THE WORD TONES OF SERBO-CROATIAN AN INSTRUMENTAL STUDY.

Per Jacobsen

This study describes the results of an instrumental investigation made at the Institute of Phonetics of the University of Copenhagen in the spring of 1967. The purpose of the investigation was to examine the phonetic character of the Serbo-Croatian word tones, i.e. to measure vowel quantities, fundamental frequency, and intensity.

By means of the commutation test, 4 contrasting types of accent are established. These 4 types are traditionally called: long falling (usually symbolized with (^))[v:], short falling (^) [v], long rising (/) [v:], and short rising (^) [v]. In syllables after the ictus only the contrast long:short is phonemic, ([v:]:[v]). Although the results of this investigation show that the traditional names of the word tones do not always correspond to the phonetic reality, they are used here to label the four types of word tones. The distribution of the four accents is limited: Falling accents occur only on the first syllable. Rising accents may occur on any syllable except on the last. Monosyllabics have falling accents.

In his book <u>Osnovi Fonetike Srpskog Jezika</u> Branko Miletić describes the Serbo-Croatian word tones as follows:

Melody. Accent (N) [V] occurs in two main types: 1) One falling, in the central dialects (Bosnia, Hercegovina, Mačva, Dubrovnik, and other regions) and 2) one level, in Belgrade and other eastern regions. The first type (the falling) is more typical and probably older.

Accent (^) [V:] is <u>falling</u> in all dialects, but in such a way that the tone ordinarily forms an arch: it rises in the beginning and then falls steeply towards the end; this fall is greater or smaller, depending on the dialect, but it is a constant feature.

Accent (*) [V] is rising in all dialects; it is either rising from the beginning to the end or partly rising, so that part of the tone is level.

Accent (/) [V:] is the most stable tone in all dialects: it is always <u>rising</u>. There are, however, dialect differences: in certain dialects the tone is more rising than in others.

INTENSITY. The falling accents have parallel tone and intensity movements: The intensity rises in the beginning or stays constant for some time and then falls. On the contrary there is no such complete parallelism between tone and intensity in the rising accents: the intensity rises only in the beginning of the vowel and then falls towards the end.

QUANTITY. Accentuated long vowels are twice as long as accentuated short vowels, and accentuated vowels are usually longer than unaccentuated ones. Although "[V] and [V] are short accents and [V:] and [V:] are long, there are some differences between them: [V:] is usually somewhat longer than [V:], while the typical [V] is longer than [V], often considerably longer. -

This description covers the traditional conception of the phonetic nature of the Serbo-Croatian word tones.

Examples of tone and quantity contrasts:

sedi	['s&:di:]	'he paints grey'
sédi	['sɛ́:di]	'grey'
sédi	['s&:di:]	'his hair is getting grey'
sèdī	['sédi:]	he sits down!
sèdi	['sédi]	'sit down!'
kiika	['kůka]	'a hook'
kuka	['kûka:]	'he laments'
kùka	['kúka]	of the hip!
Rósa	['rí:sa]	the girl's name Rosa!
ròsa	[rjsa]	'dew'

Material, informants, and recording.

A number of monosyllables and disyllables were selected for the investigation. As the physical duration usually depends on vowel quality, the material is arranged so that long a is compared with short a, long e with short e, etc. In the selection of the examples various factors, which might influence the duration or tone of the accentuated vowels, have been considered. Thus the following consonant might play a role for vowel duration. Therefore, the vowels have been analysed when preceding different consonants: nasal/liquid, voiced and unvoiced fricatives, and plosives. Vowel duration might be influenced also by the quality of the following syllable.

For this reason, the disyllabic material consists of both long and short second syllables.

The material can be arranged as follows:

Table 1 a. Monosyllables

Following	Nasal/	Nasal/ Fricative liquid		Plosive		
consonant:	riquia	unvoiced	voiced	unvoiced	voiced	
long fall-	sam	pas	paž	sat	s කිd	
short falling	sần	pås	säv	čák	b <i>g</i> s	

Table 1 b. Disyllables

				1				
Intervocalic		Liquid	d/	Fricative				
consonant:		nasal		unvoi	ced	voice	d	
Tone label:			fall-					
		ing	ing	ing	ing	ing	ing	
lst syllable	2nd syllable short: 2nd syllable	čari	táman	spâsa	kásu	kažu	fáza	
10116	long:	para	sámīm	pasa	páse	kâže	fáza	
lst syllable short		pàra	pàra	päša	pàša	päzi	tàvan	
21101.0	2nd syllable long:	pamet	tàman	päšē	pàsūlj	päzi	fàzan	

Table 1 b. Disyllables - continued

Intervocalic		Plosive						
consonant		unvoiced		voiced				
Tone label:	falling	rising	falling	rising				
lst syllable	2nd syllable short: 2nd syllable long:	kāpi skāčē	pápa pácov	tåbor sädim	sábor nádā			
lst syllable short	2nd syllable short: 2nd syllable long:	šåpat påköst	šàtor kàpūt	Sabac pada	sàda kàdet			

Similarly for each of the syllabic sounds a, e, i, o, u, r.

The tone usually starts at a lower point when the initial consonant is voiced than when it is unvoiced. In order to get as uniform a material as possible the majority of the examples have an unvoiced initial consonant, whereas a smaller number, for comparison, have a voiced initial consonant.

