables; 3. a "rising" vs. "falling" tone opposition in accented syllables, both with long and with short vowels; and 4. sentence intonation and accentuation. Several other systems are acceptable as a "literary" language. For example, people from Belgrade tend not to have a tone opposition in short vowels in disyllabic words before a short postaccentual vowel (e.g. Belić 1971 (originally 1948): 93; for phonetic data see Lehiste 1961); natives of Zagreb speak without tones altogether (e.g. Lehiste and Ivić 1986: 146); etcetera. Thus, the notion "standard language" refers to a maximal, ideal system. This language is based on the dialect group called Neostokavian; in the dialects of this group the so-called "Neostokavian accent retraction" has taken place (see 1.3 below). For a survey of Serbocroatlan dialects, including Neostokavian ones, see Ivić (1958).

Neostokavian prosody has been studied by linguists and phoneticians. Acoustic and perceptual aspects have been investigated most extensively by a couple consisting of one linguist, Pavle Ivić, and one phonetician, Ilse Lehiste. Their collaboration has resulted in a series of publications superseding everything else written on these aspects of the problem. Recently, they have summarized and expanded their work in Lehiste and Ivić (1986); this book also contains an up-to-date survey of experimental work done by others and an extensive bibliography.

In the preface (1986: xiii), Lehiste and Ivić express the hope "that the publication of this book will advance the dialogue and that the result will be a gradual convergence of views toward recognition of the true essence of Serbocroatlan accentuation [tone]". But anybody reading the book, and earlier publications, cannot but recognize that the available data do not allow conclusions different from those already drawn by Lehiste and Ivić. Since the present article contains no new data, it has nothing to add to their work. Yet, it is an attempt to "advance the dialogue", although not between different views on Serbocroatlan accentuation. Rather, it tries to indicate some issues for further, perceptually oriented research, by commenting on Lehiste and Ivić's work against the background of a somewhat different tradition in prosodic research, viz. the approach to sentence intonation developed at the Institute for Perception Research in Eindhoven, the Netherlands (e.g. 't Hart and Cohen 1973; 't Hart and Collier 1975; Collier and 't Hart 1981), in the present author's interpretation. It seems to me that a dialogue between the two traditions could advance our understanding of prosodic phenomena.

During the years of their collaboration, Ivić and Lehiste have gradually shifted the main focus of their attention from the acoustic side of the problem to its perceptual aspects; in this connection the replication of some older listening tests and the spectacular results of new experiments reported in Lehiste and Ivić (1986: 92-128) deserve wide attention. Also, the last twenty years have seen revolutionary improvements in technical facilities available for speech research, culminating at present in analysis-by-resynthesis systems (Lehiste and Ivić 1986: 105). Remarkably, the new developments have not invalidated the older, pioneering work. Therefore, my discussion also includes the older acoustic research, although such retrospection unavoidably gives rise to questions which could not have been asked, let alone answered, at the time when the research was conducted. My remarks must certainly not be seen as an attempt to call the results into question. Rather, I hope that they facilitate the access to phonetic data which are important but whose assessment is, I think, not always easy for the non-specialist. My notes are, as
far as I can tell, fully compatible with Ivrić and Lehiste's intentions, but they are, of course, my own responsibility. The formulation is intended to serve intonologists/accentologists without a knowledge of Serbocroatian, and Serbocroatists not accustomed to "phonetic" discussions about prosody, so that an individual reader may wish to skip one paragraph or other from time to time.

The article is organized as follows. Chapter One introduces some terminology which is indispensable, in my view, for classifying perceptual and linguistic contrasts involving prominence (accent), tone and intonation in a pitch accent system of the type found in Serbocroatian. Chapter Two discusses some aspects of prominence perception and accent placement, the latter mainly in order to connect up the "word" problem mentioned above with Keijsper 1985 and 1986; extensive discussion of the issue is not undertaken here. Chapter Three is about the forms of the "rising" and "falling" tones, as described by Ivrić and Lehiste (henceforth I., L, I&L, LSI, according to who is the author or first author of the publication cited; where no particular publication is referred to, the choice between I&L and LSI is arbitrary). In Chapter Three, only the forms as they appear in "neutral" sentence positions are discussed, i.e. when the distorting effect of sentence intonation is minimal. Finally, Chapter Four deals with sentence intonation and its relation to tone, again on the basis of I&L's work. The division of the material over three different chapters is somewhat arbitrary, as the matters discussed interact in various ways.

The discussion does not embrace all data in I&L's work. The selection is personal in the sense that it concentrates on issues whose further investigation would be desirable in the present author's opinion.

CHAPTER ONE: PRELIMINARIES

1.1 MODE

Consider the following visual pattern:

We can perceive this pattern in two ways: either as a sequence of upward obtrusions or as a sequence of downward obtrusions:

The "upward" way of looking is more obvious, which suggests that we tend to see things in accordance with the idea that figures stand on a ground rather than hang down from it.

Pitch obtrusions used in speech for lending prominence can likewise be divided into basically upward and basically downward types. This fact has been stressed by Bolinger in publications from 1958 onwards. Its importance can hardly be overestimated: disagreements about the relationship between pitch movement and prominence perception largely amounted, in 1958, and still amount, in 1987, to disagreements between those who recognize only upward obtrusions and those who also take into account downward obtrusions.

The two basic types of pitch obtrusion were called mode-1 and mode-2 by Van Katwijk (e.g. 1974: 151):

"There are two modes of intonation: the more common mode 1 is characterised by the fact that the lower pitch level is the basis from which excursions are made to the upper level; the less common mode 2 has the upper pitch level as the basis from which excursions are made to the lower level".

Thus, mode-1 and mode-2 excursions basically have the following form:
In these pictures, and in the further illustrations following below, the vertical lines indicate approximate syllable boundaries. A star is placed on the syllable perceived to be prominent, irrespective of the cue(s) which lead to this perception. Until Chapter Three, pitch movements are drawn schematically.

Not all pitch movements occurring in speech consist of a rise&fall or a fall&rise. In order to accommodate certain perceptual and semantic similarities between different forms in Russian and Dutch intonation, Keijsper (1980: 211-216; 1983: 143-146, 150-152 (note 5)) proposes to apply the notion of mode in such a way that we can abstract away from the level from which an obtrusion actually starts, in the following way. Starting from the basic mode-1 form given above (rise&fall), we replace the rise by a high fragment in the same place:

This form is perceptually similar to the basic form containing a rise, although the only movement in the prominent syllable is a fall; the similarity is created by the fact that the fall does not start already at the borderline of the prominent syllable (as it would in mode-2). In the same way, the falling part of the basic mode-2 form (fall&rise) may be replaced by a low fragment in the same place:

This form groups perceptually with the basic mode-2 form containing a fall, because the rise, which is now the only movement in the prominent syllable, does not start already at the borderline of the prominent syllable (as it would in mode-1).

Then, we may change the second parts of the basic rise&fall and fall&rise. In e.g. Russian and Dutch, the resulting contours are then perceptually as well as semantically different from the basic forms. Thus:

These manipulations produce the correct perceptual and semantic groupings. For example (Keijsper 1980), the Russian contour traditionally called IK-4 (type D in Odé 1986) is schematically:

The initial low fragment or fall to the prominent syllable can, perceptually as well as semantically, be equated with the fall in the contour traditionally called IK-1 (type A in Odé 1986):

The IK-4 contour shares (perceptually and semantically) the high part following the obtrusion with the contour traditionally called IK-6 (type E in Odé 1986):

But the IK-6 contour lacks the low/falling fragment of IK-4, as well as its meaning, the rise being early.

Thus, as schematically indicated in the pictures, in prominent
syllables mode-1 rises occur earlier than mode-2 rises, and mode-1 falls occur later than mode-2 falls.

As we will see below in the discussion of Serbocroatian tones, there is yet another way of keeping track of the mode distinction in different forms. In the basic rise&fall and fall&rise, this other way can be understood by noticing the point where the obtrusion ends:

After the obtrusion, pitch need not remain in the same place, of course. Thus, the following is also a mode-2 movement:

Traditions in intonation research differ as to the part of a contour to which a name is given, which results in different names for the same contour and the same name for different contours. For example, in treatments of English intonation, a movement like the following is often called a "fall" (or "high fall"):

This name does not take into account the initial rise or high fragment. The Dutch (IPO) name "rise&fall" for the corresponding movement in Dutch does take into account the first part. In IPO-terminology, the following movement is a fall:

(British: "low fall"). The name "rise&fall" when used within the British tradition, in contrast, refers to the following movement:

Thus, the British name "fall" refers to a mode-1 accent, the IPO name "fall" refers to a mode-2 accent, the British name "rise&fall" refers to a mode-2 accent, the IPO name "rise&fall" refers to a mode-1 accent. As to Serbocroatian tones, in so far as the names "falling" and "rising" are appropriate at all, they must be understood in the British rather than in the IPO sense.

1.2 THREE TYPES OF PERCEPTUAL CONTRAST

In dealing with pitch phenomena it is essential to separate the physical properties of the speech signal from the perception of the signal; the latter, in its turn, must be kept apart from categories introduced by linguistic analysis.

Given a string of syllables and a pitch obtrusion aligned with it, we may simultaneously perceive 1. prominence, 2. a certain type of pitch movement. Three variables, viz. (physical) alignment of a pitch obtrusion, prominence perception, and perception of the type of pitch movement, define the following types of perceptual contrast (the particular form of the pitch movement used here is only an example):
Here, the alignment of the pitch movement changes. The pitch movement is perceived to be the same in both cases, because the perception of which syllable is prominent changes together with the alignment of the pitch movement. For example, in

\[ \text{you did it} \quad \text{and} \quad \text{you did it} \]

we have the same type of accent (a "(high) fall") on you and did, respectively.

Here too, the alignment of the pitch movement changes. But now the pitch movement is perceived to be different in the two cases, because the perception of which syllable is prominent does not change. For example, in

\[ \text{you did it} \quad \text{and} \quad \text{you did it} \]

we have two different types of accent on you (a "(high) fall" and a (late) "rise\&fall", respectively).

The fact that in this particular example, the different types of accent can be made audible by aligning the same pitch obtrusion differently is a phonetic fact. Linguistically, it is irrelevant whether the difference between two types of accent consists (physically) in a different alignment of a certain pitch obtrusion or in a different pitch obtrusion. Thus, the following pair also has different types of accent on you (a "(high) fall" and a "(high) rise", respectively):

\[ \text{you did it} \quad \text{and} \quad \text{you did it} \]

(The statement that the difference between the two cases has no linguistic status would not be endorsed by autosegmental intonologists.)

Here, the alignment of the pitch movement remains (almost) the same. (Almost: in clear realizations, the obtrusion in the lefthand case will be somewhat earlier than that in the righthand case.) Nevertheless, a different type of pitch movement is perceived in the two members of the pair, because the perception of which syllable is prominent, changes. For example, in

\[ \text{you did it} \quad \text{and} \quad \text{you did it} \]

we have a (late) "rise\&fall" on you and a "(high) fall" on did, respectively.

Since a given physical signal may be ambiguous (especially outside its context) as to which syllable is intended to be prominent, type 3 contrasts are not exceptional (for English see Bolinger 1958: 130-136; 1961). Normally we are not aware of them, because we directly compute the probable intention of the speaker.

Although the members of a type 3 pair are directly opposed to each other perceptually, the example does not represent a direct linguistic contrast. Linguistically it consists of 1. an accent on you versus an accent on did, and 2. a (late) "rise\&fall" versus a "(high) fall". Thus, linguistically the example is a combination of types 1 and 2. Perceptually, one member of a type 3 pair becomes the other by a mode shift.

1.3 THE NEOŠTOKAVIAN ACCENT SHIFT
The Neoštokavian "rising" tone has its historical origin in a process by which an accent on a non-initial syllable of a word was transferred to the preceding syllable. It is unknown when exactly the transfer took place; one hypothesis is that it started in the fifteenth century (e.g. L&I 1986: 1). In the terminology introduced above, the transfer is a shift from a mode-1 perception of a given physical signal to a mode-2 perception:

Note that the pictures illustrating the mode phenomenon here are not intended to be exact representations of the forms of the tones; the latter will be discussed in Chapter Three below.

The mode terminology basically follows Isačenko (1939: 178), who said that

"sich phonetisch an dem Tonprofil ändert. Mehrsilber mit Ultimabetonung [...] zur Zeit der besprochenen "Akzentverschiebung" nichts geändert hat. Das Tonprofil blieb auch nach vollzogener Verlegung der Iktusstelle objektiv das gleiche. Was sich änderte, war die Richtung der Aufmerksamkeit der Sprecher".

I&L have shown that the cues for prominence in the newly accented syllable are mainly duration and vowel quality, whereas "the fundamental frequency and intensity patterns that may be assumed to have been present before the shift still appear to be associated with the originally stressed syllables" (L&I 1963: 134);

see also Chapter Two below.

If before the shift the \( n \)th syllable of a word was accented, after the shift the \( n-1 \)th syllable was, in syllable \( n-1 \) a mode-2 accent is then opposed to a mode-1 accent in other words. This opposition can be eliminated by shifting the mode-1 accents in syllable \( n-1 \) into mode-2 accents in syllable \( n-2 \), and so on. Serbo-Croatian dialects differ from each other in the extent to which they have participated in this chain reaction. Across all dialects the following hierarchy obtains (I. 1958: 105; 1959: 34):

1. The accent is retracted from the final syllable of a word more easily than from a non-final syllable, i.e. if in some dialect the accent has been retracted from non-final syllables, it has also been retracted from final syllables (but not vice versa).
2. The accent is retracted from an open final syllable more easily than from a closed final syllable, i.e. if in some dialect the accent has been retracted from closed final syllables, it has also been retracted from open final syllables (but not vice versa).
3. The accent is retracted from a short vowel more easily than from a long vowel, i.e. if in some dialect the accent has been retracted from long vowels, it has also been retracted from short vowels (but not vice versa).
4. The accent is retracted to a long vowel more easily than to a short vowel, i.e. if in some dialect the accent has been retracted to short vowels, it has also been retracted to long vowels (but not vice versa).

In Neoštokavian dialects, retracted accents are "rising" accents (mode-2), non-retracted accents are "falling" accents (mode-1). Point 3 and 4 of the hierarchy amount, of course, to a consistent tendency to treat length as a cue for accent, and shortness of the vowel in the syllable which would carry the accent in mode-1 as a cue for non-accent, i.e. as an inducement to shift to mode-2. In this connection it is essential that in the whole area, tone oppositions presuppose length oppositions, i.e. if some dialect has tone oppositions, it also has length oppositions (but not vice versa) (e.g. I. 1959: 26-27).

In the so-called standard language the retraction process has affected all syllables of a word. In effect, textbooks for Serbo-Croatian give the following rules:

1. The last syllable of a word cannot be accented (the accent has been retracted to the preceding syllable).
2. A monosyllabic word can only have a falling accent (there is no preceding syllable to which the accent could be retracted).
3. A polysyllabic word can have a falling accent only on the first syllable (idem).
4. A polysyllabic word can have a rising accent on any syllable except the last one (the accent comes from the next syllable).

With respect to rule 3, this textbook formulation is, as a synchronic statement, prescriptive rather than descriptive. As
Vermeer (1984-1985: 149) puts it, it must be doubted whether "outside a narrow circle of linguists, actors, radio announcers, bigots, and similar unreliable informants" speakers can be found who authentically speak the normative system in this respect: since the Neostokavian accent shift, new non-initial falling accents have been introduced, also on the last syllable (cf. rule 1), e.g. in loanwords. But (Lel 1986: 292) "the question whether such examples can be considered acceptable in the standard language has provoked long and bitter discussions".

A further complication concerns the fact that the accent retraction can cross a word boundary. For example, a form like ū kuču ('into the house', short rising accent on ū) is not regarded as an exception to the rule (2.) that monosyllabic words cannot have a rising accent: ū must not be seen as the last and only syllable of the "word" ū, but as the first syllable of the "word" u kuču. It is a special class of clitic words which are not words in the sense of the accent rules; they have no inherent tone and form one accent unit with another word. But, on the other hand, not regarding ū kuču as an exception to rule 2 creates a new exception, as u kūču (same translation; short falling accent on ku-) also occurs, which is an exception to rule 3 if u kuču is regarded as a single "word". It follows that one can either bring cases like u kūču under the same heading as other "exceptions" to rule 3, or regard u kuču as one "word" in some cases and as two "words" in other cases. Other "exceptions" to rule 3 will not be discussed here; for pairs like ū kuču vs. u kūču see 2.3 below.

1.4 PROMINENCE, TONE, INTONATION

All configurations mentioned in the foregoing contain one syllable perceived to be prominent; also, in all of them a certain type of pitch movement is perceived. Prominence does not equal any particular type of pitch movement, nor pitch obtrusion per se. It is a perceptual category, not a part of the physical signal. The pitch obtrusions present in the physical signal may, alone or in combination with other properties of the signal, induce us to perceive an accent (prominence) in a certain place; they are, then, cues for accent (but not accent).

Intonologists nowadays seem to agree that accent in the sense of perceived prominence must be kept apart from stress (although not everybody uses the same terms). Stress is what is marked in a dictionary; the stress mark indicates which syllable will (normally) receive the accent if the given word is accented; in Bolinger's formulation, stress is potential for accent. The differentiation does not seem to be common in the literature on Serbocroatian tones; see 2.2 below for some problems in this connection.

