UNIVERSITY OF NIŠ
The scientific journal **FACTA UNIVERSITATIS**
Series: **Linguistics and Literature** Vol.1, No 4, 1997 pp. 283 - 290
Editor of series: Nedeljko Bogdanović, e-mail: nedbog@filfak.filfak.ni.ac.yu
Address: Univerzitetski trg 2, 18000 Niš, YU, Tel: +381 18 547-095, Fax: +381 18 547-950
http:// ni.ac.yu/Facta

FUNDAMENTAL FREQUENCY CHANGE IN TWO-SYLLABLE WORDS WITH LONG RISING ACCENT

UDC:808.61-612

Mirjana Sokolović

Faculty of Philosophy, Niš, Yugoslavia

Abstract. *Fundamental frequency change in two-syllable words with long rising accent is investigated in this paper. Experimental results show that long rising accent is realized in similar way by all speakers. Fundamental frequency in stressed vowel rises during the whole vowel till the end, where it has peak, or near the end, where after peak it has slight fall. Beginning f_0 value in unstressed vowel is close to ending value of preceding syllable. Peach starts with peak, or rises very little in the beginning, before reaches the peak, and then falls till the end of the vowel. According to average results for four male speakers proposition for f_0 change which should be included in speech synthesis programme is given.*

INTRODUCTION

Speech communication between man and machine is not only brave visionary idea, it is becoming reality of the modern world. Within the most actual interdisciplinary areas which investigate human language are speech synthesis and automatic speech recognition. Such programmes have already been developed for most of the world languages. Work on TTS (Text-To-Speech) system for the Serbian language is in progress. The most important assignments of these devices are so called "voice" cards for post telegram and telephone services, which are supposed to give information about exact time, telephone numbers etc., and document reading machines for blind people and disabled people.

Current TTS system for the Serbian language does not have prosodic features.[1] Experiments with synthetic speech have shown that it is very intelligible, which was the aim of the first phase. However, this speech is monotonous and not natural enough.[2] The

Received, December 20, 1996

[1] Detailed description of the system is given in T. Tomić et al. (1986), and some of the result (synthesis of consonant) in S. Petar et al. (1986).

[2] As the authors say in S. S. Ilić and M. D. Savić (1995), synthetic speech with constant f_0 was monotonous, so

aim of the next phase is to add prosodic features, at first word prosody and than utterance prosody. This paper is included in those investigations. It is a part of voluminous investigation which aim is to explore possibilities for improving current TTS system by adding parameters which can be controlled and changed in simple way. In this paper word prosody influence on f_0 change is investigated and proposition how to improve the TTS system is given.

METHOD

Numerous investigations have been devoted to the examination of Serbian accents. Previous studies have been carried out firstly to explain their nature. The most important experimental results are given by P. Ivić and I. Lehiste (1963; 1965; 1967), A. Peco and P. Pravica (1972), J. Jokanović-Mihajlov (1983) and I. Lehiste and P. Ivić (1996). Starting with their results, chosen corpus of two-syllable words with long rising accent[3] is investigated in this paper and it is found out how f_0 changes in relation to long rising accent. According to the results, average values, which could be the basis for prosodic word features, are given.

The test material consists of next words (structure CVCVC is chosen as the most convenient): (/∪) lavež, belac, miran, ljubav, brlog, krpež; (/–) Majom (I sg), vezem, lopov, vučem, krčim (<krčati). Words are chosen so that in stressed syllable occur all vowels and syllabic *r*, as well as words with short unstressed syllable and unstressed length.

Words are pronounced by four male speakers: 1. GP, a student of the Serbian language and literature, born in 1972 in Vareš. Primary and secondary school finished in Kalesija. 2. GM, a teaching assistant, born in 1963 in Foča. Primary and secondary school finished in Foča, Faculty of Philosophy in Sarajevo. 3. AS, a student of medicine, born in 1974 in Užice. Primary school finished in Čajetina, secondary school in Užice. 4. MM, a student of civil engineering, born in 1977 in Tuzla. Primary school finished in Tuzla, secondary school started in Tuzla and finished in Niš.[4]

The speakers pronounced isolated words, mixed with other words in order to avoid monotony. Recording and analysis are accomplished using PC 486 with sound blaster Vibra 16 (Creative Labs). Recoding is repeated two or three times and the best recordings are chosen for analysis[5]. Software "Swell" is used for f_0 analysis. The software provides speech signal analysis, and one of its parts gives f_0 curve. In this paper, f_0 curve is done for every word, and next values at characteristic points are read from the curve: at the beginning of vowel, at the end of vowel, f_0 peak, also vowel duration and the moment when f_0 is at its peak.

