

# Bisyllabicity and Tone

*Ilse Lehiste*

The Ohio State University  
ilsele@ling.ohio-state.edu

## 1. Introduction

This paper deals with three instances in which the emergence of contrastive tone is tied with bisyllabicity. The emerging tonal patterns presuppose the presence of two syllables. In other words, the domain of the contrastive prosodic pattern is a sequence of two syllables. There are no tonal contrasts on monosyllabic words, and in a polysyllabic word, two syllables determine the prosodic identity of the word. The three languages considered here are Serbocroatian, Swedish, and Estonian. It will be shown in each case how the prosodic structure is connected with bisyllabicity.

## 2. Serbocroatian

The prosodic system of standard Serbocroatian comprises both quantity and tone. Quantity is independent of tone: contrastive vowel length can occur in any position, regardless of stress or tone. Vowels in monosyllabic words can also be contrastively short or long. Tonal opposition is associated with stress in the following way and for the following historical reasons.

Serbocroatian is one of the South Slavic languages. It has several dialects, some of them relatively archaic, others more innovating. The Štokavian dialect has become the basis of the literary language. Štokavian has four tones – or rather accents – short falling, short rising, long falling, and long rising. The

present state of the prosody of Štokavian is due to an accent shift, called the Neoštokavian accent shift and reconstructed to have occurred around the 14th century (Popović 1960, p. 403). There is no convincing explanation as to **why** the accent shift occurred, but the **how** is quite clear: stress shifted by one syllable toward the beginning of the word.

As a result, the final syllable of a word is never stressed. Disyllabic words are always stressed on the first syllable. The contrast in tone depends on whether the first syllable was stressed originally, or whether it acquired its stress due to the Neoštokavian accent shift. The pitch curve on the originally stressed syllable is falling; the pitch curve on the syllable bearing the shifted accent is rising. But this is not the only difference between the two disyllabic words: there is also a difference in the second syllable. After an originally stressed first syllable, the second syllable has low pitch; after a syllable with shifted accent, the second syllable has high pitch. The phonetic process of the accent shift left the high pitch associated with stress on the originally stressed syllable, while moving the other characteristics of stressedness to the preceding syllable.

Figure 1 is an illustration of fundamental frequency curves associated with Serbocroatian falling and rising accents. This is based on extensive studies that I have carried out with my Serbian colleague Pavle Ivić, and is taken from one of our early publications (Lehiste and Ivić 1963).

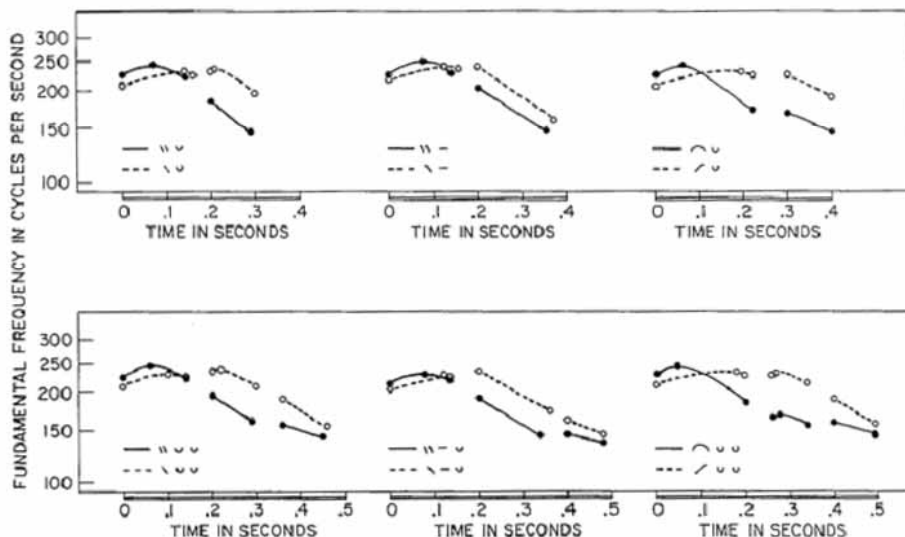


Figure 1: Average fundamental frequency patterns in Serbocroatian words with falling and rising accents.