Besides being a perceptual category, accent is also a linguistic category. In Keijser 1985 I have proposed that it is the formal side of a linguistic sign, the semantic side of which is "not not". This meaning is shared by all types of accent. Types of accent (i.e. perceived types) are also linguistically relevant; they may add a meaning to the meaning of accent. For example, in several languages the following configuration signals, in addition to the meaning of accent, that a further thought is to follow (one type of "continuation"):

The latter meaning is an intonational meaning.

Further, there may be a type of accent, in my view, which receives an intonational implication only when it is used in final position in a stretch of information. In Dutch, this type has the following form:

(See Keijser 1984 for the semantic facts leading to this proposal.)

In Russian, the type must now (see Odé 1986) probably be assumed to be:

It is very important to see that these forms are also types of
(pitch) accent. A proposal that they signal only the meaning of accent is a semantic proposal; the analysis has nothing to do with the properties of the physical signal.

Another linguistic analysis which has nothing to do with properties of the physical signal is a proposal that a certain type of pitch accent is a tone, i.e. that it serves to distinguish one word from another. Thus, the Serbo-Croatian falling and rising tones are types of pitch accent, just as are all examples given so far. Only, whereas e.g. Dutch and Russian have one type of pitch accent which signals the meaning of accent without signalling an intonational meaning, Neotokavian Serbo-Croatian must probably be assumed to have two such types (abstracting from length here), the choice depending on which word the speaker wishes to convey (but the choice of a word with a rising accent seems to restrict the possibilities as to the scope of the accent - see 2.3 below). In addition, a certain type of pitch movement may signal an intonational meaning, so that a given configuration maximally signals: 1. the meaning of accent, 2. which word is being conveyed (tone), and 3. an intonational meaning.

It is very important to see that the physical signal itself, and its perception, are indifferent to linguistic functioning as intonation and/or tone. Thus, English

\[
\begin{array}{c}
\text{\textasteriskcentered you did it} \\
\text{\textasteriskcentered you did it}
\end{array}
\]

signal the meaning of accent (in both cases) and an intonational meaning (a different one in the two cases). But as far as the physical signal and its perception are concerned, it could as well be, e.g., the meaning of accent (in both cases) and the meaning of the word you, versus the meaning of the word you.

Now, assuming that everybody agrees that the English example just given, involves accent (prominence) and intonation rather than accent and tone, there may still be disagreement about which analysis is correct. For example, there can be disagreement about the meaning of the fact that you rather than did is prominent, and there can be disagreement about the meanings of the two types of pitch movement. But there is one thing about which no disagreement can be imagined: that the pair just given and the following pair are different pairs:

\[
\begin{array}{c}
\text{\textasteriskcentered you did it} \\
\text{\textasteriskcentered you did it}
\end{array}
\]

(i.e. the same type of accent on you and did, respectively). No intonologist would ever propose that the latter pair is an analysis of the former pair, or vice versa, because, given the fact that the two pairs are different, one of them cannot be an analysis of the other. For an intonologist it is, therefore, very shocking to see that this type of analysis has been proposed for comparable pitch phenomena with a tonal function. Garde (1966: 50), for example, proposes that the Neo-Stokavian rising tone is the realization of an accent on the next syllable, and (1966: 47) that the rising tone in the Cakavian dialect of Novi (which has a different origin from the Neo-Stokavian one, as well as a different form - see L&I 1973(a) and 1986: 75-79) is an accent on the second mora of a (long) vowel. It seems that this type of analysis is based on the idea that, given two types of pitch movement which do not have an intonational function, we can or even must ask the question whether the pitch movements are due to tone or to accent implementation (cf. Gvozdanović (1983: 69), who argues along the same lines, although her particular proposals differ from Garde's). It is unclear to me what the terms "tone" and "accent" are intended to mean in this question. They certainly cannot mean, say, "perceived type of pitch movement" and "perceived prominence", because then the question would be nonsensical, perception of the type of pitch movement and perception of prominence depending on each other. Possibly, they are names for some abstract categories which have no relation to perception; in that case it would be useful to choose different names. Alternatively, they are relics of proposals like that of Jakobson (1931) and Trubetzkoy (1939; cited here from the 1977 edition), but then they are incompatible with the present state of knowledge about pitch phenomena.

Although both Jakobson and Trubetzkoy were aware of the possibility that pitch obtrusions may lend prominence, they had evidently not yet come to understand the implications of the fact that various types of pitch movement may lend prominence in the same place. Thus, Jakobson thought that accent lent by pitch equals an upward pitch obstruction, or high pitch: "Durch die Erhöhung [emphasis
mine, C.K.] oder Verstärkung des Stimmtones heben wir einen Abschnitt in einer Reihe stimmhafter Abschnitte der Rede hervor. Diese Hervorhebung wird "Betonung" genannt" (Jakobson 1931: 164; see also Trubetzkoy 1977: 180). In other words, Jakobson thought that a mode-1 accent is an accent (prominence) per se, not a particular type of accent. He consequently analyzed a type of accent which has the highest point in the second part of a long vowel, i.e. a mode-2 accent, as an accent per se, i.e. as a mode-1 accent, on the second part of the vowel (1931: 167). This idea obviously creates problems for Neo-Stokavian (as well as for all languages where the same phenomena have an intonational function), because the mode opposition applies to short vowels as well. So Jakobson retained the mode-2 accent in Neo-Stokavian and eliminated the "Betonung" as "ausserphonologisch" (1931: 176). It is then true, of course, that the "Betonung" is on the syllable with the mode-2 accent; only, "Betonung" can no longer mean "mode-1 accent" (i.e. a particular type of accent) here, it probably means "prominence", i.e. accent per se. Assuming that falling accents occur only on the first syllable of a word, it is also true, of course, that, if there is no syllable with a mode-2 accent, the "Betonung" is on the first syllable, and that this accent is then a mode-1 accent (1931: 175-176). But it must be doubted whether Jakobson would have proposed this analysis if he had been aware of the fact that he was using the term "Betonung" in two different senses, and of the fact that mode contrasts exist by virtue of the fact that the "Betonung" can remain in the same place while the "Erhöhung des Stimmtones" occurs in different places.

Trubetzkoy even thought that the Neo-Stokavian falling accent is "vorwiegend expiratorisch" (1977: 191). Interestingly, Appel (1950: 63) mentions that Trubetzkoy was acquainted with Appel's (limited) phonetic research on Serbocroatian tones (conducted in 1927 (Appel 1950: 51)). On the basis of his measurements, Appel indeed draws the conclusion that the falling accent is an "expiratory" accent, but his data do not support this conclusion. Appel found that in a vowel with the falling accent, pitch can be level, fall or rise, depending on sentence intonation. He concluded that pitch movement depends on sentence intonation and that, therefore (?), the falling accent is an expiratory accent. The implicit assumption seems to be that an accent called "falling" should fall during the vowel of the accented syllable in order to be a "musical" accent, and/or that, if pitch movements have an intonational function, they cannot be responsible for prominence. At least the latter assumption is shared by Trubetzkoy, who thought that for the "Satzbetonung" in e.g. Russian "die exspiratorischen Stärkeunterschiede zwischen den betonten Silben einzelner Satzglieder phonologisch relevant sind" (1977: 203), and that then only "Satzintonation" (1977: 199-201) involves pitch (see also Jakobson (1931: 169)). In other words, Trubetzkoy and Jakobson had no idea of what we now call various types of pitch accent, also in languages without tones.

This line of reasoning is no longer valid. One should be aware of the fact that people still working along these lines use terminology which is hardly comprehensible to others. Thus, Garde (1966: 45, 49) calls accents in Russian Intensity accents, and Gvozdanovic (1985: 108) associates concepts like "old" and "new" information with types of pitch movement (intonation), although the literature on the subject is about prominence placement (accentuation)(and word order).

It seems to me that the notion of pitch accent, as it has been developed during the last thirty years, should make us aware of the fact that Jakobson's question "Tone or Accent?" amounts to the question (in my terminology, but other terminology would have to express the same): "Shall we analyze the given mode-1 versus mode-2 contrast as a mode-1 versus mode-2 contrast, or as a mode-1 accent in two different places?" This question does not make sense, because the alternatives are not alternative linguistic analyses of the same perceptual reality but different perceptual realities. Thus, if one insists on deriving tone from accent or vice versa, one can, given a physical signal

sensibly do one of two things: either one gives a rule specifying how the pitch obstruction is perceived, so that the place of the accent follows automatically, or one gives a rule specifying where the accent is perceived, so that the perception of the pitch obstruction follows automatically. To give a more useful example, consider a tone contrast resulting from contracting two syllables into one:
Referring to this development, one can sensibly say that the contraction results in a neutralization of a distinction between two accent positions, so that a distinction between the same type of accent in different syllables becomes a distinction between two types of accent in the same syllable. But it does not make sense to say: 1. that the members of the first pair have only accent, not a particular type of accent, and 2. that the difference between the members of the second pair consists in a difference of accent position.

It may be useful to compare a Jakobson/Trubetzkoy type of analysis with a superficially similar autosegmental analysis of mode-2 accents. Ladd (1983: 729-732) proposes relating e.g. the English (late) rise&fall to the (early rise&fall) by means of the feature [+ delayed peak]; thus, schematically:

[-delayed peak]  [+delayed peak]

In this proposal, an (early rise&fall) and a (late) rise&fall are considered to be related in function (see also Gussenhoven's (1983(a): 35-39) slightly different formulation). In my view, the proposal fails to account for the semantic facts, but what is relevant here is that it is understandable. It is so because it recognizes both prominence (here indicated by the star) and pitch obstruction, in both the [-delayed] and the [+delayed] form; thus, the proposal accounts for the fact that different types of pitch movement are perceived in the two cases, the star remaining where it is while the alignment of the pitch obstruction is changed. Ladd suggests (1983: 731) that the same analysis could be used for the Serbocroatian tones. This suggestion is, of course, somewhat odd, because the "delayed" type happens to be an "anticipated" type (see 1.3). A better proposal would be one which derives a "rising" accent from a "falling" accent on the next syllable, schematically:

[-retracted star]  [+retracted star]

One can, of course, question the usefulness of such hocus-pocus-phonology (as a synchronic analysis), but the point is that one can derive one type of accent from another type of accent by either moving the pitch obstruction or the star. But one cannot derive the pitch obstruction from the star, nor the star from the pitch obstruction.

Different yet is a proposal like that of I. (1965), who represents rising tones by means of an abstract accent mark on a syllable boundary (which is phonetically more precise than an accent mark on the next syllable - e.g. I. 1976: 38). This proposal differs from proposals like that of Garde by virtue of the fact that the abstract accent mark is explicitly not intended to be an accent in the sense of prominence. The latter (cf. the star in my pictures) is called a "culminative accent" (e.g. I. 1976: 39; L&I 1986: 252). The accent mark on the syllable boundary corresponds to the highest point of the pitch obstruction; it carries the "distinctive function" (e.g. I. 1976: 39; L&I 1986: 252). The differentiation between "culminative function" and "distinctive function" keeps things understandable, but care must be taken, in my view, not to call the highest point of the pitch obstruction "accent", because this name may easily be associated with the idea that prominence is perceived in the place where the highest point of the pitch obstruction occurs. This does not hold true for mode-2 accents. And, contrary to L&I's suggestion (1986: 256), it is not at all a specific feature of the NeoStokavian rising accent that the distinctive "accent" (the highest point of the pitch obstruction) and the culminative accent (prominence) do not coincide.

In English intonation, the phenomenon is also known as C accent (Bollinger 1958 etc.), Dip and/or Scoop (e.g. Vanderslice and Ladefoged 1972: 821-822; Ladd 1983: 731).

Finally, it can be argued, of course, that the English case
used for comparison here differs from the Serbocroatian case, because in Serbocroatian words a mode-1 accent is excluded in many places (see 1.3). An analysis to the effect that the mode contrast is neutralized in favour of mode-2 in such cases is compatible with a pitch accent approach.

1.5 NOTATIONAL CONVENTIONS

The following symbols represent the various prosodic possibilities:

- "short falling": falling accent on a short vowel
- "short rising": rising accent on a short vowel
- "long falling": falling accent on a long vowel
- "long rising": rising accent on a long vowel
- : unaccented short vowel
- : unaccented long vowel

For example (e.g. L&I 1986: 259-271): сёла (villages, nom.pl.), сёла (village, gen.sg.), мать (mother, nom.sg.), матери (grandmother, nom.sg.), боля (inflammation, nom.sg.) боля (set fire to, 3.sg.aor.), уст (set fire to, 2.sg.imper.), у города (to the town, acc.sg.), у города (idem), города (town, dat.sg.), города (town, loc.sg.).

CHAPTER TWO: PROMINENCE

2.1 CUES FOR ACCENT AND FOR TYPE OF ACCENT

Since the fifties it has become clear that intensity is a very unreliable cue for accent. For Serbocroatian it is especially relevant to note that the syllable following a rising accent may (but need not) have the intensity pattern normally associated with an accented syllable, but that intensity differences alone are unable to distinguish between the falling and the rising types of accent. This has been proved conclusively by Rehder (1968: 163-178)(cf. L&I 1986: 143). For detailed intensity data see L&I (1963: 22-25, 68-75) and the Priolsi series (= I&I 1963-1972), especially I&I 1963 and 1965.

Vowel quality may contribute to prominence perception: short postaccentual vowels, especially /a/ and /u/, and, in the speech of I&I's informant 1. also /o/, are centralized. Postaccentual long vowels, in contrast, resemble accented long vowels more than they resemble postaccentual short vowels. The formant structure of /e/, /o/, and /a/, differs according to phonological vowel length (for /e/ and /o/ I&I's data reflect eastern pronunciation). There is no difference correlating with accent type. For details see I&I (1963: 81-130) and the Priolsi series, especially I&I 1967.

This leaves duration and pitch as the most interesting parameters involved. I&I found that duration is the most important single cue for accent: vowels in accented syllables are consistently longer than vowels of the same phonological quantity in unaccented syllables; on the average, the ratio is approximately 1.5 : 1 (other authors sometimes mention 2 : 1). Phonologically long vowels are also approximately 1.5 times longer than the corresponding short vowels (i.e. accented long vs. accented short, and unaccented long vs. unaccented short). As a consequence, it is not the case that the phonetically longest vowel in a word necessarily occurs in the accented syllable: in a sequence "accented short plus unaccented long" the two vowels may have an equal duration, and the unaccented long one may be longer than the accented short one in individual cases. In " the pitch pattern of the configuration suffices to bring out the first syllable as accented (cf. e.g. L&I 1963: 25), but there are realizations of " for which one must simply know that the first syllable must be taken to be the accented one (L&I 1986: 164), i.e. that the mode contrast is neutralized in favour of mode-2, because the durational and pitch relationships can in principle be identical to those in a hypothetical "'. Within a given dialect only one of " or " can occur because, according to the hierarchy cited in 1.3 above, if the accent has been retracted in " (resulting in "') it has also been retracted in " (resulting in "'), and if the accent has been retracted in " (resulting in "') it has also been retracted in " (resulting in "'). In other words, " and " are not opposed to each other. Evidently, a system where different meanings were correlated to these two forms would meet with considerable problems of communication, precisely because there is no consistent perceptual
cue distinguishing between the two. Details about vowel duration can be found in L&I (1963: 25-29; 75-78) and in the Prilozi series.

L&I's conclusions as to the relevance of duration have been questioned by Gvozdanović (1980: 34 ff.), who, on the basis of L&I's data, concluded that the durational clue is redundant. The author's assumption seems to be that, in order for duration to be a cue for accent, e.g. the first vowel in "--" should be phonetically longer than the second one. This assumption is unwarranted, in my view, the relevant point being the configuration as a whole (as a 'Gestalt'), as opposed to other configurations possible in the same system. It is not the case that a speech signal must contain cues which keep the signal apart from signals which cannot occur in the given system. Thus, a speaker of standard Serbocroatian uttering a word "---" with the longest vowel in the third syllable (Gvozdanović 1980: 34) need not fear misunderstanding because "---" is impossible (as is "---", at least in the opinion of the same author). (In the particular example cited, but not in all examples, there is a pitch cue - see below.) It is possible, of course, that a given signal may be perceived differently by users of different systems, because they listen to the signal with their own system in mind. For example, a vowel with a duration of, say, 155 msec may be judged as short by listeners from Vojvodina and as long by listeners from other areas, because vowels tend to be longer in the Vojvodina region (L&I 1973(b): 188-189). In the same way, it is quite conceivable that a given signal intended to be a realization of "---" can be heard as a realization of "---", or even "---", by people whose system allows these possibilities, or by foreigners. Also, the fact that in some types of realization of the rising accent, pitch is also a cue, does not affect the systematic relevance of duration. In fact, the whole system could hardly exist without duration as a cue for accent, because people would constantly be in doubt about the mode in which the signals must be listened to. (For a somewhat different formulation see L&I 1986: 162-164).

There is one point in the duration data which, I think, is worth investigating further. It concerns the finding (L&I 1973(b): 186-188) that "the same durations are more likely to be assigned to the 'short' category when the word has a rising accent than when the word has either a monotone fundamental frequency or a falling accent. Conversely, durations that are already judged as long with monotone or with falling accents may still be perceived as short when they have a rising accent".