---

they tried to break monotony by changing f_0 for ±1.25%. Such speech was more natural, but hoarse.

[3] Analysis of one-syllable words with long falling accent can be found in M. Sokolović (1997) and

[4] Since it has been worked mainly with male voice in speech synthesis in our country, results for male speakers are given in this paper.

[5] In reading there were not mistakes in word accent. Only some unstressed lengths were shortened, which can be seen in average values.

RESULTS

Fundamental frequency change in words produced by GP

Average f_0 values measured at characteristic points (at the beginning, at the peak, at the end) are given in table 1. Next marks are used: f_{0b} - f_0 value at the beginning of the vowel, f_{0p} - peak value, f_{0e} - f_0 value at the end of the vowel; t_p - peak position measured from the beginning of the vowel, t_e - vowel duration (the same marking is used for other informants).

Table 1. Average f_0 values in words with (/) accent produced by GP

GP		f_{0b} (Hz)	f_{0p} (Hz)	f_{0e} (Hz)	t_p (ms)	t_e (ms)
/∪	stressed vowel	107.55	123.46	120.98	186.00	223.67
	unstressed vowel	116.45	119.36	103.33	25.33	113.00
/–	stressed vowel	109.24	121.31	119.27	177.50	196.25
	unstressed vowel	121.94	123.48	103.32	8.25	124.25

Word type (/∪). In those words f_0 rises in the largest part of the stressed vowel, with slight fall at the end of the vowel. The average rising interval is 15.91 Hz, and average falling interval is 2.48 Hz. Tonal peak is reached on average at 83.16% of the vowel duration, but in some words these values are between 60 and 95%.

Unstressed short vowel begins with f_0 value which is under the end of the stressed vowel (on average for 4.53 Hz). Relation between rise and fall is reversed: after average rise of 2.91 Hz, f_0 falls for 16.03 Hz - movement is similar to movement in stressed vowel, but it is shorter and it is situated at lower f_0 range. Average peak position is 22.42%, but in different words it is between 18 and 88% of the vowel duration.

Word type (/–). F_0 in stressed vowel has long rise (on average 12.07 Hz), and then falls for 2.04 Hz, which is very similar to fo movement in word type (/∪). F_0 peak is on average at 90.45% of vowel duration, and measured values vary between 67 and 100% of vowel duration. It is interesting to point out that GP has only one word with peak at the very end of the stressed vowel (*vučem* - this is also the shortest stressed vowel for informant GP).

Unstressed long vowel starts at f_0 level which is above the final value of the stressed syllable (on average 2.67 Hz). However, in some words relations are different: in words *lopov* and *vučem* f_0 peak is at the beginning of the second syllable and above the whole stressed syllable; in word *vezem* f_0 peak is above f_0 peak of the first syllable, and in word *Majom* all characteristic values are under the lowest f_0 value of stressed syllable. Average f_0 rise in unstressed long vowel is 1.54 Hz and average fall is 20.16 Hz. All words (except *vezem*) begin with the highest f_0 value for the vowel, so that average peak position is 6.64% of the vowel duration.

F_0 change can be shown as in the figure 1 (according to average values). On comparative figure both accent patterns are shown. Type (/∪) is marked by solid line, and type (/–) with broken line (the same marking is used for other informants). F_0 movement is as in the figure, somewhere with more straight line at both peak sides (for example *miran*, *Majom*). The only exception is word *vučem* where f_0 contour is concave.

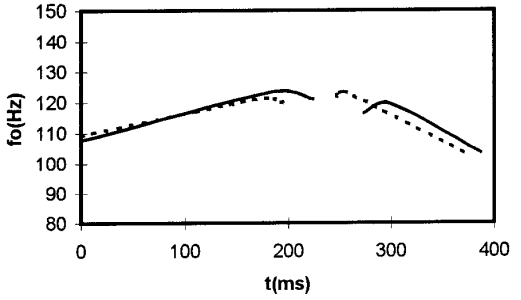

Fig. 1. Graphic presentation of f_0 change in words with (/) accent produced by GP.