On the figure, the solid line represents F0 movements in falling accents, and the dashed line shows the pitch curves associated with rising accents. Three disyllabic accentual patterns are shown on the top half, and three trisyllabic patterns on the bottom half. The first pattern on the left represents words with a short accented syllable followed by a short posttonic syllable. The second pattern (the one in the middle) shows the pattern found on words with a short accented syllable and a long posttonic syllable (recall that vowel quantity is independent of accent). The pattern on the right shows the pitch pattern on words with a long first syllable and a short second syllable. The trisyllabic patterns represent words where an extra unstressed short syllable has been added to the disyllabic pattern.

As can be seen from the figure, there is a difference in the pitch curves on the stressed syllables, depending on the length of the syllables. The long falling and long rising accents are clearly differentiable on the basis of the shape of the F0 on the stressed syllable itself. The pitch peak occurs during the first third of the long falling syllable nucleus, and almost at the end of the long rising syllable nucleus. The difference in the pitch level of the posttonic syllable is of course present, but in the case of long accents, one might claim that it is conditioned by the pitch of the preceding syllable and thus not independently contrastive. It is the short accents that depend on the height of the second syllable to manifest the contrast between falling and rising accents. The pitch peak in the short syllable with a falling accent is approximately in the middle and the difference between the two contours appears minimal. Here it is clearly the high pitch on the posttonic syllable that differentiates the short rising accent from the short falling accent.

There are certain distributional restrictions on the occurrence of the two accents that are due to their origin. Falling accents occur only on the first syllable of a word. Rising accents can occur on any syllable except the last syllable of a word. It follows also that monosyllabic words have only falling accents, since there was no syllable to which the accent could have been shifted, producing a rising accent. The falling pitch contour on the syllables now said to bear the falling accent was originally just a characteristic of their stressedness; stressed syllables had high pitch and a falling pitch movement. That falling pitch movement became a contrastive tone after the Neostokavian accent shift had created a rising tone. The contrast between falling and rising requires two syllables for its manifestation.

### 3. Swedish

Swedish and Norwegian are two North Germanic languages that have developed contrastive accents. They are usually referred to as “Accent 1” and “Accent 2”. These are actually cover terms for contours that may have more than one phonetic manifestation. To simplify it a great deal, Accent 1 is a one-peaked accent and Accent 2 is a two-peaked accent. Accent 1 became contrastive with the development of the pattern now called Accent 2 (this is comparable to the Serbocroatian situation, where the falling accent became contrastive with the development of the rising accent.) There is a great deal of research dealing with the dialectal manifestations of the accents; I will follow the outline given in the most recent treatment available to me (Riad 2003).

Riad claims that the two-peaked accent originated in the context of stress clash. The two pitch peaks originated as two stress peaks. In clash situations the secondary stress was reduced, but the pitch part was retained. Where stress reduction became permanent, the tonal information became lexical – in other words, the difference between one-peaked and two-peaked contours became a tonal opposition.

This development is associated with syncope and other prosodic developments in late Old Norse, the parent language of all Scandinavian languages, and its beginnings are dated about a thousand years ago. The stress clash developed in three environments – in compounds, nouns with derivative suffixes, and nouns with inflectional suffixes. The second element of the compound and the derivative suffix carried secondary stress in Old Norse. Old compounds now mostly carry the two-peaked Accent 2. Many suffixes that were stressed in Old Norse are no longer stressed, but they induce Accent 2 in the forms to which they are attached. Inflectional suffixes are all unstressed at the present time. Riad says that “the process of replacing stress with lexical tone is productive in the same way today as it was a good millennium ago” (p. 95).

In the present context, Swedish illustrates the association of tone with bisyllabicity: a sequence of two syllables is necessary for the manifestation of contrastive tone.