Besides the Vojvodina vs. non-Vojvodina distinction cited above, there is yet another issue which is potentially relevant here. Van Katwijk (1970) instructed a Dutch speaker to pronounce the nonsense utterance /papapap/ with a basic mode-1 accent (risefall) and with a basic mode-2 accent (fallrise) in various syllables. The measurements showed that, in the first two syllables, the mode-1 type lengthens /p/ (in the beginning of the accented syllable) more than does the mode-2 type, and that the mode-2 type lengthens /a/ (in the accented syllable) more than does the mode-1 type (1970: 88). Since the NeoStokavian rising accent can be regarded as a mode-2 accent (see 1.3), the outcome of the duration experiment reported in L&I 1973(b) suggests to me that relative duration of vowels and consonants could be an additional cue for accent. Also, it is not inconceivable that auditory impressions like that of Maretić (1963: 124), viz. that rising accents are longer than falling accents, which impression is not consistently confirmed by measurements of vowel duration in natural speech, can be more adequately explained if vowel duration is systematically related to consonant duration.

Finally, as mentioned above, the pitch configuration in words with falling accents is such that little doubt can arise about which syllable is accented (for illustrations see Chapter Three below). In some types of realization of the rising accents the pitch configuration contains cues which detract from the a priori obviousness of a mode-1 perception (see the beginning of 1.1). Starting from the schematic mode-1 risefall used in Chapter One, these detracting cues can be introduced by retracting the pitch obtrusion somewhat and/or by making the rising movement less steep:

As long as the obtrusion is not retracted considerably (cf. 3.7 below), these pitch cues are probably negative rather than positive, i.e. they make it less obvious that the second syllable in the example is accented rather than evident that the first syllable is.
Additional cues, such as duration and vowel quality, are needed to turn the scales. As we saw in Chapter One, a decision as to which syllable is accentuated automatically involves a decision as to which tone is used, and vice versa.

On the whole, it seems unrealistic to try and find one single cue for accent. Although the situation in NeoStokavian is relatively clear, I think it would be useful to devote further perceptual research to the question of whether pitch obtrusions, whose prominence-lending role in mode-1 is at present no longer seriously doubted, must systematically be accompanied by other cues in order to enable us to interpret a given signal in mode-2.

2.2 WORD ACCENT VERSUS SENTENCE ACCENT

Magner and Matejka (1971) investigated whether, in various places in the NeoStokavian area, the classic prosodic system as described in grammars of the language, is indeed being spoken and recognized. They repeatedly mention the minimal pair I danas \( \ddot{p} \text{ara} / \ddot{p} \text{ara} \) igra ulogu (Even today steam/money plays a role) (1971: 79, 136, 151-152, 163, 171, 172, 183-184). The pair, which is a classic example of the opposition between the short falling and the short rising accents, is one of their examples putatively showing that the prosodic system described by grammars of Serbo-Croatian cannot be taken seriously: even specialists were unable to hear the difference consistently (1971: 171-172). L&I (1986: 147-148) question the correctness of the example: in many places \( \ddot{p} \text{ara} \) means both "steam" and "money". A further problem with the example is, I think, that the sentences were possibly spoken with the last sentence accent on the word danas: Even toDAY steam/money plays a role, because in that case the words under investigation precede the word which probably carries the last sentence accent (danas).

Surprisingly little attention has been paid to the relationship between sentence accent and word tone. According to Hodge (1958; 1965: 536-537) rising tone (for Hodge: disyllabic (double) stress) occurs only under primary stress. De Bray (1959/1960: 393) found that "absence of sentence stress causes a less distinct realisation in pitch of the word accent". Among native authors, Matešić (1970: 38) remarks that "ohne Betonung kann das ortotonierte Wort im Satz vorkommen, wenn es für den Satzinhalt unwichtig ist und sich an das Wort anlehnt, auf welchem der Schwerpunkt der Information liegt. Dies geschieht oft bei gewissen Pronomen wie bei jedan (unbestimmter Artikel)"

(see I. 1965: 138). It is unclear to me how widespread this phenomenon is. Consider, for example, the following sentence discussed by Miletić (1960: 72-73): Prošle nedelje prođe jedan "putnik kroz naše selo (Last week, a traveller came through our village). The sign, (with prošle last), prođe (came through), jedan (one, a), and naše (our) indicates a normal word accent, no sign (with kroz (through)) indicates no accent (a clitic word), the sign ' (with nedelje (week) and selo (village)) indicates a secondary sentence accent, and the sign " (with putnik (traveller)) indicates the main sentence accent. It is implied, as far as I understand, that all words except kroz have a tone. Now, interestingly, if people investigating the meaning of sentence accent placement were to discuss the English translation, they would discuss the sentence Last-LAST WEEK, a TRAveller came through our village, with accents on week and traveller, and possibly one on last: the rest of the sentence would be regarded as unaccented. The Serbo-Croatian sentence is slightly different from the English one because word order is different, but in, e.g., a Russian sentence with the same word order as in Serbo-Croatian, I would add an optional accent on the counterpart of prodje only, leaving kroz naše selo unaccented. Miletić, however, indicates three further degrees after the "rheumatic" accent on putnik: no accent (kroz), a word accent (naše), and a secondary sentence accent (selo). If tones can be realized on naše and selo here, and if the common point of view that Serbo-Croatian has tones only in accented syllables is correct, a question arises as to the sense in which naše and selo are accented. I wonder whether it could not be the case that...
the tone marks given in e.g. a dictionary indicate "potential for tone", i.e. the tone (type of accent) which appears if the given word is accented in a given sentence. To be sure, it is possible to have accents after the "rhematic" accent, also, e.g., in English. If such accents involve the use of pitch obtrusions (rather than, e.g., only duration), the excursions have to be smaller than that of the "rhematic" accent, otherwise the "rheme" becomes the "theme" and the last accented item following it becomes the "rheme" (see Bolinger 1986: 50-51, 111, 182, and the pages listed in the index under the heading of "relative height"). But if Miletid's example means what it seems to mean, it is not a case where accents would appear after the "rhematic" accent in, e.g., English or Russian (cf. also Mahnken 1964: 80-82).

For another example, this time of an unaccented (?) stretch in a "theme", consider the following (Mahnken 1964: 49): pa Bosanaca bilo je nas / cetvoricu (approximately: As to people from BOSnia who were among us, there were FOUR of them). Since the main accent in the first part of the sentence is probably on Bosanaca, and since probably only nas carries a "secondary" accent, it is not self-evident, in my view, that a tone can be realized on bilo, as seems to be suggested by Mahnken.

My second point is related. The list of sentences used for the study of sentence intonation (I&L 1969: 118-123; see also L&I 1968: 274-288) includes series such as the following:

1. Marko gradi pravu barkeru. (Marko builds a real boat.)
11. Marko gradi pravu barkeru!
12. Marko gradi pravu barkeru.
15. Marko gradi pravu barkeru.

The quantitative data for the last word of these sentences (I&L 1969: 125) are given separately for examples 11 and 15, but the realizations of barkeru in 1, 12, 13 and 14 (and 20, which is comparable to 1) are averaged. The same holds true for all other comparable series in the list. I think it is unfortunate that type 1 has not been systematically treated as different from 12, 13 and 14 (but see I&L 1969: 133 for some data of individual cases). The word barkeru in 1 was probably spoken with a so-called "neutral" sentence accent, which is also a sentence accent (see Keijser 1985: 19 ff.). If tone oppositions are affected in this type of sentence, this must be ascribed to the type of sentence accent used here (cf. 4.6 below). In contrast, if the "emphatic" accent in 12, 13 and 14 is a "rhematic" accent, there is a preliminary question as to whether there is any accent on barkeru at all (see above). I think that I&L's (1969: 141-142) observation that there were fewer cases of neutralization of the distinction between the two short accents in sentence final position in Tada se potpuno pribijio ... (Then he came completely close to
... than in series like 1, 12, 13 and 14, should make us aware of a possible difference between words with and words without sentence accent. Thus, the appearance of a word in sentence final position without "emphasis" (as in 15) and without "exclamation intonation" (as in 11) is no sufficient reason, in my view, to treat the remaining cases on a par, because any more or less sophisticated treatment of sentence accent in a language without tones would call barking accent in 1 but unaccented in 12, 13 and 14 (if the "emphatic" accent is "rhythmic"). It has not yet been shown, so far as I know, that the situation in Serbocroatian is different because of the existence of pitch accents with a tonal function, so for the time being we must reckon with the possibility that the situation is the same.

2.3 WORDS

The problem as to the relation between word stress and sentence accent signalled in 2.2 is the more important because, from the point of view of what is known about the functioning of sentence accent in languages without tones, the rising accent does not seem to behave in a normal way.

In application to, e.g., English, I have proposed (1985: 109 ff.) that a distinction be made between accent (prominence) per se and the scope of an accent. An accent, by meaning "not not", basically introduces an idea into the fund of knowledge shared by speaker and listener (a "rhythmic" accent) or selects an idea from this fund (a "thematic" accent). The unaccented elements which are introduced or selected simultaneously with the accented item are said to fall inside the scope of the accent; unaccented elements which are understood to have been introduced or selected prior to the accented item are said to fall outside the scope of the accent. For example, in JOHN died, died may be included in the scope of the accent on John. In that interpretation "John" and "died" are introduced together, as a complex chunk of information; the sentence might be an answer to the question: What happened? Died may also fall outside the scope of the accent on John. In that interpretation the sentence might be an answer to the question: Who died?, i.e. the sentence introduces only the idea "John", whereas "died" is understood as having been introduced earlier, before the sentence was spoken. In John DIED, in contrast, "John" cannot be introduced together with "died", i.e. John cannot fall inside the scope of the accent on died. In the example, JOHN died has two interpretations (the accent may have broad or narrow scope), and John DIED has only one interpretation (the accent can only have narrow scope). This asymmetric relationship is, in the functioning of sentence accent, a regularly recurring phenomenon. In my treatment, its cause is the fact that the elements of a sentence are linked together in one direction or another; in the example, "died" is linked to "John" (not "John" to "died"): "John died". Let me call this asymmetric patterning the "sentence accent pattern"; the essential point to note is that an accent which can be interpreted as having broad scope (in the example: JOHN died) can always be interpreted as having narrow scope, but the converse does not hold true (John DIED has only narrow scope). The separation of accent per se and the effect of linkage (scope) implies that, e.g., JOHN died is opposed to John DIED indirectly, via JOHN DIED (with two accents); thus, the minimal pairs are JOHN died vs. JOHN DIED, and JOHN DIED vs. John DIED. In other words, it is not the place of an accent which is semantically relevant (accent on John vs. accent on died), but the presence vs. absence of an accent in a given place (see Keijser 1985: 109-115; 192-195).

When the asymmetric pattern sketched above is destroyed, a new type of construction arises. For example (Bolinger 1986: 71; my notation), people start out saying He JOINed the Boy SCOUTS. Then, other groups are organized: GIRL Scouts, SEA Scouts, etc.; here, an accent is placed on the first element in order to contrast this element with Boy; there is no accent on the second element, in order to signal that this element is the same as in the already existing combination Boy Scouts. Finally, it becomes normal to say GIRL Scouts, SEA Scouts, and also BOY Scouts, also when no contrast is in mind. This is a process of compound formation. In my notation, the steps of the process are: "Boy SCOUTS" (broad and narrow scope), "GIRL Scouts" (only narrow scope), "GIRL Scouts" (broad and narrow scope), "GIRL/BOY Scouts". The last step is, as indicated, that the components of the construction are no longer linked together in one
direction or another at the moment they enter into the speech chain; instead, they enter as a single complex. A semantic difference exists between combinations whose parts are linked together at the moment of speaking and combinations which are treated as a single complex; this difference has been described in Keijsper (1986) in application to Russian ne vs. ne-. For example, in "ne prijatel' priechal" (not a FRIEND arrived), "ne" and "prijatel'" make a separate contribution to the total information: somebody who arrived, and who was first thought to be a friend, in reality lacks this property. In "neprijatel' priechal" (a not-FRIEND (enemy) arrived), in contrast, "ne" and "prijatel'" together function as a single concept, which directly evokes the thought of the thing meant: a person identified by the property that he lacks the property of being a friend.

Now, the "sentence accent pattern" as given above may also apply to single words in the sense "combinations whose parts (if any) are not linked together at the moment of speaking". For example, the accent in BLACKbird may include the entire word in its scope (broad scope), but it may also be interpreted as focusing on the element black- only, for example in order to contrast the word to BLUEbird (narrow scope). We may occasionally say blackBIRD or BLACK-BIRD if we wish to contrast -bird with e.g. (black)BOARD (narrow scope). In this sense, a word may behave as a sort of mini-sentence. The stress of a word marked in a dictionary indicates the syllable where an accent may include (at least) the entire word in its scope. There are many cases of lexical stress position which are clearly not arbitrary from the point of view of sentence accent. Thus, it would be very odd if the stress of the word blackbird were on the second part; as it is, on the first part, the lexical stress position reflects the process of compound formation sketched above. Also, we may put an accent in a lexically deviating place in a word in order to signal a meaning which is normally signalled on the level of a sentence only; for example, we may put an accent as near as possible to the end of a word in order to signal a meaning which, if word order is flexible, is signalled by putting an accented word at the end of a sentence. As long as it is clear which word is being conveyed, accent placement is extremely flexible (see Bolinger 1986 for numerous examples).

The reason for bringing up all this here is that the scope story does not apply, as far as I understand the situation, to the Neoštokavian rising accents, while it does apply to the falling accents. From a synchronic semantic point of view this is very remarkable. My attention was drawn to the rising accents by Kravar 1963 and 1968, who discusses a use of falling accents which is obviously remarkable from the point of view of students of word stress. Kravar describes the "sentence accent pattern" indicated above, used on the level of a single word (in the sense: "combination whose parts (if any) are not linked together at the moment of speaking"). For example, he explains (1963: 215-216) how the word nepušči (non-smokers) may become nepušči for semantic (as opposed to historical phonological) reasons: although one can say Ovdje smo sví nepušči (We are all non-smokers here), namely if one does not think of pušči (smokers), people will be inclined additionally to mark the contrast between the two words: Ovdje smo dva pušča i tri nepušča (Among us are two smokers and three non-smokers). Finally, only the initial falling accent is used, also if no contrast is intended: Ovdje smo sví nepušči (We are all non-smokers here). Other examples (Kravar 1968) include the following. Besides instrumental people also use instrumental, because, according to Kravar, they put it in the associative series instrumental, kondicional, pálatl etc. Likewise, potencijal is used for the grammatical term, because of kondicional, išestl, etc., but potenciál is used in the sense "power, strength", because there is no such associative series in the latter case. And some dictionaries give aktiv for the verbal form, because of the association with pasXv; they give aktiv for "group of activists". In other words, the falling accent, but not the rising accent, can focus upon only a part of a word.

Kravar's observations agree with Belić's (1971 (originally 1948): 95-96) remark that words may be spoken with non-retracted accentuation (devojka instead of devojka (girl), etc.) if special emphasis is needed. Also, Van Katwijk (1974: 151-152) noted that mode-2 intonation does not allow for emphatic accentuation (which has a physiological basis). (This fact is reflected in the meanings of mode-2 accents with an intonational function, but it has no consequences for scope there, although it is often associated with it, for example when people do not indicate a "neutral" sentence-final fall (mode-2) because it is not "emphatic" and is felt to be an
What we need, then, in addition to the "sentence accent pattern", is the "word accent pattern". It deals with cases where we can only speak of the place of an accent (rather than of the presence vs. absence of an accent in a given place), i.e. when the minimal opposition is "accent on x versus accent on y". It seems that we must associate this type of opposition with accents which can only have broad scope. (Accents which have broad scope appear, then, in both patterns.) Kravar's discussion suggests that the rising accent, in contrast to the falling accent, can function only in the word accent pattern. It, indeed, has its origin in this pattern: an accent on the nth syllable turned into an accent on a n+1st syllable (see 1.3) rather than: an accent on the n+1st syllable is added to the accent on the nth syllable, and the accent on the nth syllable disappears, as in Kravar's nepušači-example. It seems that the mode shift which gave rise to the rising tone has had the semantic consequence that the rising tone, being a mode-2 accent, cannot be interpreted as having narrow scope.

The difference between the "sentence accent pattern" and the "word accent pattern" must be kept apart from the difference between combinations with a link and combinations without a link: as we saw above (blackbird, nepušači), the sentence accent pattern may also apply to combinations whose parts are not linked together at the moment of speaking. The complementary issue for the word accent pattern is then, of course, the question of whether this pattern may also apply in combinations with a link between the elements. It seems that combinations like u kucu ('into the house'), i.e. where the NeoStokavian accent retraction has crossed a word boundary (see 1.3), are examples of this possibility. Thus, the relevant question to ask here is: if one wishes to exclude kucu from the scope of the accent on y (the house, and not, e.g., over the house) must the rising accent be replaced by a falling accent? (cf. Kravar 1968: 211). If it must, we can expect a difference to exist between the NeoStokavian type Ũ kucu and the older Slavic type u vodu ('into the water'; this type already existed at the time when the NeoStokavian accent retraction took place): so far as I understand, the accent in Ũ vodu can be interpreted as having narrow scope. (Basically, the difference derives then, from the fact that one can say Ũ vodu, but not Ũ kucu.)