Fundamental frequency change in words produced by GM

Word type (/∪). The long rising accent in this informant's pronunciation is realized with long and big f_0 rise of 16.33 Hz and little fall of 1.72 Hz (average values). Tonal peak is reached at 93.77% (on average) of the vowel duration. Majority of words has f_0 peak at the beginning of the stressed vowel, and f_0 peak is in the 83-100% range of vowel duration.

The next syllable beginning is lower than the end of the stressed vowel (for 4,01 Hz on average). There are some words where starting value (i. e. peak) is higher than the ending value of stressed vowel (words *lavež*, *miran*). During the short unstressed vowel f_0 rises for 4.48 Hz on average, and then falls for 14.4 Hz, although in majority of the words f_0 peak is at the beginning (these are the same words which have peak at the end of the vowel in the first syllable).

Word type (/–). F_0 in stressed vowel rises during the vowel (in half of the words reaches the end of the vowel), and then falls. Average rise is less than in type (/∪) and equals 13.24 Hz. Average fall is 1.86 Hz. F_0 peak is reached at the end (as in type (/∪)), on average at 91.54% of the vowel duration (measured values are between 79 and 100%).

Unstressed long syllable is different from short syllable. F_0 movement starts with value which is 0.61 Hz higher then ending value of the stressed syllable, and peak is equal as starting value (in all words). Pitch falls from the beginning to the end (11.37 Hz on average). Such contour occurs in the majority of the analyzed words, for example *vezem*, *lopov*, *vučem*, where the beginning of the second vowel is higher or near the end of the first vowel. The second syllable begins at lower level only in word *Majom* (about 2 Hz).

Average f_0 values are shown in table 2 and f_0 movement in figure 2. In all words pitch rises almost in straight line, but in word *miran* slope is less at both sides of the peak. In word *vučem* curve is concave.

Table 2. Average f_0 values in words with (/) accent produced by GM

GM		f_{0b} (Hz)	f_{0p} (Hz)	f_{0e} (Hz)	t_p (ms)	t_e (ms)
/∪	stressed vowel	100.04	116.37	114.65	216.00	230.33
	unstressed vowel	110.64	115.52	101.12	11.670	123.33
/–	stressed vowel	103.03	116.27	114.41	243.50	266.00
	unstressed vowel	115.02	115.02	103.65	000.00	135.50

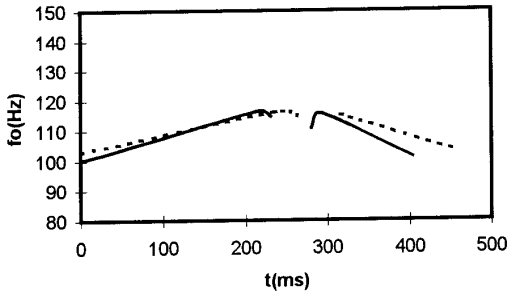

Fig. 2. Graphic presentation of f_0 change in words with (/) accent produced by GM.

Fundamental frequency change in words produced by AS

In general, everything which is told about f_0 change in the vowel with long rising accent for previous speakers is the same for AS. Peach is rising during the largest part of the stressed vowel, tonal peak is placed at the end of the vowel, where it is followed by slight f_0 fall, or at the very end, where final value is also peak value. Average values are shown in table 3 and f_0 movement in figure 3.

Table 3. Average f_0 values in words with (/) accent produced by AS

AS		f_{0b} (Hz)	f_{0p} (Hz)	f_{0e} (Hz)	t_p (ms)	t_e (ms)
/∪	stressed vowel	101.90	123.58	120.47	163.40	187.8
	unstressed vowel	114.74	121.41	107.62	18.02	97.8
/–	stressed vowel	102.75	126.47	123.31	180.20	191.6
	unstressed vowel	122.30	122.75	108.04	9.40	117.4

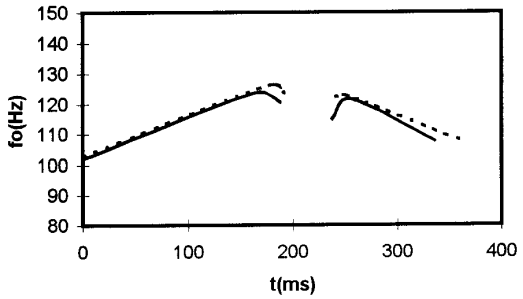

Fig. 3. Graphic presentation of f_0 change in words with (/) accent produced by AS.

Word type (/∪). Average f_0 rise is 21.08 Hz, fall is 3.11 Hz. Peak is on average at 87.01% of the vowel duration, and values are in small range of 78-100%.