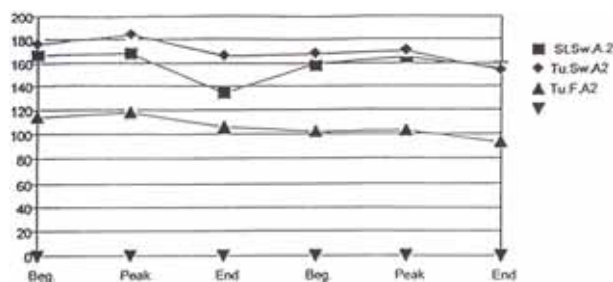


Figure 2 : Average fundamental frequency patterns in Swedish disyllabic words with Accent 2, produced by Swedish speakers in Turku, Finland, and Stockholm, Sweden.

Figure 2 is an illustration of the Swedish accents from a study that I carried out myself (Lehiste 1997a). This was a comparison of standard Swedish with the form of Swedish spoken in Turku, Finland. There has been a Swedish population in Western Finland since prehistoric times. One characteristic of their form of Swedish is absence of lexical tone. The figure presents fundamental frequency profiles for the two-peaked accent, namely Accent 2.

The data points represent the F0 values at the beginning, peak, and end of each of the two syllables constituting the domain of the accent. The top line is the average for five speakers from Turku, and the lower line is the average for five Stockholm speakers. The patterns were averaged from 51 words in which Accent 2 was expected. (The difference between the average fundamental frequency levels is due to the fact that four of the five Turku speakers were women.)

There is one dramatic difference between the Turku speakers and Stockholm speakers, and that is in the two-peaked pitch contour on Accent 2 in the words produced by

Stockholm speakers. The Turku speakers have just a slight rise on the first syllable, followed by gradual decline in pitch; the Stockholm speakers have a fall at the end of the first syllable and a rise in the second syllable. There is no lexical tone in the Turku dialect, but the two accents are clearly differentiated in the Stockholm form of the language, and two syllables are required for manifesting the distinction.

#### 4. Estonian

In the preceding two examples, I have shown how the emergence of tone was connected with bisyllabicity, with the result that the domain of the manifestation of tone is a disyllabic sequence. In my third example, I would like to present a case where tone is in the process of developing at the present time, and how the development of tone depends on the emergence of disyllabic sequences as the basic units within the prosodic system of the language (Lehiste 2003).

The language in which this process is under way is Estonian. The language belongs to the Baltic-Finnic branch of the Finno-Ugric language family, and it is a close relative of Finnish. Tone has not been reconstructed for the ancestor language of Baltic-Finnic languages, but all languages of this group have some form of contrastive quantity. Certain sound changes took place in Estonian less than a thousand years ago, when the protolanguage of Finnish and Estonian began to split into the two sister languages. The sound changes relevant in this context are syncope and apocope – loss of a short unstressed vowel after a long syllable.

Apocope – loss of a final vowel – is dated in the thirteenth century, and syncope – loss of a medial vowel – is dated in the sixteenth century (Kask 1972). Both processes affected short vowels of an open succeeding syllable after a long preceding syllable. For example, the word *laulma* ‘to sing’ had an earlier form *\*taulamahan*; the /a/ of the second syllable was lost through syncope. An earlier form *\*jalka* ‘leg’ became *jalg* through apocope. (Vowel loss after a long syllable did not take place in Finnish, where the contemporary forms are *laulamaan* and *jalka*.)

The loss of the second syllable vowel was accompanied by compensatory lengthening of the originally long syllable, which now acquired the extra characteristics of overlength. As a result, Estonian now has three degrees of quantity: short, long, and overlong. This applies both to vowels and to consonants. There are certain restrictions as to where the three contrasts may occur which I will not go into at the present moment.

An additional change took place at the same time: the development of disyllabic sequences, or metric feet, into units with an internal structure manifested in the durational relationship between the two syllables. Metric feet most probably existed already in the protolanguage of Finnish and Estonian. Both Finnish and Estonian words have the property that there is a primary stress on the first syllable of a word, and secondary stresses on successive odd-numbered syllables, i.e. on third, fifth, seven, ninth etc. syllables. The sequence of stressed syllable + unstressed syllable constitutes the disyllabic metric foot. The specific Estonian development is the emergence of temporal compensation between the syllables, so that when the first syllable becomes longer, the second syllable becomes shorter. This results in a kind of isochrony

of metric feet. The three kinds of metric feet – short, long, and overlong – are characterized by the durational ratios between the two syllables of the metric foot, which are 2/3 for the short quantity, 3/2 for the long quantity, and 2/1 for the overlong quantity. These durational ratios between the syllables have been established by extensive acoustic-phonetic studies (for a review, cf. Lehiste 1997b), and been confirmed by other researchers (e.g. Krull 1993).