As to accent position, it is important to keep apart the scope of an accent from the normal semantic effect of putting an accent or not putting an accent somewhere. Consider, for example, the following sentences (ISL 1967: 75-76): Pogao je u grad (He went to the city), and Tekovine napretka prodiru u na selo, a ne samo u grad (The results of progress come also to the countryside, not only to the city). The difference is possibly that in Pogao je u grad, going to the city is opposed to not going to the city, whereas in the u grad sentence, the city is viewed as one among other places. The fact that in u grad no thought of other places is evoked, is the normal effect of the absence of an accent on grad, and the fact that this thought is evoked in u grad is the normal effect of an accent on grad. Given these accentuations, a narrow scope interpretation is only one of the possibilities. In u grad, it applies, roughly, if one thinks of various combinations of a preposition and grad, and in u grad, it applies if one thinks of various combinations of y and a noun. As illustrated by the example, u grad can be opposed to na selo, i.e. the accent on grad may include the preposition in its scope (oppositions need not only be made with y plus noun). Another question is whether the accent also includes the verb in its scope; the (few) examples I have seen suggest that the verb tends to be included in the scope of the accent in retracted types (u grad, u kucu), whereas it tends to be excluded from the scope of the accent in the non-retracted type (u grad, u kucu). Thus, in ISL's example, the u grad sentence conveys the fact that he went, but in the u grad sentence the fact that the results come is treated as "given" information. The semantics of such combinations can obviously not be described without much material from various dialects and reliable
native intuitions. I would suggest that an adequate description will need, beside the normal regularities of accent, the difference between the word accent pattern and the sentence accent pattern, and the difference between combinations with a link and combinations without a link.

Finally, it must perhaps be mentioned that neither the distinction between the word accent pattern and the sentence accent pattern, nor the distinction between link and no link, is intended to save the textbook rule according to which a falling accent on a non-initial syllable of a word is impossible (see 1.3). But in application to Ĳ kuću/kuću I would suggest that Ĳ kuću consists of one word in the sense that the accent can (presumably) only function in the word accent pattern, and consists of two words in the sense that the elements are linked together at the moment of speaking (in contrast to "adverbial" combinations); but exclusive functioning in the word accent pattern obviously implies that we cannot oppose to each other various types of link (as a forward link would give narrow scope for the accent). Examples like Ĳ kuću consist of two words both in the sense that the accent can function in the sentence accent pattern and in the sense that there is a (forward) link between the elements; given this configuration, a narrow scope interpretation is only one of the possibilities. As a consequence, neither of my distinctions (concerning pattern and link) corresponds to the notion "prosodic word" as applied by Gvozdanović (e.g. 1980: 40 ff.) to the same combinations. I find the application of this notion highly confusing because, if the dialectal differentiation postulated by Gvozdanović (1980: 40 ff.) existed (Gvozdanović 1984-1985: 177-178; L&I 1986: 168), one would, for Ĳ kuću, prefer to have two words where the author has one, and vice versa. For example, in a hypothetical dialect where Ĳ kuću can only have a "contrastive" interpretation, the elements are merged to such a degree that this accentuation is felt to be a deviation from the lexically normal Ĳ kuću. Compare: if "black BIRD" changes, via "BLACK bird" and "BLACK bird", into BLACKbird", blackBIRD is interpreted "contrastively". One could say that "black BIRD" (which has a "non-contrastive" interpretation) consists of two words, and that "blackBIRD" consists of one word, but hardly the converse. Possibly, Gvozdanović's intention can be expressed if, instead of the notion "prosodic word", a notion such as "focus domain" (Gussenhoven 1983(b): 391 ff.) were used. Thus, instead of saying that the accent in black BIRD may include black in its scope, Gussenhoven would say that the phrase may be one focus domain, and instead of saying that the accent in blackBIRD does not include black- in its scope, Gussenhoven would probably say that the word consists of two focus domains here. I find Gussenhoven's use of the notion "focus domain" objectionable too (for reasons irrelevant here), but the name is less confusing than one associating contrastive/non-contrastive readings with word boundaries, whether or not "prosodic" ones.

CHAPTER THREE: TONE

3.0 INTRODUCTION

This chapter examines some aspects of the forms of the falling and rising tones as they appear in I&L's data in circumstances in which the influence of sentence intonation is (considered) minimal. Also, a few "technical" points are discussed which have to do with the relation between acoustic data and perception. In the history of phonetic research on the Serbocroatian tones, detailed acoustic measurements preceded perceptual evaluation of the findings. I think that it is useful to make explicit how this history is reflected in I&L's data. As L remarked in her review of Purcell (1973)(L 1976: 144), a parameter-by-parameter presentation tends to obscure the fact that a listener probably bases his identification of the physical signal "on a Gestalt-like overall impression that integrates all perceivable parameters and compares the total accentual pattern with a kind of stored overall image". My discussion is intended to expound the "overall impression" rather than the acoustic detail.

3.1 LINEAR AND LOGARITHMIC SCALES

Data as to fundamental frequency (F0) are often given in hertz (Hz). For linguistic purposes, a statement like "here we find a rise
of 50 Hz" is practically useless, because for the human ear, a rise from, say, 50 to 100 Hz is approximately four times as large as a rise from, say, 250 to 300 Hz. This is because the Hz scale is linear, whereas human pitch perception is (usually regarded as) logarithmic. Therefore, perceptually oriented data usually employ a logarithmic scale, for example semitones (ST). Data given in ST directly express the size of a movement as it is for the human ear. For example, a rise from 50 to 100 Hz is 12 ST, a rise from 250 to 300 Hz is 3.2 ST. The following formula translates pitch excursions given in Hz into ST (De Pijper 1983: 14):

\[
\text{size} = \frac{12}{\log 2} \cdot \log \left( \frac{f_2}{f_1} \right)
\]

where size = size of a frequency change in semitones
f2 = end frequency of a frequency change in hertz
f1 = start frequency of a frequency change in hertz

Positive outcomes indicate rises, negative outcomes indicate falls.

The musical scale is a logarithmic scale. An octave is 12 ST, a seventh is 11 ST, a major sixth is 9 ST, a fifth is 7 ST, a fourth is 5 ST, a major third is 4 ST, a minor third is 3 ST, a second is 2 ST.

Most frequency data in the phonetic literature on Serbo-Croatian tones (and intonation) are expressed in Hz, but sometimes a musical indication is given, and L&I 1986 sometimes use ST. In the present chapter, all excursions are given in ST (derived from the original Hz data by means of the above-mentioned formula).

The acoustic orientation of the literature is also reflected in the classification of informants. Thus, L&I's informants consist of a group of speakers with a low-pitched voice, a group of speakers with a medium-pitched voice, and a group of speakers with a high-pitched voice (e.g. L&I 1963: 38 ff.). Although this classification is unavoidable if one wishes to average Hz data of different speakers, it is not necessarily the best possible classification. Since it has now become clear that regional differences exist (see especially L&I 1986: 111-120), one could consider the possibility of classifying informants according to their dialectal background, irrespective of whether they have, e.g., a low or a high voice. In that case average figures can be given in ST but not in Hz.

When one wishes to give a visual impression of a pitch movement, Hz data can be used directly if they are plotted on a logarithmic scale. In this respect, the illustrations given in e.g. L&I (1963: 18) represent perception (but not, e.g., L&I 1986: 201-208). In the present chapter all illustrations will be plots on the following scale, unless it is stated otherwise:

\[
\begin{array}{c}
500 \\
450 \\
400 \\
350 \\
300 \\
250 \\
200 \\
150 \\
100 \\
100 200 300 400 etc. \\
\end{array}
\]

\[
\begin{array}{c}
500 \\
450 \\
400 \\
350 \\
300 \\
250 \\
200 \\
150 \\
100 \\
\end{array}
\]

\[
\text{TIME (milliseconds)}
\]

The F0 scale (vertical) is logarithmic, the time scale (horizontal) is linear.

The slope of movements (indicating whether they are steep or gradual) can be expressed in ST per second. Differences as to slope must be rather large in order to be perceptually relevant; as a rule of thumb one can say that a movement must be twice as steep/gradual as another one (expressed in ST per second) in order to have a clear perceptually different slope ('t Hart, pers. comm.). During the present discussion almost no slope data can be given, as no data are available specifying consonant duration. Since slope appears to be relevant in some cases, I sometimes use formulations which in a roundabout way refer to slope.

3.2 CONSONANTS
Acoustic data concerning pitch movement are often given for vowels (and syllabic /r/) only, either because no pitch can be measured in voiceless consonants or because of the theoretical point of view that prosodic phenomena pertain to vowels only. Thus, LSI give the F0 level in the beginning of a vowel, at the highest point reached during a vowel, and at the end of a vowel. Except for Peco and Pravica (1972) (see below) all phonetic publications on the subject which I have seen make the same choice (some pictures in Miletic (1926) give additional information). Considerable care is taken that vowels are separated from the surrounding consonants (e.g. Purcell 1973: 40-43). I think this choice is unfortunate. As I have tried to suggest in Chapter One, by indicating approximate syllable boundaries, pitch obtrusions are often not aligned with the segmental sequence within the boundaries of a syllable, let alone within the boundaries of a vowel. It is, in my view, one of the most important innovations of (most of) the prosodic research conducted at the Institute for Perception Research at Eindhoven (see the Introduction) that it describes entire movements, irrespective of the segments with which they are aligned, treating virtual movements (in voiceless stretches) on a par with measurable ones. There appears to be no perceptual basis for doing otherwise. If, for example, pitch jumps from the end of a preaccentual vowel to the beginning of the vowel in the accented syllable, the jump is heard as if there were a real movement: our mind "fills in" the missing part (see Odé 1986: 433-434 for some examples). The issue is, in my view, of the utmost importance, because phonetic descriptions which take into account movements during consonants are different from descriptions based on measurements of pitch movements made during vowels only. For example, as mentioned in 1.1 above, a Dutch rise-fall corresponds to what is called a fall in the British tradition: the (mode-1) rise meant here is a prominence-lending rise which starts at ca. 50 msec before the onset of the vowel in the prominent syllable (e.g. Van Katwijk 1969: 70); the British name "fall" disregards this early movement.

ISL (1963: 36) devote a footnote to consonants. They observe that interesting phenomena occur during consonants, and that it would be useful to study them. They leave them out of consideration, however, because (ibid.) F0 movements taking place outside vowels are redundant, these movements being absent in voiceless consonants, and because movements in a vowel alone are not equivalent to movements in a sequence of e.g. a vowel plus sonant. This footnote announces a strictly acoustic approach. From a perceptual point of view it would be more advantageous to treat virtual movements in voiceless stretches on a par with real movements in voiced consonants than to disregard the latter because the former cannot be measured. We hear entire configurations; thus, a pitch movement in the first syllable of a sentence always presents special problems, exactly because there is no preceding syllable with respect to which the movement can be evaluated perceptually.

Fortunately, rising accents often occur in a non-initial syllable of a word and do not occur in the last syllable of a word (see 1.3): since Hz data are available for all vowels in a word, the information which interests me can be inferred. Sometimes, ISL mention explicitly what happens during a consonant (e.g. ISL 1963: 57). In the case of falling accents the available data are more problematic, because falling accents mostly occur in the first syllable of a word (see 1.3). Since most data concern separate words, they omit a part of the potentially perceptually relevant information, namely about what happens before the beginning of the first vowel of a word. The few data concerning sequences of a proclitic word and a word with a falling accent (type u k'úcu) clearly illustrate my point (ISL 1963: 39, 51; 1967: 77-78). For example, in the type "spoken by I. (ISL 1963: 39) the pitch movement in Hz is as follows (average values given by ISL throughout):

<table>
<thead>
<tr>
<th></th>
<th>first vowel</th>
<th>second vowel</th>
<th>third vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>begin</td>
<td>189</td>
<td>223</td>
<td>187</td>
</tr>
<tr>
<td>highest point</td>
<td>193</td>
<td>244</td>
<td>187</td>
</tr>
<tr>
<td>end</td>
<td>186</td>
<td>226</td>
<td>150</td>
</tr>
</tbody>
</table>

This means that in the consonant(s) between the first and the second vowels the distance from 186 Hz to 223 Hz is covered, i.e. there is a (virtual or real) rise of 3.1 ST. The rise continues in the second vowel, from 223 Hz to 244 Hz (1.6 ST). Thus, in total, a rise of 4.7 ST is made, of which only the last part (1.6 ST) occurs in the vowel of the accented (second) syllable. Exactly which part of such a (mode-1) rise is placed in the vowel itself may well be inaudible
(Odé 1986: 431-432). The pictures of whole sentences given in L&I 1986 (201-208) suggest that the initial early rise is regularly present in falling accents (also if the acoustic data give only the last part of it). (Possibly, the correct starting point to take in the example is 193 Hz, or, say, 190 Hz, but this does not affect the issue.)

After the rise in the example, pitch begins to fall. In the vowel of the second syllable it falls from 244 Hz to 226 Hz (-1.3 ST), in the consonant(s) between the second and third vowels it falls from 226 Hz to 187 Hz (-3.3 ST), and in the third vowel the distance from 187 Hz to 150 Hz (-3.8 ST) is covered. Thus, from the highest point of the configuration (244 Hz) to the end of the third vowel, pitch falls: 8.4 ST. As will appear further on, it is probably relevant to note that at the beginning of the third vowel, pitch has returned to the level from which the obtrusion started in the first vowel.

A picture of the example, on the scale given in 3.1, taking into account the data as to the location of the highest points in the vowels (I&L ibid.) appears as follows:

![Diagram of pitch configuration]

In this picture, as well as in all illustrations following below, (virtual or real) movements during consonants are represented by interrupted lines. Since the duration of consonants is unknown, the slope of these movements in the pictures does not necessarily correspond exactly to reality.

If one looks at I&L’s data in the way suggested here, the discussion in the literature as to whether the location of peaks is relevant appears in a somewhat different light. Since there has been a lot of confusion about the issue, it will be treated in detail in the following sections.

There has been an issue about consonants, too. Purcell (1973: 92-103) found that, normally, the fundamental frequency curve in the vowel of the accented syllable starts higher with falling accents than with rising accents. In his data, trisyllabic words with long accents constitute an exception; here, the situation is the reverse. L (1976: 143) adequately explains the exception by calling attention to the consonants preceding the accented vowel in Purcell’s trisyllabic long cases. Unfortunately, balance was lost here: the consonants were mostly voiceless in the long rising category and mostly voiced in the long falling category; and, as L points out, “it has been known for some time that the fundamental frequency curve starts at a low frequency after voiced consonants, whereas after voiceless consonants, the curve starts practically at its peak value [Peterson and Lehiste, 1960].” L’s reasoning is confirmed by Jacobsen (1967: 98-99), who gives separate data for vowels after voiced and after voiceless consonants (cf. L&I 1986: 140). According to L (1976: 143), “the influence of initial consonants on the starting frequency of the fundamental frequency curve makes it improbable that the starting frequency could play a significant part at the suprasegmental level”. This is correct as long as one takes the starting frequency to be the frequency at the beginning of the vowel, but not if one adopts the suggestion that the relevant starting point is located earlier, in the preceding syllable (if any). If one takes this earlier point, the problem of the automatic variation caused by the type of consonant can be eliminated, because the same starting point can be taken before voiced and voiceless consonants (’t Hart, p.c). (The effect of consonants must, of course, be looked into if one studies tones having a consonantal origin, but here it can safely be disregarded.) Finally, L&I (1986: 149) note that Peco and Pravica (1972: 212-214) present tables containing measurements of F0 movements purportedly taking place during voiceless consonants, which, as remarked by L&I, is impossible. Although one cannot, of course, measure things which cannot be measured, this is not a sufficient reason for disregarding virtual movements in voiceless consonants, because one can imagine to hear things which cannot be measured.