Next vowel begins at lower frequency (for 5.73 Hz on average) compared to the end of stressed vowel. As in the majority of unstressed vowels, f_0 movement is rising-falling, with average rise of 6.67 Hz and average fall of 13.79 Hz. F_0 peak values are between 0 and 43% of the vowel duration, which gives average value of 18.67%.

Word type (/–). Relations are similar in this word type. Average rise is 23.72 Hz, average fall 3.16 Hz. F_0 peak is close to the end of the word (on average 94.05%), because peaks are concentrated in the 85-100% range of the vowel duration.

The beginning of the second vowel is lower than the lowest point of the stressed vowel (1.01 Hz on average), with similar f_0 movement. Rise at the beginning is slight (0.45 Hz on average), and fall is bigger (14.71 Hz on average). These values appear because peak is at the beginning in most words, and all values are between 0 and 30% of the vowel duration.

Fundamental frequency change in words produced by MM

Neither this information's long rising accent pronunciation has any individual features. Average values are shown in table 4.

Table 4. Average f_0 values in words with (/) accent produced by MM

MM		f_{0b} (Hz)	f_{0p} (Hz)	f_{0e} (Hz)	t_p (ms)	t_e (ms)
/∪	stressed vowel	111.56	124.31	122.54	149.25	167.25
	unstressed vowel	119.94	126.64	118.34	32.75	96.00
/–	stressed vowel	107.19	122.82	121.67	179.25	199.50
	unstressed vowel	123.68	124.21	115.45	10.00	94.25

Word type (/∪). Tonal rise is sharp (12.75 Hz on average) and ends at 89.24% of the vowel duration (on average), after which f_0 slightly falls till the end (1.77 Hz on average). Peak is not at the very end in any word, and measured positions are very compact and gathered between 84 and 94% of the vowel duration. In the rising part f_0 curve is mostly slightly convex, and in some words straight (*lavež*, *miran*).

In majority of the words the beginning of the next vowel is lower than the end of the preceding one, on average 2.6 Hz (except in word *brlog*, where it is higer), and further course is rising-falling with similar rise and fall (6.7 Hz average rise, 8.3 Hz average fall), but the fall is three times longer, so that it is slighter. Average peak position at 34.11% involves values from 17 to 50%. F_0 peak in the second vowel is higher than f_0 peak in the first vowel, but the end is under the first syllable's end in all words.

Word type (/–). In these words average relations are mostly unchanged (average rise is 15.63 Hz, average fall is 1.15 Hz). F_0 peak is between 80 and 100% of the vowel duration (on average 89.85%), and f_0 changes to the peak at concave line (except *lopov* where it is convex).

The greatest difference between this type and type (/∪) is relation between first vowel's end and the second vowel's beginning. The second syllable is higher for 2.01 Hz, and this occur in all words. The highest f_0 value is mostly above the first syllable's peak (except in word *Majom*), so that the rise at the beginning is slighter (on average 0.53 Hz), although in many words it does not exist because f_0 starts from the highest value, and fall is similar to fall in type (/∪) and equals 8.76 Hz. The second vowel ends with peak f_0 value under ending f_0 value in the stressed vowel.

Comparison between f_0 movement for both types is shown on figure 4. F_0 changes at concave line at rising, except in words *lavež* and *miran* where line is straight, and *lopov* where it is convex.

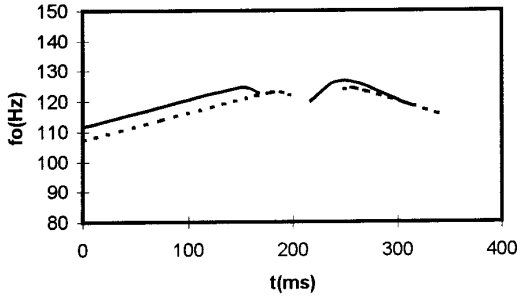

Fig. 4. Graphic presentation of f_0 change in words with (/) accent produced by MM.

CONCLUSION

Experimental results show that long rising accent is realized in similar way by all speakers, with insignificant variations. Fundamental frequency in stressed vowel rises during the whole vowel till the end, where it has peak, or near the end, where after peak it has slight fall. Beginning f_0 value in unstressed vowel mostly is close to ending value of preceding syllable. Peach starts with peak, or rises very little in the beginning, reaches the peak, and then falls till the end of the vowel. Because of pronunciation uniformity of long rising accent for all speakers, it can be concluded that for speech synthesis of this accent in two-syllable words is enough to use average values shown in table 5 and in figure 5.