The difference between the two long degrees is accompanied by a difference in the fundamental frequency contour. Words in the two original quantity degrees, short and long, have a typical rising-falling pitch contour, where the first syllable is rising and the second syllable is falling. But disyllabic sequences in overlong quantity have a rising-falling contour on the first, overlong syllable, followed by a low second syllable. I have hypothesized that this is associated with the originally disyllabic status of the overlong first syllable: together with compensatory lengthening, the pitch contour of the disyllabic unit was transferred to the overlong syllable (Lehiste 1978).

I tested the relevance of the fundamental frequency contour in a series of listening tests with synthesized stimuli, of which only a sample can be presented here. (Lehiste 1970-1975). There are many minimal triples in Estonian that differ only in their suprasegmental structure. I used the minimal triple *sada* – *saada!* – *saada* (‘hundred’ – ‘send!’ – ‘to get’), and produced the stimuli with first vowel durations ranging in 10-msec steps from 120 to 240 msec. Each stimulus was provided with three pitch contours and three second-syllable durations, and listeners had to identify the words presented in random order. Since we are interested in the role of the pitch contour on long vowels, let us look at listener responses to changes in the fundamental frequency contour.

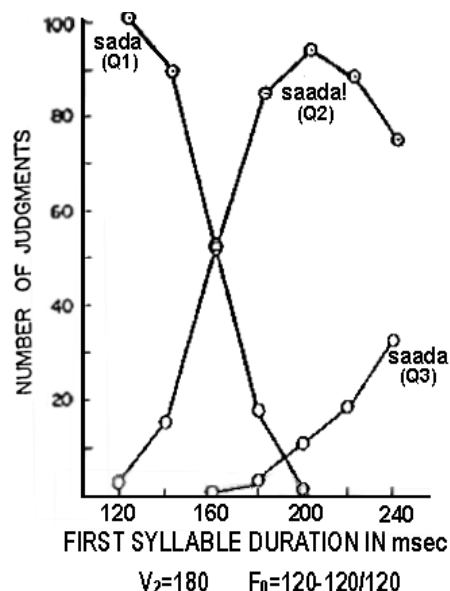


Figure 3: Listener identification of synthesized Estonian words as belonging to short (Q1), long (Q2), and overlong (Q3) category with monotone fundamental frequency.

Figure 3 shows listener responses when the stimuli were provided with monotone fundamental frequency. From 180

msec on, the listeners heard the word to be in long quantity – to represent *saada!*. Even with the longest first syllable, assignment to overlong quantity did not surpass 30%.

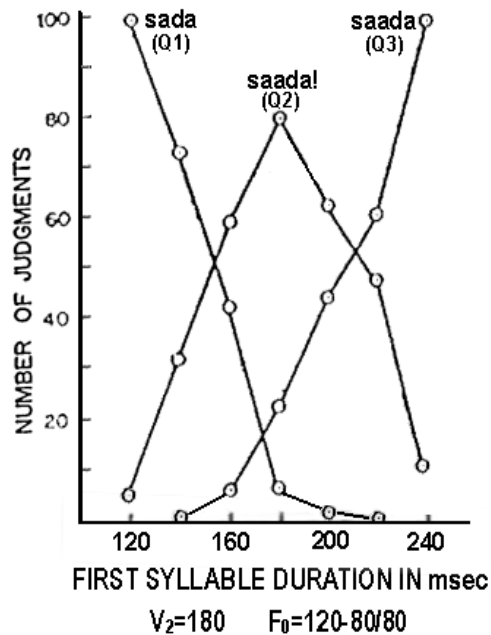


Figure 4: Listener identification of synthesized Estonian words with falling fundamental frequency on the first syllable.