3.3 VOWELS VERSUS CONFIGURATIONS
LSI have argued repeatedly that, whereas the distinction between the two long accents, ~ and ' , might "conceivably be based on information contained in the accented syllables themselves" (LSI 1963: 132), the information in the F0 contour in the vowel of the accented syllable is insufficient for keeping apart the two short accents, • and ' . For both long and short accents, "the feature which was constantly present and appeared to carry the main burden of the distinction was the relationship between the stressed and the posttonic syllable. In the case of both the short and the long falling accents, the posttonic syllable had a low fundamental frequency; in the case of both rising accents, the posttonic syllable had a fundamental frequency that was either the same or higher than that of the preceding syllable" (ibid.). In a footnote (LSI 1963: 135) the authors add that a series of listening tests showed that, when the consonants between the accented and the postaccentual vowels plus the postaccentual vowel were eliminated, identifications were completely random. Inclusion of a voiced intervocalic consonant rendered identification much better than chance. Inclusion of the complete second syllable resulted in complete identification. (Note C.K.: in the case of voiceless intervocalic consonants, at least the beginning of the next vowel is needed in order to enable us to "fill in" mentally the (virtual) movement in between.) LSI's data, which mainly concern northeastern pronunciation, indeed allow no other conclusion (the situation in (south)western pronunciation is somewhat different - see 3.6-3.7 below). The pitch movement in the vowel of the accented syllable may be the same or almost the same in the case of • and ' . For example, in the types • and ' , as spoken by informant D8, the pitch movement in the vowel of the accented syllable is as follows (average values throughout):

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant D8</th>
<th>LSI 1963: 49</th>
<th>I&amp;L 1965: 84</th>
</tr>
</thead>
<tbody>
<tr>
<td>' (entire configuration)</td>
<td>informant D8</td>
<td>LSI 1963: 49</td>
<td>I&amp;L 1965: 84</td>
</tr>
<tr>
<td>^ (entire configuration)</td>
<td>informant D8</td>
<td>LSI 1963: 49</td>
<td>I&amp;L 1965: 84</td>
</tr>
</tbody>
</table>

Thus, in both cases a predominantly rising movement was found (highest point relatively late in the syllable). Other informants

had, for example, a predominantly rising accented vowel in the so-called "falling" accent ' , and a predominantly falling accented vowel in the so-called "rising" accent . For example, during the same vowels in the speech of informant E13 the following pitch movements are made:

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant E13</th>
<th>LSI 1963: 53</th>
<th>I&amp;L 1965: 86</th>
</tr>
</thead>
<tbody>
<tr>
<td>' (entire configuration)</td>
<td>informant E13</td>
<td>LSI 1963: 53</td>
<td>I&amp;L 1965: 86</td>
</tr>
<tr>
<td>^ (entire configuration)</td>
<td>informant E13</td>
<td>LSI 1963: 53</td>
<td>I&amp;L 1965: 86</td>
</tr>
</tbody>
</table>

Thus, it seemed that the traditional names "falling" and "rising" were completely at odds with the facts.

The greater part of this mystery is resolved if we complete the pictures; we then see why LSI regarded the relationship between accented and postaccentual syllables as the relevant point:

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant D8</th>
<th>LSI 1963: 49</th>
<th>I&amp;L 1965: 84</th>
</tr>
</thead>
<tbody>
<tr>
<td>' (entire configuration)</td>
<td>informant D8</td>
<td>LSI 1963: 49</td>
<td>I&amp;L 1965: 84</td>
</tr>
<tr>
<td>^ (entire configuration)</td>
<td>informant D8</td>
<td>LSI 1963: 49</td>
<td>I&amp;L 1965: 84</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant E13</th>
<th>LSI 1963: 53</th>
<th>I&amp;L 1965: 86</th>
</tr>
</thead>
<tbody>
<tr>
<td>' (entire configuration)</td>
<td>informant E13</td>
<td>LSI 1963: 53</td>
<td>I&amp;L 1965: 86</td>
</tr>
<tr>
<td>^ (entire configuration)</td>
<td>informant E13</td>
<td>LSI 1963: 53</td>
<td>I&amp;L 1965: 86</td>
</tr>
</tbody>
</table>

Thus, in both cases a predominantly rising movement was found (highest point relatively late in the syllable). Other informants...
Thus, both informants produce a large fall during the consonant(s) between the first two vowels in "^^" (-6.3 ST in the case of D8, -5.8 ST in the case of E13). Both informants make the rising movement of the rising accent partly or entirely in the consonant(s) between the first two vowels and in the second vowel; from the end of the vowel in the accented syllable to the highest point in the postaccentual syllable D8 covers +3.0 ST, and E13 +4.3 ST).

In the text accompanying the quantitative data, I&L make very detailed comments on the pitch movement in every individual vowel. Such a discussion necessarily distracts the reader's attention from the configuration as a whole, so that it is not always immediately evident which status must be assigned to the movement described. For example, as we saw above, the pitch movement in the vowel of the accented syllable in "^^" and "^^" spoken by D8 is almost the same; but the status of this movement in the configuration as a whole is entirely different: in "^^" it is the peak of the pitch obtrusion, whereas in "^^" it is only the preparation for an upward pitch obtrusion which is yet to come. It is quite possible that listeners hearing the two configurations will have great difficulty in recognizing the sameness of the movements in the first vowels, because the perceptual evaluation of a pitch movement is made with respect to its melodic context. Also, it is quite possible that listeners will hardly be able to hear the difference between the first vowel of "^^" as spoken by D8 and the same vowel as spoken by E13, despite the fact that the highest point is reached late in the vowel in the former case and early in the latter case (see the pictures); the subsequent movements in the consonant(s) and in the second vowels may well overshadow the difference in the first vowel perceptually.

When reading I&L's detailed descriptions of pitch movements, it is very useful first to make a picture on the basis of the quantitative data, in the way explained in 3.1-3.2 above, so that one has an impression of whether one is reading about some minor detail or about an extremely important point; the difference is not always clear from the descriptions themselves. Also, one must be aware that a vowel-by-vowel presentation tends to suggest comparisons which are not necessarily the best ones. For example, when we put the movement in the vowel of the accented syllable of the type "^^" as spoken by I, beside that of E13 (repeated here for convenience), we compare the following movements:

```
type "^^" (accented vowel)          type "^^" (accented vowel)
informant I                      informant E13
I&L 1963: 19                     I&L 1963: 53
I&L 1963: 44                     I&L 1965: 86
```

Clearly, the highest point is late in the vowel in I.'s case but early in the case of E13. However, if we take into account the entire configurations, it appears to be more natural, in my view, to compare the movements indicated by the arrows in the following pictures:

```
type "^^" (entire configuration)   type "^^" (entire configuration)
informant I                      informant E13
I&L 1963: 19                     I&L 1963: 53
I&L 1963: 44                     I&L 1965: 86
```

Thus, I. uses the entire vowel of the accented syllable to gradually go to the highest point in the configuration in the beginning of the postaccentual syllable (with a dip around the intervocalic consonant(s)), whereas E13 performs a steeper rise in the intervocalic consonant(s) and in the postaccentual vowel. (The slope of the first part of I.'s rise (in the vowel of the accented syllable) is 20.3 ST per second, the slope of the last part of E13's rise (in the vowel of the postaccentual syllable) is 30.3 ST per second.) It is evidently this rise which induces people to call the accent a "rising" accent. It is also true, of course, that pitch in the postaccentual vowels in the examples starts as high as, or higher than, the highest point reached during the accented syllables.

A third type of realization of the short rising accent represented in I&L's data is one without a clear rise but with (as far as
the data go) a fall starting from the postaccentual syllable, for example:

\[ \text{type} \quad \text{informant D2} \]
\[ L&I 1963: 44 \]
\[ I&L 1965: 82 \]

Alternatively, pitch may remain high in the third syllable (e.g. type "" spoken by EI - L&I 1963: 50; I&L 1965: 85), which is possibly a different intonation. If pitch remains on the same level in an absolute sense we perceive a slight rise, because the normal case is that there is a gradual downward drift (declination).

The short falling accent also occurs in a relatively gradual and in a relatively steep variant; this time the different realizations pertain to the falling movement after the peak, for example:

\[ \text{type} \quad \text{informant I} \]
\[ L&I 1963: 19 \]
\[ I&L 1965: 38 \]
\[ \text{type} \quad \text{informant EI} \]
\[ L&I 1963: 51 \]
\[ I&L 1965: 50 \]

The effect is, of course, the same; already at the beginning of the postaccentual vowel, pitch is considerably lower than the highest point reached during the accentuated syllable (-3.9 ST in I.'s case, -8.1 ST in EI1's pronunciation). Note that the rising parts of the obtrusions must probably be seen as continuations of rises which have started earlier (see 3.2 above). The size of the fall is possibly regularly larger in the "steep" pronunciation than in the "gradual" pronunciation (cf. 3.6-3.7 below).

Since I&L give a vowel-by-vowel description, and since they do not always mention what happens during consonants, comparisons of the type suggested here are not always made explicit. But if one is aware of the possibility, they can easily be derived from their quantitative data, which then agree remarkably well with e.g. Belli's (1926-1927) auditory description.

### 3.4 Long Versus Short

If in English we first say John and then repeat the word with the same type of pitch accent while stretching out the vowel, the pitch movement will be stretched out as well: Jo-o-ohn. This is not what happens in I&L’s data when falling and rising accents appear in syllables with phonologically long vowels. Instead, the additional length is, essentially, added after the peak in the case of falling accents, and before the peak in the case of rising accents (so that the long accents are more clearly distinct from each other in the accented syllable itself than the short accents). For example, in the examples spoken by I. (average values throughout) the types "" and "" have the following form:

\[ \text{type} \quad \text{informant I} \]
\[ L&I 1963: 19 \]
\[ I&L 1963: 38 \]
\[ \text{type} \quad \text{informant EI} \]
\[ L&I 1963: 50 \]
\[ I&L 1965: 85 \]

As before, the rising part of the movement must probably be thought of as having started earlier than the first point plotted here. The first vowel in "" lasts 143 msec, the peak occurs at 68 msec; the first vowel in "" lasts 223 msec, the peak occurs at 56 msec. Thus, the peaks are approximately in the same place from the beginning of the vowel (it tends to be somewhat earlier in ""), the length being added in the falling part of the movement. It follows that it is somewhat confusing to express the location of the peak as a percentage of the duration of the vowel, although this may be useful for other reasons: if one does so, it looks as if the peaks in "" and "" are in different places also if in reality they are (almost) in the
same place. In an experiment with synthesized tones (LSI 1973(b): 184) LSI placed the peak of falling accents at 50 msec from the beginning of the vowel, irrespective of the duration of the vowel, which was varied.

As to the falling part of the movement in the examples, the relevant point to note here is that the fall does not become less steep in the accented syllable when the latter has a long vowel (cf. e.g. LSI 1967). In effect, at the end of the long vowel in the example, pitch is already -6.0 ST under the peak, whereas at the end of the short vowel it has fallen only -1.4 ST, less time being available in the latter case. The level already reached within the accented syllable in the long vowel case is, in the short vowel case, attained somewhere during the postaccentual syllable: at the beginning of the postaccentual vowel, pitch is -4.9 ST under the peak, at the end of this vowel it is -8.6 ST under the peak. At the end of the postaccentual vowel after the long accent, pitch is -8.7 ST under the peak; hence, the fall becomes less steep in the postaccentual vowel here, the size of the fall being (almost) equal to that in the short vowel case, despite the longer total duration.

As to the other informants represented in I&L's data, it is interesting to note that a speaker may have a relatively steep fall in the intervocalic consonant(s) with short falling accents, but a more gradual fall during the vowel in the accented syllable with long falling accents:

<table>
<thead>
<tr>
<th></th>
<th>type ''</th>
<th>type ''</th>
</tr>
</thead>
<tbody>
<tr>
<td>informant D8</td>
<td>L&amp;I 1963: 49</td>
<td>L&amp;I 1963: 49</td>
</tr>
<tr>
<td>LSI 1965: 84</td>
<td>LSI 1965: 84</td>
<td></td>
</tr>
</tbody>
</table>

On the whole, the pronunciation of the long falling accent seems to be more stable than that of the short falling accent in I&L's data.

Turning now to the long rising accent, here length is added before the peak. If, as in I.'s speech, the entire vowel of the accented syllable is used for gradually going to that peak, we would expect that the rising movement becomes more gradual with a long instead of a short vowel, more time being available. This effect is less than it might be in theory, because the peak tends to be retracted somewhat in the long vowel case. Thus, the highest point of the configuration is, in I.'s speech, principally in the postaccentual syllable with short rising accents (with the exception of short rising accents in non-initial penultimate syllables before a short ultimate - see 3.6 below), whereas it is often at the end of the vowel in the accented syllable with long rising accents. The difference is not spectacular, however, because pitch remains at approximately the same level until the beginning of the postaccentual vowel anyway (see also L&I 1986: 46):

<table>
<thead>
<tr>
<th></th>
<th>type ''</th>
<th>type ''</th>
</tr>
</thead>
<tbody>
<tr>
<td>informant I</td>
<td>L&amp;I 1963: 19</td>
<td>L&amp;I 1963: 19</td>
</tr>
<tr>
<td>LSI 1963: 44</td>
<td>LSI 1963: 55</td>
<td></td>
</tr>
</tbody>
</table>

Just as in short rising accents, the rising part of long rising accents may be concentrated at the end of the accented syllable, or two essentially level syllables may follow each other.

3.5 THE NOTION "PEAK" (VRHUNAC)

There has been a lot of confusion about the meaning of the word "peak", different authors applying the word in different ways, and the same authors using it in more than one sense. The senses relevant to the present discussion are: "highest point reached during the vowel of the accented syllable" (henceforth: peak₁), "highest point of the entire configuration" (henceforth: peak₂), and "highest point reached during any vowel" (henceforth: peak₃). The discussion is about whether the location of the peak is relevant.

As we saw in 3.3, the pitch movement in the vowel of the accented syllable in `-` and `-`, and hence the location of peak₁, varies
considerably. Therefore, ISL were undoubtedly correct in concluding that the location of peak₁ is irrelevant to the distinction between, at least, the short accents. Yet, Purcell (1973) found that "the peak of the fundamental frequency within the accented vowel occurs closer to the start of the vowel for a falling accent than for a rising accent" (1973: 159). The main cause of this discrepancy is very simple: in Purcell's data, peak₁ in most cases equals peak₂, i.e. the highest point of the configuration is in the accented syllable; in ISL's data, rising accents mostly have the highest point of the configuration in the postaccentual syllable, i.e. peak₁ ≠ peak₂ (Purcell 1973: 209; L&I 1986: 155). Thus, also in ISL's data, the peak₂ of rising accents is late, namely still later than in Purcell's data. I would suggest that this peak₂ is the only one whose location can be relevant, whether it occurs in the accented or in the post-accentual syllable. Thus, instead of applying a notion such as "tonal center" (L&I 1986: 68-72, 256-257) to a fixed syllable (the accented syllable with falling accents and the postaccentual syllable with rising accents), and then saying that falling accents always have the peak₂ in the tonal center, and that rising accents mostly have the peak₂ in the tonal center, although they can have the peak₂ in the syllable before the tonal center, it is much simpler, in my view, to forget about peak₁, now that it has been shown to be irrelevant, and to speak only about peak₂ (cf. L&I 1986: 169). This implies, of course, that only entire configurations have to be compared with each other, not the pitch movements in every individual vowel.

The location of peak₂ in rising accents is clearly subject to dialectal differentiation: ISL's informants were mainly from Vojvodina (northeast), Purcell's informants were from (Bosnia-)Hercegovina ((south)west)(see e.g. Belić 1926-1927: 230). However, the observation that speakers from Vojvodina have the highest point of the configuration in rising accents mostly in the postaccentual syllable, whereas speakers from (Bosnia-)Hercegovina tend to place it late in the accented syllable, does not close the issue. As we saw in 3.3, the peak₂ of short falling accents is sometimes late in the accented syllable in ISL's data. So it seems that some configurations are realizations of falling accents for northeastern speakers and of rising accents for (south)western speakers. The same is suggested by the perceptual tests reported in L&I (1986: 103-105, 111-120): it appeared that listeners from the northeast recognized the (synthesized) accents mainly on the basis of the F₀ relationship between accented and postaccentual syllables, whereas listeners from the southwest reacted to the F₀ movement in the accented syllable; in effect, the answers were different in the crucial cases. This result is quite spectacular, because it is so disastrous: it suggests that listeners from different areas use entirely different cues for recognizing the accents, raising the possibility that they often misunderstand each other; yet, this is not the case (cf. L&I 1986: 169-171, 254-255). I shall return to the issue in 3.7 below. I think that the situation is less disastrous than it seems to be: although potentially ambiguous realizations probably exist, a slight modification of L&I's test stimuli might well reduce their number greatly.

A further issue about peak location has been introduced by Gvozdanović, who, in publications from 1972 onwards (Santen 1972), bases her analyses on the location of peak₁, i.e. the highest point reached during every subsequent vowel. The author finds the location of this peak₁ in ISL's data, which, indeed, specify the highest point reached during every subsequent vowel. ISL unfortunately call this highest point also "peak", but the quantitative data, as well as the discussion, make it clear that this "peak" is not a peak in the sense of "pitch obstruction standing out from its surroundings". Thus, L&I (1963: 21) say, for example, that the postaccentual syllable after falling accents does not contain a separate fundamental frequency peak (although it has, of course, a highest point). At any rate, Gvozdanović defines: 1. a falling type of fundamental frequency: peak₁ in the first 25% of the duration of a vowel; 2. a non-falling non-rising type of fundamental frequency: peak₂ between 25% and 75% of the duration of a vowel; and 3. a rising type of fundamental frequency: peak₃ in the last 25% of the duration of a vowel (e.g. Gvozdanović 1980: 35). Since these peaks apply to both accented and unaccented syllables, they include peaks₁ and peaks₂, as well as highest points in other vowels. Thus, in the following picture of "", repeated here from 3.2 above, peaks₁ are the points indicated by the arrows:
The bubble in the first vowel here is probably automatic; it can probably be "stylized away" in a perceptually adequate representation.