Table 5. Average f_0 values in words with (/) accent for all speakers.

average		f_{0b} (Hz)	f_{0p} (Hz)	f_{0e} (Hz)	t_p (ms)	t_e (ms)
/∪	stressed vowel	105.26	121.93	119.66	178.66	202.26
	unstressed vowel	115.44	120.73	107.60	21.94	107.53
/–	stressed vowel	105.55	121.72	119.67	195.11	213.34
	unstressed vowel	120.74	121.37	107.62	6.91	117.85

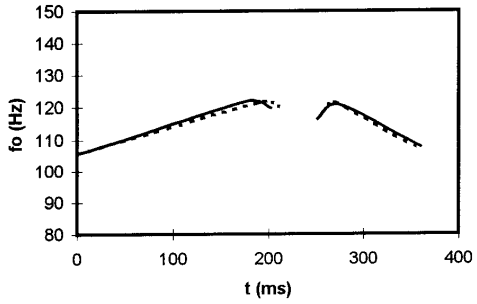

Fig. 5. Graphic presentation of f_0 change in words with (/) accent for all speakers.

Since these are values for single words, in further work on programme word accent should be coordinated with sentence prosody. While including words into sentences,

sentence prosody influence on word accent should be taken into consideration. Previous results for this word type (in S. Peter at al. 1992) show that in that case word duration shortens, and f_0 peak rises.

REFERENCES

1. Ilić S. S., M. D. Savić, *Provera kvaliteta i razumljivosti u sistemu za konverziju teksta u srpski govor*, Zbornik radova XXXIX konferencije za ETRAN, sv. II, Beograd, 1995, 485-488.
2. Ivić P., I. Lehiste, *Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku*, ZbFL VI, Novi Sad, 1963, 31-80.
3. Ivić P., I. Lehiste, *Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku*, ZbFL VIII, Novi Sad, 1965, 75-129.
4. Ivić P., I. Lehiste, *Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku*, ZbFL X, Novi Sad, 1967, 55-94.
5. Jokanović-Mihajlov J., *Priroda uzlaznih akcenata u progresivnijim štokavskim govorima*, SDZb XXIX, Boograd, 1983, 295-338.
6. Lehiste I., P. Ivić, *Prozodija reči i rečenice u srpskohrvatskom jeziku*, Izdavačka knjižarnica Zorana Stojanovića, Sremski Karlovci-Novi Sad, 1996.
7. Peco A., P. Pravica, *O prirodi akcenata srpskohrvatskog jezika na osnovu eksperimentalnih istraživanja*, JF XXIX, sv. 1-2, Beograd, 1972, 195-242.
8. Peter S., T. Tomić, J. Jokanović-Mihajlov, Z. Kašić, M. Filipović, *Neki aspekti sinteze srpskohrvatskih glasova*, Zbornik radova sa XXX jugoslovenske konferencije ETAN-a, sv. VI, Beograd, 1986, 67-72.
9. Peter S., Z. Kašić, T. Tomić, K. Simić, *Promena F_0 izolovane reči pod uticajem rečenične melodije*, Zbornik radova sa XXXVI konferencije ETAN-a, Beograd, 1992, 1l-17.
10. Sokolović M., *Promena osnovne frekvencije u jednosložnim rečima sa dugosilaznim akcentom*, Zbornik radova sa XLI konferencije ETRAN-a, sv. II, Beograd, 1997, 567-569.
11. Tomić T., S. Peter, M. Filipović, *Program za sintezu glasova u srpskohrvatskom jeziku*, Zbornik radova sa XXX jugoslovenske konferencije ETAN-a, sv. VI, 1986, 59-65.

PROMENA OSNOVNE FREKVENCIJE U DVOSLOŽNIM REČIMA SA DUGOUZLAZNIM AKCENTOM

Mirjana Sokolović

U radu se analizira promena osnovne frekvencije u dvosložnim rečima sa dugouzlaznim akcentom. Rezultati eksperimenta pokazuju da se kod svih ispitanika dugouzlazni akcenat realizuje na sličan način. Na osnovu prosečnih rezultata za četiri muška govornika daje se predlog promene osnovne frekvencije u ovim rečima koju treba ugraditi u uređaje za sintezu govora pomoću računara.