But as demonstrated in Figure 4, when the fundamental frequency was changed to falling on the first syllable, listeners began to hear the stimuli as overlong – they identified it as *saada*. And the gain in ‘overlong’ assignments was at the expense of ‘long’ judgments. Stimuli with the same duration were heard as long on monotone pitch, and as overlong with falling pitch.

In addition to having the pitch contour of a disyllabic word, the overlong syllables behave in many ways as if they were still disyllabic words. An overlong syllable is always stressed, be it primary or secondary stress. An overlong syllable may be followed by a syllable bearing secondary stress. This can be explained from the disyllabic origin of overlength. Estonian words normally carry primary stress on the first syllable and secondary stresses on odd-numbered syllables. The syllable following an overlong syllable was originally the first syllable of the next disyllabic foot.

## 5. Summary and conclusions

In this paper, I have described two cases – Serbocroatian and Swedish - where contrastive tone requires for its manifestation a sequence of two syllables. In Serbocroatian, high pitch was originally a feature of a stressed syllable; through historical development high pitch was disassociated from stress and became contrastive in its own right. In Swedish, the contrastive accent patterns developed likewise from original sequences of two stresses, of which the second was reduced in

stress, but retained the higher pitch peak formerly associated with stress, resulting in a two-peaked pitch contour. In the third case – Estonian, the prosodic system is still basically quantitative, but fundamental frequency is one of the features that distinguish between two long degrees quantity. In that restricted sense it has a distinctive function. In all three languages, there are no contrasts on monosyllabic words, and a disyllabic sequence is necessary for the manifestation of contrastive tonal patterns.

## 6. References

- Kask, Arnold, 1972. *Eesti keele ajalooline grammatika*. Tartu: Tartu Riiklik Ülikool.
- Krull, Diana, 1993. Word-prosodic features of Estonian conversational speech: Some preliminary results. *PERILUS XVII*. Stockholm: Department of Linguistics, Stockholm University, 45-54.
- Lehiste, Ilse, 1975. Experiments with synthetic speech concerning quantity in Estonian. *Congressus Tertius Internationalis Fenno-Ugristarum, Tallinn 1970*. Pars 1: Acta linguistica. Tallinn: Valgus, 254-269.
- Lehiste, Ilse, 1978. Polytonicity in the area surrounding the Baltic Sea. *Nordic Prosody: Papers from a Symposium*. Eva Gårding, Gösta Bruce and Robert Bannert (eds.). Lund: Department of Linguistics, Lund University, 237-247.
- Lehiste, Ilse, 1997a. Cross-linguistic comparison of prosodic patterns in Finnish, Finland-Swedish, and Stockholm Swedish. In *Language and its Ecology: Essays in Memory of Einar Haugen*, Stig Eliasson and Ernst Hakon Jahr (eds.). Trends in Linguistics Studies and Monographs 100. Berlin: Mouton de Gruyter, 367-377.
- Lehiste, Ilse, 1997b. Search for phonetic correlates in Estonian prosody. *Estonian Prosody: Papers from a Symposium*, Ilse Lehiste and Jaan Ross (eds.). Tallinn: Institute of Estonian Language, 11-35.
- Lehiste, Ilse, 2003. Prosodic change in progress: from quantity language to accent language. In *Development in Prosodic Systems*, Paula Fikkert and Haike Jacobs (eds.). Berlin/New York: Mouton de Gruyter, 47-65.
- Lehiste, Ilse; Ivić, Pavle, 1963. *Accent in Serbocroatian: An Experimental Study*. Michigan Slavic Materials 4. Ann Arbor: Department of Slavic Languages and Literatures, University of Michigan, 1-142.
- Popović, Ivan, 1960. *Geschichte der serbo-kroatischen Sprache*. Wiesbaden: Otto Harrassowitz.
- Riad, Tomas, 2003. Diachrony of the Scandinavian accent typology. In *Development in Prosodic Systems*, Paula Fikkert and Haike Jacobs (eds.). Berlin/New York: Mouton de Gruyter, 91-144.
- Ross, Jaan; Lehiste, Ilse, 2001. *The Temporal Structure of Estonian Runic Songs*. Berlin/New York: Mouton de Gruyter.