Next, it is not easy to translate the falling and rising types of accent into the types of pitch movement cited above without arriving at total chaos. More specifically, as we saw in 3.3 above, the location of peak is varies considerably, so that I&L concluded that it is irrelevant. Gvozdanović uses for her analysis only a part of I&L’s data, a part suggesting that the variation is less considerable than it in fact is (for one case see L&L 1986: 161-162). Then, the author treats all peaks on a par, as if they were peaks in the sense of "pitch obtrusions standing out from their surroundings". Thus, in Gvozdanović’s schematic representation (1980: 36) the picture of given above would appear as follows:

\[ \text{\begin{center}
\begin{tabular}{c}
\hline
\text{1} & \text{1} \\
\hline
\end{tabular}
\end{center}} \]

This type of representation is (ibid.) a modified reproduction from Šilić and Rosandić (1976: 32-35). The pictures referred to can hardly be called adequate, for example:

\[ \text{\begin{center}
\begin{tabular}{c}
\hline
\text{[n]} & \text{[a]} & \text{[p̩]} & \text{[n]} \\
\hline
\end{tabular}
\end{center}} \]

(Silić and Rosandić 1976: 35) This is a highly infelicitous popular representation of the finding (probably I&L’s) that with falling accents, the highest point in the postaccentual syllable is lower than the highest point in the accented syllable, whereas with rising accents, the postaccentual syllable is as high as, or higher than, the accented syllable. The representation is especially infelicitous in a schoolbook (Šilić and Rosandić 1976 is intended for the first class of secondary schools), because it is too abstract. Stylized representations of the pitch movements, along the lines of the pictures used in the foregoing here, would be much clearer. At any rate, Gvozdanović has possibly understood the vertical bars in the postaccentual syllables as stylized representations of pitch movements (whereas they can at most be bars in the sense of a histogram); she places them slantwise ("since fundamental frequency is a function of time" - Gvozdanović 1980: 36) and adds preaccentual syllables. When we look at the resulting pictures, e.g. the one derived for above, we see that it becomes difficult to understand why, in the example, the second syllable is perceived as prominent, because the picture can at most be a stylized representation of . From a listening test it appears that the problem is solved differently in different dialects. In application to the example, Gvozdanović’s (1980: 97) conclusion is that people from the (south)west (in the author’s classification, which differs from the usual one - L&L 1986: 166) perceive the second syllable as accented because they have a prosodic word boundary between the first and the second syllables, and, given the fact that there is no syllable which has the peak in the last 25% of the duration of the vowel, the second syllable (the first one of the new prosodic word) has the peak in the area of 25%-75% of the duration of the vowel. People from the northeast, in contrast, have no prosodic word boundary between the first and the second syllables (and would therefore hear the first syllable as accented on the basis of its peak location if they listened from the beginning onwards), but they nevertheless hear the second syllable as accented because it is the first syllable from the end which does not have the peak in the first 25% of the duration of the vowel.

This formulation would not be incorrect, although unnecessarily complicated, if the postulated dialectal difference as to prosodic word boundaries existed (Gvozdanović 1984-1985: 177-178; L&L 1986:
156), if prosodic word boundaries were the opposite of what one would be inclined to call word boundaries at first sight (see 2.3), if all data were identical to those in the author's selection, and if small pitch bubbles overshadowed by their surroundings, and the absence of pitch obtrusions, could be treated on a par with perceptually relevant pitch obtrusions. All this not being the case, we can safely, without any listening test, discard the notion "prosodic word", as well as the types of pitch movement cited above, i.e. drop the idea that \( \text{peak}_3 \) is a useful notion.

I am not sure whether Gvozdanović would now (1984-1985) subscribe to this proposal. On the one hand, she has changed the hierarchy in her description, to the effect that preaccentual syllables having the highest point between 25%-75% of the duration of the vowel (see the bubble in the first syllable of our example) are now separated terminologically from accented syllables having the peak in the same place. Also, it seems that accented syllables can now have the peak in the first 25% of the duration of the vowel (just like postaccentual syllables) (as I&I 1985: 162 remark, such examples are not uncommon in their data). Furthermore, the author seems to agree that the formerly postulated dialectal differentiation as to the location of prosodic word boundaries does not exist (see above). And finally, she now summarizes her findings concerning trisyllabic words as follows (op.cit. 177, my translation): "In other words, only the syllable with a clear \( F_0 \) peak can be perceived as accented", which is true if one does not regard the peak in the postaccentual syllable after rising accents as a clear \( F_0 \) peak. But on the other hand, she retains the notion "prosodic word": a peak in between 25%-75% of the duration of a vowel in an unaccented syllable signals its beginning, and a peak in the first 25% of the duration of a vowel in an unaccented syllable signals its end (ibid.). In other words, one must know which syllable is accented in order to determine whether a given type of pitch movement signals a prosodic word boundary. Also, the author summarizes her main 1980 hypothesis, viz. that accent is predictable from tone (type of pitch movement) and prosodic word boundaries, without changing it (ibid.). It seems to me that this is a bit odd, because we now need the accent in order to know the prosodic word boundaries. At any rate, if one wishes to retain the Jakobsonian idea that either tone or accent can be eliminated from the analysis, one must, in my view, formulate things in such a way that the thing predicted does not appear in the predicting parameters. Also, in my view, at least all possibilities represented in I&I's data should be included. Besides peaks, in the first 25%, in the area from 25%-75%, and in the last 25% of the duration of the vowel in an accented syllable, there are also preaccentual syllables having the highest point of the bubble not between 25%-75% of the duration of the vowel, and postaccentual syllables having the highest point not in the first 25% of the duration of the vowel. I would suggest that it is easier to discard the Jakobsonian idea: it is wrong anyway, even if it can be adequately formulated (see 1.4).

Summarizing the present section, I see no reason to talk about peaks other than peaks, henceforth: peaks.

3.6 MINIMAL REALIZATIONS

Having eliminated all peaks other than the highest point of the configurations called "falling" and "rising" accents, there are still two competing ways of describing the configurations: in terms of peak location or in terms of pitch relationships. Although the two will often amount to the same, I think I&I's option, viz. pitch relationships, remains the best one, although my motivation for this choice is somewhat different from I&I's: I would say it is needed because the movements in the configurations have a variable slope. Also, in this section and in the next one, I shall suggest that I&I's description in terms of pitch relationships between accented and postaccentual syllables is not the only option, and that a slightly different formulation might be better, especially when "(south)western" pronunciations are taken into account. Since the data are defective for my purpose, I can only give the idea, not the details. Further, we will be talking only about the "neutral" realizations under discussion in the present chapter.

I think it is useful to start by summarizing I&I's findings by introducing a "forbidden area". In the schematic rise&fall and fall&rise used in Chapter One above, this area is defined by the distance between the two points indicated in the following picture:
I have now added some declination, cf. L&I (1986: 199-204) (but the scale should probably be logarithmic, so that the relevant lines are probably more or less parallel - see Cohen et al. 1982). Let me call the distance between the two points indicated "the belt". In general terms we can say that (in I&L's data) in the first part of the postaccentual syllable rising accents must be above the belt, while falling accents must be under it; the belt itself is forbidden area.

The most important remaining questions are then, in my view:

1. With respect to which other point(s) of the configuration must the belt be defined?; and

2. How broad must the belt be, i.e. what is the minimal difference between falling and rising accents?

I&L's discussion suggests that the belt be defined with respect to the accented syllable, the upper boundary of the belt being located approximately at the level of the (highest point reached in the) accented syllable, and the lower boundary considerably under it. I think this choice creates some problems. First, if the boundaries are defined with respect to the accented syllable, an entirely different description must be given when "(south)western" pronunciations enter the discussion (see the next section). Further, for falling accents, if the lower boundary of the belt must be located at a certain distance under (the highest point of) the accented syllable, it becomes difficult to make allowance for obtrusions of various size. But it is doubtful whether we can separately compute the size of the falling part of an obtrusion, independently of the size of the rising part preceding it. In contrast, it is known that our ear is sensitive to whether or not pitch resumes a low reference line (see e.g. Cohen et al. 1982: 268). Therefore, it seems to me that it is better to give a description directly in terms of the wider perspective offered by L&I (1986: 199-204) when they discuss the intonation of entire sentences. To this end, we may draw the low line belonging to the picture given above and locate the belt with respect to this line:

Theoretically possible falling accents are then, for example:

For monosyllabic words (which can only have a falling accent - see 1.3) the belt can safely be placed at the end of the accented syllable (see also the next section).

Next, we may make a guess at the minimal breadth of the belt (cf. question 2 above). I shall do so by discussing some less clear-cut realizations of the short accents, in the examples spoken by I., than have been discussed in the foregoing. We would need the melodic contexts of the words under discussion if we were to be precise (cf. 3.2 above); since these contexts are not available for most examples of falling accents, I can only give an approximation.

"Minimal" realizations of short rising accents in I.'s speech are cases which have the highest point of the configuration in the accented syllable, so that pitch has already started to fall at the beginning of the postaccentual syllable. These are, in the available examples (see also L&I 1986: 46), non-initial short rising accents in penultimate syllables before an ultimate with a short vowel, i.e.
types "" and "". (These cases are especially interesting because
there are dialects which have no tone opposition in penultimate
syllables before a short vowel; there seems to be a conflict between
the wish to be low at the end of the word and the wish to have the
peak of rising accents late.) The pictures representing the (average)
quantitative data are the following:

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;&quot;</td>
<td>I</td>
<td>I&amp;L 1963: 44</td>
</tr>
<tr>
<td>&quot;&quot;</td>
<td>I</td>
<td>I&amp;L 1963: 45</td>
</tr>
</tbody>
</table>

From the last point in the preaccentual syllable to the highest point
pitch rises 3.1 ST in "", and 3.0 ST in "". The first point in the
postaccentual vowel is -0.4 ST under the peak in "", and -0.6 ST in
"". At the end of the postaccentual vowel (125 msec and 113 msec
later, respectively), pitch is -5.3 ST under the peak in "", and
-4.4 ST in "". So we can guess that at the end of the postaccentual
vowel pitch has returned to the level from which the obtrusion
started. (The fact that the falling part of the obtrusion is larger
than the rising part can probably be ascribed to declination.) We see
here that the retraction of the peak into the accented syllable is
accompanied by a return to the level of the preaccentual syllable
within the postaccentual syllable. This fact also makes the cases
"minimal". When more postaccentual syllables with short vowels are
present, pitch returns to the level of the preaccentual syllable
(postaccentual syllable) in the post-postaccentual syllable; when the
postaccentual syllable has a long vowel, pitch returns to the level
of the preaccentual syllable in the postaccentual syllable, but then
the peak is also in the postaccentual syllable (more time being
available) (as far as can be guessed on the basis of the available
examples (I&L 1963: 44-45), of course).

Further, some examples of short rising accents are "minimal" in
the sense that the highest point of the configuration (now in the
postaccentual syllable) is less than 2 ST above the last point in the
preaccentual syllable (I take only the case where the latter is
available): "" (+1.0 ST), "" (+1.3 ST), "" (+1.5 ST). But I
think we must reckon with the possibility that the (gradual) rise
starts still earlier. Thus, in "" the highest point of the
configuration is 2.3 ST above the average level of the pre-pre-
accentual syllable; for "" this figure is 3.3 ST (for "" the
figure cannot be given).

On the whole, my best guess is that in I.'s speech the upper
boundary of the belt (below which rising accents may not be in the
first part of the postaccentual syllable) is at least 2 ST above the
level from which the obtrusion starts. If the highest point of the
configuration is in the accented syllable in I.'s speech, this
highest point is not more than 0.6 ST above the highest point in the
postaccentual syllable (as we will see in the next section, this
issue is relevant for distinguishing between different dialects).

Next, we would like to compare these figures with those of
"minimal" realizations of the short falling accent, but for the
latter the preceding melodic context is not available; realizations
for which the preaccentual syllable is available (combinations with a
proclitic word) are not "minimal" realizations. So we must make a
guess on the basis of a combination of data.

First, "non-minimal" realizations of the short falling accent
where data are available for the preaccentual syllable (and for the
postaccentual syllable):

<table>
<thead>
<tr>
<th>Type</th>
<th>Informant</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;</td>
<td>I</td>
<td>I&amp;L 1963: 39</td>
</tr>
</tbody>
</table>

In "" (repeated here from earlier sections) pitch rises 4.7 ST from
the last point in the preaccentual vowel to the peak. It falls -4.6
ST to the first point in the postaccentual vowel; at the end of this
vowel (99 msec later) pitch is -8.4 ST under the peak. In "" the
rise is 6.2 ST from the last point in the preaccentual syllable
to the peak. The fall is -4.7 ST to the first point in the postaccentual
syllable and -10.2 ST to the end of the postaccentual vowel (120 msec
later). Thus, in both cases pitch returns to the level of the
preaccentual syllable (plus declination) in the first part of the postaccentual vowel.

Next, we may consider some realizations of the short falling accent which are "minimal" in the sense that the interval between the highest point of the configuration and the first point in the postaccentual vowel is not very large, and/or in the sense that in the beginning of the postaccentual vowel pitch is not yet considerably under the starting point of the vowel in the accented syllable (and hence certainly above the earlier point from which the obtrusion has probably started - the latter being unknown). In the available examples (see also L&I 1986: 46) these are the following:

- **type** ---
  - informant I
  - ISL 1963: 38

(I omit the type ---, for which the data are incomplete.) In --- pitch rises 1.4 ST from the beginning of the first vowel to the peak. It falls -1.7 ST to the first point in the postaccentual vowel and -4.7 ST to the end of this vowel (105 msec later). Although the level of the preaccentual syllable (unknown) is probably resumed during the postaccentual vowel here, the part of the fall completed before the postaccentual vowel is small as compared with other examples spoken by I. (where this part tends to be at least ca. 3 ST, except for --- (-2.3 ST))(see also L&I 1986: 46). Type --- is deviating in a number of respects: the rise is relatively large (2.4 ST from the beginning of the first vowel); the fall is relatively small (-1.4 ST to the beginning of the postaccentual vowel, -3.6 ST to the end of this vowel); the peak is relatively late (at 83 msec from the beginning of the vowel in the accented syllable (at 73.5% of the duration of this vowel)). Since all potentially problematic realizations of the short falling accent in the examples spoken by I. have a long postaccentual vowel somewhere (but not all examples with a long postaccentual vowel are problematic), one is inclined to think that it is not entirely natural for I. to pronounce postaccentual length (cf. L&I 1963: 30; L&I 1986: 35). At any rate, I would guess that the lower boundary of the belt is approximately at the level from which the obtrusion starts (plus declination), and that pitch resumes this level in principle in the first part of the postaccentual vowel; but --- is an exception in the sense that the first two syllables are close to the "minimal" realizations of the short rising accent given above (which are, however, clearly distinct from ---, i.e. from a short falling accent before a short ultimate). On the basis of the available data alone, however, we can only say that with short falling accents in I.'s speech (frame sentence environment), the highest point of the postaccentual vowel is at least -1.4 ST under the peak (mostly ca. -3 ST).

On the whole, then, I would guess that in the examples spoken by I. there is a belt in the postaccentual syllable of at least ca. 2 ST. For other informants represented in I&L's data, no reasonable guess at the breadth of the belt can be made, because information about preaccentual syllables is absent also for rising accents. But it is clear that the belt tends to be broader than in I.'s speech (making the difference between the accents clearer perceptually). To give an extreme example, type "" spoken by E10 has the highest point of the postaccentual syllable -6.3 ST under the peak; type "" spoken by the same informant has the peak in the postaccentual syllable; this peak is 4.0 ST above the highest point reached during the accented syllable (pitch does not return to low before the end of the word)(L&I 1963: 50; I&L 1965: 85). It seems to me that relatively steep realizations of the accents have a broader belt than relatively gradual realizations (see also the end of the next section).

A special case is E14, whose realizations of "", "", and "", are not very convincing if we try to describe them in terms of the distance between the highest point of the accented syllable and the highest and last points of the postaccentual syllable (preceding context is not available). Thus, in E14's "", pitch rises 2.8 ST from the beginning of the first vowel; it falls only -1.9 ST to the beginning of the postaccentual vowel, and -4.5 ST to the end of this vowel. In his "" pitch rises 3.6 ST from the beginning of the vowel; it falls -0.5 ST to the beginning of the postaccentual vowel, and -3.1 ST to the end of this vowel. Moreover, the peak is late in the
accented syllable in both cases (so that there is no compensation for a narrow belt). Type "" is probably fine as compared with "" pronounced by the same speaker, but minimal otherwise: in "" there is a rise of 4.1 ST from the beginning of the first vowel, a fall of -1.2 ST to the first point in the postaccentual vowel, and of -2.1 ST to the end of this vowel (plus a late peak). But in "" the peak is in the postaccentual syllable (5.2 ST above the beginning of the first vowel), and pitch falls -0.8 ST from the peak to the end of the postaccentual vowel. It is possible that this speaker is losing the tone contrast in dimyllabic words with short vowels. Alternatively, it is relevant that the pronunciation of E14 "shows clearly his dialectal background (western branch of the central dialect), which has been only slightly altered by the Northeastern environment" (L&I 1963: 38; quantitative data L&I 1963: 54; I&L 1965: 87).

Before we turn to the problem of dialectal differentiation, it must perhaps be made explicit that I assume that the belt is not a line, i.e. that the transition between rising and falling accents is not categorical. I am not sure whether I&L would agree with this assumption. On the one hand, their experiments reported in L&I 1972 (see also L&I 1986: 93-105), although not devoted precisely to a belt in my sense, clearly suggest that there is a "random response" area between rising and falling accents. But on the other hand they mention (e.g. L&I 1986: 38) that their measurements had an accuracy of ±1 Hz, and that, according to L 1970, this is close to the just noticeable difference for the perception of fundamental frequency in spoken language. Isačenko & Schärdich (1970: 17-18) and Rossi & Chafcouloff (1972) suggest that L is somewhat too optimistic here about what people can hear in normal communicative situations (for discussion see 't Hart 1981). At any rate, the minimal difference between rising and falling accents has still to be determined, in my view.

3.7 NORTHEAST VERSUS (SOUTH)WEST?

We saw in 3.5 above that speakers from (south)western areas tend to pronounce rising accents in such a way that the highest point of the configuration is in the accented syllable rather than in the postaccentual syllable. In itself, this fact is unproblematic; after all, there are also realizations of rising accents spoken by, e.g., I., which have the peak in the accented syllable. What makes the issue problematic is that the highest point of the postaccentual syllable can already be considerably under the peak. Purcell's (1973) speaker R.G., for example, whose native status is not questioned by L&I (1986: 155-156), has the following distances between the highest points of the accented and of the postaccentual syllables in frame sentence environment (Purcell 1973: 243): "" : -5.9 ST (-7.3 ST to the end of the postaccentual vowel), "" : -6.0 ST (-7.9 ST), "" : -4.4 ST (-6.3 ST), "" : -1.7 ST (-5.1 ST), "" : -5.0 ST (-6.4 ST), "" : -3.7 ST (-6.4 ST), "" : -3.7 ST (-5.7 ST), "" : -1.2 ST (-4.4 ST). This is, of course, incompatible with the idea that with rising accents, the postaccentual syllable starts approximately as high as, or higher than, the highest point reached in the accented syllable.

L&I's tests devoted to this issue (1986: 103-105, 111-120) suggest that listeners from the northeast and listeners from the southwest base their identification of short accents on different cues, viz. the $F_0$ relationship between accented and postaccentual syllables (northeast) and the movement in the accented syllable (southwest). I think it would be preferable to try and find a formulation which describes the dialects in the same terms; not all possibilities for doing so have been exhausted, in my view.

Since the peak of rising accents has clearly been retracted into the accented syllable in southwestern pronunciations, a possibility which immediately suggests itself is that the belt, in the sense of the preceding section, is in principle placed at the end of the accented syllable here, rather than in the postaccentual syllable. Thus, the figures for Purcell's speaker R.G. become much better if we take the distance from the peak to the end of the accented syllable, probably especially if we take into account the level from which the obstruction starts. (Here, if anywhere, we would need a preceding context, but, unfortunately, Purcell does not give it.) Thus, in R.G.'s "" the end of the vowel in the accented syllable is -4.4 ST under the peak and -3.8 ST under the beginning of the vowel (an earlier point is not available). For the other types the figures are the following: "" : -3.9 ST (-3.5 ST), "" : -1.3 ST (+0.9 ST), "" : -1.3 ST (+0.1 ST), "" : -4.2 ST (-3.4 ST), "" : -1.0
ST (+0.6 ST), "": -0.7 ST (+1.4 ST), "": -0.9 ST (+0.7 ST). Here, the difference between "" and "" remains questionable, of course (see below).

Next, the question arises of whether a belt formulation does not amount to the same as a formulation in terms of peak location: if, with falling accents, there must be an obtrusion such, that at the end of the accented syllable, pitch has returned to the level of the preaccentual syllable (presumably), the peak can hardly be elsewhere than in the beginning of the accented syllable; and if, with rising accents, there must be an obtrusion such, that at the end of the accented syllable, pitch is still at a certain distance above the level of the preaccentual syllable, the peak can hardly be elsewhere than at the end of the accented syllable. And why not say that for northeastern pronunciations both peaks are in principle later, say in the middle of the accented syllable for short falling accents and in the postaccentual syllable for short rising accents? Although this is, of course, the obvious relation between belt and peak location, I think the belt formulation is better, because there are examples showing that the highest point can be elsewhere than in the most obvious place, provided that the slope of the movements involved is adjusted to the belt. For example, Purcell's R.G. has a relatively early highest point in type "" (at 41.3% of the duration of the vowel) and a relatively late peak in "" (at 66.7% of the duration of the vowel). In "", as compared with "", which has the highest point at 35.7% of the duration of the vowel, the speaker succeeds in making a difference despite the comparable peak location by delaying the fall until the end of the vowel in "" but not in "":

Thus, if we take the entire fall, from the highest point to the end of the postaccentual vowel (-5.7 ST in "", -6.4 ST in ""), we can say that the perceptual effect of this different distribution is sufficiently close to a late and early peak, respectively, for the two movements to be realizations of a rising and falling accent, respectively.

As to R.G.'s "" (peak at 66.7%), here the speaker does not succeed in making the movement different from "" (peak at 72.2%) before the end of the accented syllable. But slightly later, the same difference can be observed as in "" versus "":

In "" the fall from the highest point to the end of the postaccentual vowel is -4.4 ST; 27.3% of this fall (-1.2 ST) is completed before the postaccentual vowel. In "" the fall from the peak to the end of the postaccentual vowel is -6.4 ST; but now 57.8% of this fall (-3.7 ST) is completed before the postaccentual vowel. Thus, again we see that the almost identical peak location is compensated for by a different distribution of the fall, i.e. by adjusting the slope of the movements.

To give an example from I&L's data, consider the following pair, both with the highest point late in the accented syllable:

This pair can be described in terms of distances between the highest point of the configuration and the highest point of the postaccentual syllable: as we saw in 3.6, this distance is -0.4 ST in I.'s ""; it is -7.0 ST in D8's "". But we can also say that at the beginning of
the postaccentual vowel, I., has completed only 7.5% of the total fall (-5.3 ST), whereas D8 has covered already 75.3% of the total fall (-9.3 ST) at this point. The almost identical peak location (at 86.9% of the duration of the vowel in I.'s case, at 82.5% in D8's case) is obviously irrelevant.

On the whole, then, although there is an obvious correlation between peak position and belt, I would suggest abstracting away from peak position in both "northwestern" and "southeastern" pronunciations. Further, I would expect that potentially ambiguous realizations can be made unambiguous in the following way. Consider the following three configurations with a peak placed 1. early in the accented syllable; 2. late in the accented syllable and 3. in the postaccentual syllable:

Movement 1 will probably be recognized as a falling accent by all listeners; movement 3 will probably be recognized as a rising accent by all listeners. But movement 2 is potentially ambiguous. Listeners from the northeast, who have the highest point of rising accents "prototypically" in the postaccentual syllable (a3), will be inclined to hear movement 2 as a falling accent, because at the beginning of the postaccentual syllable, pitch is already falling, so that it will soon reach the low boundary of the belt (b3). In order to turn movement 2 into a clearly rising accent, we must make the fall gradual, at least in its initial part, so that it, in effect, takes the line of movement 3:

Listeners from the southwest, in contrast, have the peak of rising accents "prototypically" late in the accented syllable (a2). They will hence be inclined to hear movement 2 as a rising accent. In order to turn it into a falling accent, we must make a steep fall after the peak, so that it is clear that at the beginning of the postaccentual syllable, pitch has returned to low, i.e. that it has taken the line of movement 1 despite the late peak:

Turning now to L&I's (1986: 111-120) test presumably showing that listeners from the northeast attach the greatest importance to the F0 relationship between accented and postaccentual syllables, whereas listeners from the southwest react to the movement in the accented syllable, I would suggest that two additional types of stimulus are needed before such enormous differences can be assumed to exist. My first modification amounts to replacing a vowel-by-vowel way of thinking by a configurational point of view. Consider the following stimulus:

This stimulus is intended to combine a relatively early highest point in the accented syllable with a relatively high postaccentual syllable. I think it is possible that southwestern listeners, who identified the combination as a realization of a falling accent, may have simply disregarded the postaccentual syllable, because the configuration probably sounds unnatural to them (having, in effect, two peaks). I would suggest that it is more natural, retaining the same idea, to connect the highest point of the accented syllable with...
with, say, the first point of the postaccentual syllable:

L&I 1986: 112
stimulus 6 (modified)

This modification might also evoke "rising" responses from southwestern listeners.

The opposite stimulus type, viz. a relatively late peak in the accented syllable plus a relatively low postaccentual syllable, also gave rise to a different response pattern from northwestern and southwestern listeners. For example:

L&I 1986: 112
stimulus 5

The highest point of the postaccentual syllable is -4.0 ST under the peak and -2.1 ST under the starting point of the rise. So north­

eastern listeners say that the configuration is a realization of a falling accent. Southwestern listeners say that it is a realization of a rising accent because, according to L&I, the peak is late. But also, I would suggest, because the stimulus lacks a cue enabling such listeners to discard the idea that the configuration is a rising accent. Such a cue could be introduced by changing the movement in the postaccentual syllable, retaining the same end-point, in the following way:

L&I 1986: 112
stimulus 5 (modified)

As has been explained above, this modification has the effect of placing more of the total fall before the postaccentual syllable. I think this may counteract the late position of the peak. As it is, unmodified, the gradual fall in the postaccentual syllable can probably be evaluated as a movement bringing pitch back to low, whereas for southwestern listeners, the perceptual effect should be that pitch has already returned to low at the beginning of the postaccentual syllable.

Along these lines a uniform description can possibly be arrived

at.

Finally, I would like to discuss a case which turned out to be ambiguous when it was used in an experiment:

example: nôsi
informant: I
L&I 1986: 106

The word had been spoken and recognized as a realization of a short falling accent by I. independently (L&I 1986: 108-110); yet, it turned out to be too much like a rising accent for others. According to the belt criterion developed in 3.6 for I.'s speech, the example is indeed a falling accent: the highest point of the postaccentual syllable is -4.2 ST under the peak and -0.9 ST under the starting point of the rise in the vowel (the consonant is not available). But, in comparison with the relatively large rise (L&I 1986: 108), the fall in the accented syllable is very gradual: after the rise of 3.3 ST (plus consonant) pitch falls 2.0 ST in the vowel of the accented syllable. In effect, the fall, which was probably intended to be a fall bringing pitch gradually back to the low level, can possibly be interpreted as a high level fragment (with declination being responsible for the downward drift), so that the real level-changing fall only starts at the end of the accented syllable. Since the fall is gradual also in the postaccentual vowel, the example resembles a rising accent with a relatively early peak. Two types of modification might eliminate this effect. First, we may connect the highest point in the accented syllable directly with the end of the postaccentual syllable:
Since the fall in the accented syllable is somewhat less gradual in the modification than in the original, it is probably clearer that the fall is indeed the fall bringing pitch back to the low level. Alternatively, changing the postaccentual syllable, we can possibly turn the example into a steep type of realization of a falling accent by lowering the beginning of the postaccentual syllable:

Example: nősi
Modification: + informant I
L&I 1986: 106

In this modification, the level-changing fall starts at the end of the accented syllable. In that case the (steep) fall in the consonant should be larger than 3.3 ST (the size of the rise): at least the 2 ST must be added which have been "lost" during the accented syllable. This is why I think that the lower boundary of the belt is farther below the starting point of the obtrusion with steep realizations of the falling accent than with gradual realizations. (The belt described for I. in 3.6 pertains to "gradual" realizations of the falling accent.) As it is, unmodified, the example has a somewhat hybrid character. I hope that nobody will try to explain this fact by observing that the peak is exactly in the middle of the accented syllable.

4.0 INTRODUCTION

More than 25 years ago, De Bray (1959-1960: 381) expressed the opinion that "while Serbo-Croatian possesses greater varieties of word accent than, say, Russian and English, its sentence intonation is much simpler". It must be doubted, alas, whether this is correct. It is true that far more is known about intonation in, say, Russian or English. For Serbocroatian there does not seem to be the sort of basic fund of knowledge comparable to, say, Bryzgunova's (1977) inventory of Intonation Constructions for Russian, which could be taught to foreigners learning the language and which could be referred to by a linguist in such a way that everybody more or less understands what he is talking about. The neglect of the subject is obviously a consequence of the fact that much linguistic effort is needed for the description of the great dialectal diversity in word tone systems alone; also, of course, there is no mass teaching of the language to foreigners which could have given rise to a didactic tradition. The discussion in this chapter is, therefore, mainly intended to convey the idea that further research is indispensable.

4.1 PRELIMINARY REMARKS

L&I (1986: 179-237) present a survey of research on sentence intonation. The main contribution is that by the authors themselves (summarized in L&I 1986: 179-220). They note (1986: 179) that "a substantial amount of research in the intonation of tone and accent languages has recently been carried out [...] if we were to start our investigation now, it would benefit from the results achieved in these studies".

In retrospect it can be said that two main factors make the reading of the data fairly complicated:
1. As we saw in 3.2-3.3 above, a detailed analysis of the pitch movement in every subsequent vowel of a word tends to distract our attention from the pitch configuration as a whole, and omits potentially relevant information, most importantly about what happens before the vowel of an accented syllable. The issue is even more important in the case of sentence intonation, because at least two types of intonation exist, as far as I can tell, for whose descrip-
tion this information is indispensable (see 4.5-4.7 below); unfortunately, all data about intonation concern words with an accent on the first syllable. Also, it is hardly possible to determine whether an acoustic difference between words with rising and words with falling accents ensures that the tone contrast is preserved perceptually in a certain type of intonation if we do not know the pitch movement in the larger stretch into which the word enters.

2. A second point which makes the understanding of the data difficult is the classification of the material. Basically, two criteria have been used for classification: the sentence position occupied by a given word (initial, medial, final), and the type of sentence in which the word occurs (statement, wh-question, yes/no-question, exclamation, etc.). This makes it very difficult to interpret the data in the light of what is known about intonation in other languages, because the classification implies that the following issues have not clearly been separated from one another: 1. the type of pitch configuration encountered, irrespective of the sentence position in which it occurs, and irrespective of the type of sentence in which it occurs; 2. potential variations in the realization of the tones due to sentence position, with a given type of pitch configuration; and 3. the meaning of the configuration. If, for example, the first word of a sentence is spoken with a "minor continuation" type of intonation (see for this term e.g. Delattre et al 1965), and if this intonation affects the realization of word tones, the latter must be ascribed to the type of intonation chosen by the speaker, not to the fact that the word is the first word of the sentence, if the effect is absent in a sentence where the first word is spoken with, say, an "emphatic" "rhemetic" accent. It is true, of course, that the first word of a sentence is more likely to be spoken with a "minor continuation" type of intonation than the last word of a sentence, but this is because the intonation means that something is yet to follow, and this meaning is bound to be most frequently signalled if indeed something follows. As to meaning, it is no longer a common point of view among intonologists that types of pitch configuration entertain a one-to-one relationship with categories like "statement", "question", "exclamation", etc., let alone that such categories are signalled by sequences of pitch accents.

It is not quite clear to me to what extent the classification of the material has influenced IGL's average data. On the one hand, they clearly describe types of pitch configuration, such as the "reverse pattern" used in questions (see the following sections). But on the other hand, they separately average, e.g., first words in statements, first words in yes/no-questions, first words in exclamations, etc. (if the last accent is elsewhere). This leaves some doubt about whether the averaged pronunciations were indeed all pronunciations of the same type of pitch configuration, and about whether the reader would do much harm if he averaged the averages.

Because of these difficulties met while trying to interpret the data, I have not been able to determine with a certain minimal degree of reliability how many discretely different configurations with an intonational function are represented in IGL's material. But I have found no reason to doubt the following conclusions:

1. When a conflict arises between tone and intonation, it is tone, not intonation, which tends to be sacrificed (e.g. LSI 1986: 235-237).
2. There is one case where the difference between the word tones becomes clearer than in the "neutral" environment discussed in Chapter Three above, namely when the words involved are spoken with an "emphatic", "rhemetic", "non-interrogative" accent, especially in non-final position. In this case all parameters, including pitch excursions, are enlarged (e.g. IGL 1970: 246).
3. There is also one case where tone oppositions disappear altogether, namely when the words involved are spoken with the so-called "reverse pattern" (to be discussed below) (e.g. IGL 1969: 162; IGL 1970: 246).
4. The generalization for the remaining cases seems to be that tone distinctions can be retained with different intonations, and that the distinction between the long accents is more often retained than the distinction between the short accents (e.g. LSI 1986: 236).

Point 2 ("emphatic accents") will not be discussed here, as it adds no essential new information to that presented in Chapter Three; for details see IGL (1970: 237-240). As to the remaining cases, I shall not directly discuss IGL's data, for the reasons indicated above; there would be a long list of "maybe"s and "possibly"s, which
would be, I think, of little use. Rather, I shall try to approach the problem of the relation between word tone and sentence intonation in more general terms, along the following lines. In the foregoing I have tried to make it plausible that falling accents can be regarded as mode-1 accents and that rising accents can be regarded as mode-2 accents, but that the tonal mode distinction differs from known intonational mode distinctions in the property of being expressed by movements with a variable slope. At the same time, reading I&L's data concerning sentence intonation, one does not get the impression that intonation is drastically different from, say, Russian intonation, where mode distinctions of a better known type are found. It seems to me that it is not unlikely that the problem of the relation between word tone and sentence intonation amounts to the question of where two ways of expressing mode distinctions are in conflict with each other. Therefore, I shall list the most obvious cases of potential conflict below. It goes without saying that the list is only intended as an inventory of cases which are, in my view, especially interesting for further experimental research. Also, it will be clear that, if tone distinctions were to be preserved in all intonational circumstances, the most obvious way of doing so would be by reserving mode-1 types of intonation for falling accents and mode-2 types of intonation for rising accents. Since the various types of intonation are semantically different in languages without tones, the choice of this option would imply that intonational distinctions are sacrificed in order to retain tone distinctions. This is not what happens in I&L’s data; as indicated above, the main tendency is the other way round.

4.2 THE REVERSE PATTERN AND RELATED CONTOURS

The reverse pattern is reminiscent of the realization of the rising tone which has a relatively steep rise at the end of the accented syllable (see 3.3). I give two examples from I.'s speech, both with an "emphatic" accent on the given word:
The following sections discuss a number of issues concerning these configurations; a fifth type, resembling II, but not identical with it, will be given in 4.5 below.

4.3 TYPE I VERSUS TYPE III

The reverse pattern has been compared with the Russian "question intonation" (De Bray 1959-1960: 382; Leed 1968: 335-336; Nikolaeva 1973: 180-181), but there remains some confusion as to whether the two languages have the same or a different "question intonation". Type III differs from the intonation which is regarded as the normal question intonation in Russian (IK-3 in Bryzgunova's (1977) classification, type C in Odé 1986): the Russian type has the peak approximately one syllable earlier, i.e. the Russian type is type I given above. A semantic difference is correlated with the formal difference between I and III. I&L (1969: 116) note that questions without a lexical sign of interrogation (and hence with III/IV) occur mainly as an echo of a preceding utterance, if the speaker is not sure whether he has heard the preceding utterance correctly, or if he is surprised about or does not believe the content of the preceding utterance. It would be interesting to know whether this "echo" effect is also present in other sentences with III/IV, and whether it is present in both III and IV. The impression that the information is "known" in a certain sense derives, in my view, from the meaning of the initial low/falling fragment in the contours (cf. Keijser 1980: 227-233; 1983: 122-123; 1984: 26-28). It is absent in the Russian type I, because no initial low/falling fragment occurs there (the rise being early)(Keijser 1980: 220-222; 1983: 126-128).

However, the "Serbocroatian" type III also occurs in Russian, with, as far as I can tell, a meaning similar to that noted by I&L (Keijser 1983: 134, point 6; possibly cf. type G in Odé 1986). But the use of type III may be different in the two languages. (When I once tried to ask for another cup of tea, using in Serbocroatian my "Russian" type I, I was corrected by a Dutch Serbocroatist: she used type III, which could be impertinent in Russian in these circumstances.) Type IV, in contrast, is common in Russian (IK-4 in Bryzgunova's (1977) classification, type D in Odé (1986)). This raises the question of whether type I also occurs in Serbocroatian. I did not meet it in I&L's data. Leed (1968: 332) gives two examples with a rise in the accented syllable, which is the last one of the sentence; this could be a type II in the last syllable or, if type I exists, a neutralization of I and II. (The Russian example given by Leed (ibid.) is a neutralization of I and II (type F in Odé 1986)). Type III/IV (possibly neutralized) can be realized in a last syllable, at least in Russian. Further, some pictures in Mahrken (1964: 91-104) can be interpreted as representations of type I, but I do not find the evidence conclusive.

Before we proceed, it may be useful to mention that the name "question intonation" is unfortunate. For example, although type I is regarded as "the" question intonation in Russian, in reality all four types occur in questions, and type I does not only occur in questions, "question" not being an intonational category. It would not be surprising if, on closer inspection, the Serbocroatian reverse pattern occurs elsewhere than in questions alone.

4.4 TYPE III VERSUS TYPE IV

As indicated above, I&L subsume both configurations under the heading of "reverse pattern". They ascribe the falling last part of type III to sentence-final intonation. Thus, the example bambah given above (with a falling end) was the last word of a sentence, whereas zakoni (with a high end) was the first word of a sentence. In cases like the latter, the terminal fall occurs in the last syllable of the utterance (e.g. I&L 1978: 110-111). But the authors mention that speaker DI, more often than I. (it is mainly the intonation of these two persons which has been investigated), uses the high ending type in sentence-final position (I&L 1972: 126). I would suggest keeping...
the two types apart, at least in the present phase of investigation, for two main reasons. First, it would not be surprising, in the light of what is known about other languages, if the distribution of the two were different. Thus, in e.g. Russian, type IV does not only occur in questions, but is also used as a "continuation intonation" (alongside I and II); type III is not used in the latter function (as far as I know). It seems to me that Miletic (1960: 81-82) mentions type IV in this function. Secondly, if the falling last part can occur either immediately after the peak even if the word involved is not the last word of the sentence, or in the last syllable of the sentence, this is very interesting typologically, again for two reasons: 1. the corresponding sentence-final rise does not seem to exist (see 4.8 below), and 2. the difference as to the location of the fall has been associated with different languages; thus, the corresponding English contour has the fall immediately after the peak, whereas the corresponding Dutch configuration postpones the fall to the last syllable of the sentence (cf. Ladd 1983: 732, 749).

In general, it appears to be important to keep apart movements made immediately after the peak and movements made in the last syllable of an utterance, even if the difference is neutralized in sentence-final position; more specifically, the use of type III (Serbo-croatian) or type I (Russian) as a "question intonation" might be related to the impossibility of a rising movement in the last syllable (see also 4.8 below).

I&L found that tone distinctions are neutralized in III/IV, and that length distinctions are optionally neutralized (e.g. I&L 1972: 126, 129). Leed (1968: 334-335) heard a difference between long rising and long falling accents in some examples of III, the rise of the intonation contour starting earlier in the former case (in the last mora of the vowel in the accented syllable rather than in the consonant(s) between accented and postaccentual vowels). I&L's data do not confirm this observation. It would be interesting to know whether the difference between III and IV is neutralized in a last syllable. (Type IV can be realized in a last syllable, at least in Russian; the difference between I and II, in contrast, is neutralized in this position.)

The fact that the mode-2 configurations III and IV are in conflict with the tone distinction is not surprising. As Bolinger

(1978: 494) remarks, "a C accent [corresponding to III/IV here] would face definite obstacles in any tone language".

4.5 TYPE II (AND V) VERSUS TYPE IV

As indicated in the schematic pictures in 4.2, a type II intonation (mode-1) has an early, steep rise. In this respect it is comparable to the (virtual or real) rise from the preaccentual syllable to the highest point of the accented syllable which we found in 3.2 for falling accents in a "neutral" environment. In effect, a considerable part of the rise is absent in quantitative data starting from the vowel of the accented syllable. Possibly, it is the rise which can be detected in the following pictures (LI 1986: 205; note that the frequency scale is linear here):
These pictures show a large rise from the second word to the highest point in the first vowel of the third word (indicated by the arrows). I have added to the pictures). This could be the rise of a type II intonation. The type is used, inter alia, as a "continuation intonation" (cf. e.g. Miletic 1960: 81-81).

However, I&L's quantitative data concerning these cases suggest to me that it is not, or is not only, the rise of type II which is under discussion: the rise in the vowel of the accented syllable is too gradual, and the highest point is reached too late, for the rise to be the (steep) rise of type II. Also, from the figures given in I&L (1965: 109), which concern individual words taken from a single sentence, it can be deduced, so far as I understand, that there is no large interval between the preceding words and the starting point of the vowels of the words immediately before the comma (which also have a gradual rise in the accented syllable). The gradual rise is comparable to the rise we found in 3.3 in the gradual variant of the rising accent (rather than to the steep rise in the intervocalic consonant(s) before falling accents). Elsewhere, for words immediately before a comma, a figure is given which averages over four sentences; as I&L indicate in their tables, individual cases deviate considerably from the average. This raises some suspicion that different types of intonation have possibly been lumped together.

At any rate, it seems to me that, possibly in addition to the rise of type II, a further rise, say type V, is represented in the data. Compare:

(The differences are reminiscent of the Russian intonation types E, H, and D, respectively, in Odé 1986.) Mahnken (1964: 84-85, 87-90) seems to associate type II with falling accents and type V with rising accents. Type II is a mode-1 type of intonation. With respect to type V, perhaps we must say that it is indifferent to mode in the intonational sense (the rise being so gradual that we cannot say whether it is early or late).

The issue is obviously highly relevant to the question of whether tone distinctions are preserved with various types of intonation. As we saw in 4.4 above, I&L found that tone distinctions are neutralized in III and IV (mode-2 types of intonation). In type V (and II?), in contrast, the tone distinction can be preserved:

"The contrast between falling and rising short accents was neutralized very often, the contrast between the two long accents less frequently. The contrast, when it existed, was manifested in a different relation between Ff peaks of the accented and the first posttonic vowel. In words with rising accents, it was rising, and in words with falling accents, it was falling (the same is basically valid also for intersyllabic intensity relations). In addition, the fundamental frequency movement within the vowel under the long rising (·) accent was usually rising up to the end of the syllable nucleus, whereas the vowel under (·) showed a rising-falling movement, the initial rise being followed by a moderate fall in the second half of the vowel". (I&L 1972: 131-132).

This conclusion pertains to words at the end of the first clause in a complex sentence. The different relationship between accented and postaccentual vowels which I&L found in cases where the tone distinction is preserved could be brought under the heading of the "belt" introduced in 3.6. But now the location of the belt must not be related to the level from which the obtrusion starts, but, I think, to the end of the accented syllable. In effect, the belt would be raised in its entirety, schematically:
The raised belt corresponds to what I&L call a "rising coefficient" (e.g. I&L 1972: 131). As I have tried to indicate above, I have some doubts about whether this story applies to type V only, or also to type II. (The perceptual distinctions involved here are strikingly similar to the distinctions being investigated by Odè in Russian intonation (1986: 427, and to appear); here, I am leaning heavily on Odè’s research.)

Finally, although I&L’s data do not suggest anything to the effect, I think it would be interesting to count whether in normal running speech there is an association between type II and falling accents, and between type IV and rising accents. The possibility occurred to me when I heard a speaker in a school radio programme explain to children the existence of tones in their language by saying the form sjènica twice, first with type IV, then with type II (but with a “half” fall in the last syllable). I guess the children were supposed to hear first sjènica (titmouse) and then sjènica (bower).

4.6 FINAL FALLS

An intonational mode distinction which is bound to be relevant for words carrying the last accent before a major (sentence) boundary is the difference between a fall which is preceded by an early rise or initial high fragment (“high fall”, mode-1) and an earlier fall, starting before the accented syllable (“low fall”, mode-2) (see 1.1). The former is more “emphatic” than the latter. (See, in application to Russian, Odè 1986 (types B and A, respectively) and Keijsper (1983: 118-125).) The “non-emphatic” type is sometimes omitted in descriptions (cf. 2.2.2-2.3 above), but it must be kept apart from the absence of an accent.

The following picture reproduced from L&I (1986: 202) possibly gives an example from Serbocroatian (note that the frequency scale is linear here):

The picture shows that the highest point in the first syllable of the last word is higher than the end of the preceding word with the falling accent (bàbama) but lower with the rising accent (tràvica). The same holds true for the corresponding picture for the long accents (I&L 1986: 201). If the difference recurs consistently, it may mean that an intonational distinction is being sacrificed here. However, only in some of the corresponding quantitative data can I detect the (last part of the) rise suggested by the picture for falling accents (see also e.g. I&L 1969: 137); and I&L (e.g. 1969: 161) mention that the descent of the fundamental frequency curve starts, as a rule, before the last word (see also Mahnken 1964: 73-75). Possibly, then, words with falling accents occur with both types. It is conceivable that rising accents are incompatible with the "emphatic" character of the mode-1 type. Note that the intonational mode distinction under discussion here differs from that dealt with in the preceding sections: we are now talking about the difference between an (early rise)fall (mode-1) and a fall starting in the preceding syllable (mode-2); in the reverse pattern this mode-2 fall is combined with a late rise (fall), which also has a mode-1 partner. The mode-2 fall alone seems to be combinable with the tonal mode distinction (expressed in the next syllable): “The end of the utterance is signalized by low, falling fundamental frequency and intensity. This transforms rising relations into falling ones, thus diminishing the physical distance between the
realizations of rising and falling accents, and frequently leads to a neutralization of the contrast between them, especially between the two short accents (• and '). As to the two long accents, the Ff fall in the vowel under (•) was usually smaller than in the vowel under (•), and laryngealization in the first posttonic syllable, which was regular in words with (•), occurred only sporadically in words with (•)* (I&L 1972: 131).

(In all cases, there are also intensity data.) This seems to be compatible with Mahnken's (1964: 73-75) observation that with falling accents, the (remainder of the) fall occurs in the accented syllable itself, whereas with rising accents, the postaccentual syllables contribute considerably to the total fall (cf. 3.6-3.7 above). Both descriptions allude to a "belt", this time a lowered one, schematically:

```
   *  
```

It seems to me that the issue requires further research, especially because it is unclear whether all data pertaining to this case concern accented words (see 2.2 above).

4.7 MINOR CONTINUATIONS?

Types II, IV and V mentioned in 4.2-4.5 above can be expected to be used as, e.g., "comma intonations" and "colon intonations". Elsewhere, for example at the beginning of a sentence or immediately after a comma (to take I&L's type of classification), there may be less outspoken signals that more information is to follow. In these cases I&L also found a "rising coefficient" (e.g. I&L 1972: 132), but it is smaller than in the "major continuation" types. As compared with the "neutral" cases discussed in Chapter Three, the upward effect is probably especially relevant for the recognition of falling accents, because the falling interval between the accented and the first postaccentual syllable tends to become smaller. For example, in the type "" cited in 3.4 from I.'s speech, pitch was already 6.0 ST under the peak at the end of the accented syllable (8.7 ST at the end of the postaccentual vowel). The average for the same type in the beginning of a (statement) sentence, spoken without emphasis (I&L 1970: 226), has the end of the accented vowel only 1.4 ST under the peak; in the beginning of the postaccentual vowel pitch has fallen 2.7 ST, and at the end of this vowel 5.0 ST. In itself, smaller excursions need not affect the relevant relationships, namely if the preceding rising part of the obtrusion becomes smaller as well. However, some pictures in L&I (1986: 201-207), as well as a part of the quantitative data (I&L 1970: 226-233; 1972: 110-115), suggest that, with falling accents, the low line indicated in 3.6-3.7 may be resumed in the post-postaccentual syllable. This means that such realizations are close to realizations of rising accents with a relatively early peak, along the lines of the long example given in 3.6. But I would not dare to say anything about the issue without further melodic context.

Another issue which requires further context is the position of the accented syllable, in rising accents, with respect to its surroundings. In a pilot investigation with nonsense words, Jacobsen (1977) found that with rising accents, the accented syllable is considerably below its melodic context. The same is suggested by the schematic representation in L&I (1986: 204) and by I&L's (1965: 109) information concerning individual words from a single sentence (short rising accent in the word immediately after the comma). This would be a logical realization of a mode-2 accent (cf. type J in Odé 1986). But a discussion of this possibility here would be pure guesswork.

4.8 NO SENTENCE-FINAL RISE

Finally, I would like to mention that a non-prominence lending rise in the last syllable of a sentence does not occur in the available data. Miletic (1960: 86) mentions that speakers of Serbo-Croatian have difficulty with a German question like Kann ich ein Zimmer haben?, spoken with the main accent on Zim- and a rise in the last syllable (-ben). The absence of this rise groups Serbo-Croatian with Russian rather than with, e.g., German, English and Dutch (see on this point Odé (1986: 436) and Keizer (1983: 134-135, 144-147)).
For a comparison of the question intonations of Serbocroatian, Rumanian and Albanian see LSI 1980.

4.9 CONCLUSION

It will be clear from the foregoing that my attempt to interpret I&L's data about sentence intonation has not been very successful. It seems better to obtain first the indispensable further information and then to fit I&L's data into the picture rather than to try and predict details of sentence intonation on the basis of what is now known. The technical impediments which existed at the time when I&L started their investigations have been eliminated since then, so that nothing seems to prevent further research on Neo-Kotkovian prosody. If it is conducted with the same care that is present on every page of I&L's work, it will be an appropriate continuation.

Leiden University

* I would like to thank J. 't Hart and W. Vermeer for their constant willingness to share their expert knowledge of, respectively, pitch perception and Serbocroatian tonal systems.

REFERENCES

Appel, W.

Bacić, J.

Belić, A.

Bolinger, D.L.


Bray, R.G.A. de

Bryzunova, E.A.
1977 Zvuk i intonacija russkoj reči, 3-e izdanje, pererabotannoe. Moskva.

Cohen, A., R. Collier, J. 't Hart

Collier, R.

Collier, R. J. 't Hart

Delattre, P., E. Pienack, C. Olsen

Garde, P.

Gussenhoven, C.

Gvozdanović, J.


J. 't Hart, J.

J. 't Hart, J., A. Cohen


Rossi, M., M. Chafcouloff 1972 "Recherche sur le seuil différentiel de fréquence fondamentale dans la parole", Traavaux de l'Institut de Phonétique d'Aix 1, 179-195.

Santen, J. 1972 "O fonetičeskoj i fonologičeskoj prirode udarenija v serbko-chorvatskom jazyke", Sovetskoe slavjannovedenie 8, 4, 100-105.