At last the structure of the stressed syllable is considered. All disyllables are of the type sada or Sabac, i.e. with open first syllable. *

A number of the words investigated (16 monosyllables and 132 disyllables) form minimal pairs in which tone and/or length determine the meaning.

The words were written on cards, one word on each card, and arranged in quasi-random order, the only restriction being that minimal pairs did not occur adjacent to each other. The words were placed in a frame sentence: piše...na karti ('...is written on the card'). By this method a complete uniformity in the placement of the words in a sentence context is obtained, and the sentence intonation does not, at the place in question, disturb the word tone too much. The utterances were recorded in a silent room on a professional tape recorder at the Institute of Phonetics.

Four informants participated in the investigation:
MA (female) was born in 1938 in Sremski Karlovci in Vojvodina. She has a university degree in Serbo-Croatian language and literature.
Her material consists of 47 monosyllables and 242 disyllables, 289 utterances in all.

CJ (female), born in 1930 in Valjevo in Serbia. Her material consists of 43 monosyllables and 205 disyllables, 248 utterances in all.

SM (male), born in Sabac in Serbia in 1941. Material: 48 monosyllables and 196 disyllables, 244 utterances in all.

DA (male) born in Zlarina near Sibenik in Dalmatia and grown up in Bosnia and Vojvodina. His material consists of 37 monosyllables and 164 disyllables, 201 utterances in all.

The complete material consists of 982 utterances.

^{*} The material was intended to consist of 60 mono- and 240 disyllables, a number I did not reach. OJ, MA, and SM have spoken more than 240 disyllables, but this is because there are doublets of several examples.

ently distinguished by the speakers the minimally contrastive words were rerecorded on a special tape where they were placed immediately after one another. The informants were then asked to identify the utterances. This listening test showed that all 4 persons distinguished between the four types of accent, easily and without making errors. On the other hand they did not in all cases distinguish long and short unaccented vowels. The only exception was the short a of nom. sing. fem. in contradistinction to the long a of gen.plur. but here, too, there was some vacillation.

Therefore, I have desisted from any conclusion concerning the influence of the subsequent syllable on the quantity of the stressed syllable. In the calculations on vowel quantities I have combined disyllables assumed to have long and short second syllables into one group. If no phonemic distinction is found between long and short unstressed vowels the possible differences in duration are too accidental to be considered in the calculations.

Recording of acoustic curves.

The utterances were subjected to acoustic analysis by means of an intensity meter (KTH type) and Frøkjær-Jensen's pitch meter and recorded on the mingograph. Four synchronous traces were recorded on the ink writer: an intonation curve, a logarithmic intensity curve with an integration of 2,5 ms for female voices and 5 ms for male voices and a high-pass filtering (-3dB at 500 cps) for all voices. The high-pass filtering in connection with the logarithmic scale conditions that the consonants stand out rather distinctly. The third curve is an intensity curve (linear scale) with an integration of 5 and lo ms, respectively, and with linear frequency response. The fourth curve is a 'duplex oscillogram', which is a combination of an ordinary oscillogram and a high-pass filtered intensity curve.

The logarithmic intensity curve and the duplex oscillogram were used to delimitate the sounds before measurements were made on the intensition curve and the linear intensity curve. The speed of the paper was loo mm/sec, (one mm corresponding to 1/loo sec).

The mingograph recordings were made by cand.art. Hans Peter Jørgensen of the Institute.

Table 2. Vowel duration in cs.

_			1.0					
2	à [à:]	e [E:]		:]	: 1 G	[ù:]	£ [r̀:]	
MA	20,7	20,4	19,2	_			20,0	
OJ	21,0	- 1					17,3	
SM		21,4					17,0	
DA		18,1				3,7	_	
average	22,4	20,1					18,1	
average								
	a [a]	e [8]	î [ì]] % [ð] ü	[ù]	[f] ¥	
MA	13,3	14,9	10,1	12,	3 12	, 5	12,0	
OJ	13,0	11,0	10,0	11,	1 10	,6	9,8	
SM	12,6	11,7	8,7	10,	9 9	,6	7,7	
DA	14,5	12,5	10,3	11,	2 11	,3	11,0	
average	13,4	12,5	9,8	11,	4 11	, 0	10,1	
	AND							
2		syllable						
	â [à:]	á [á:]	ê [E:]	é [É:]	î [ì:]	í [í:] 8 [3:]	δ [J:]
MA	22,4	21,9	20,8	21,4	18,8	18,8	20,9	20,6
OJ	21,5	21,7	17,5	20,3	16,0	17,1	18,6	20,0
SM	24,4	20,9	21,0	21,1	18,2	18,5	19,8	19,5
DA							21,7	21,4
average	22,6	21,1	19,3	21,2	17,6	17,9	20,3	20,4
	û [ù:]	ú [ú:]	r̂ [r:]	ŕ [ŕ:]	à [à]	à [á]	[3] °	è [٤]
MA	20,3	_	_	_		_		14,9
ОЈ	18,0	16,8	16,3	18,7	12,7	11,8	11,0	11,6
SM	20,5	19,3	21,3	18,8	11,6	10,6	10,9	11,1
DA	19,8	19,8	16,5	19,5	13,2	12,8	12,0	11,4
average	19,7	19,0	18,6	19,3	13,1	12,5	12,3	12,3
	i [1]	ì [1] à	6 [6]	٥ [6]	ù [ù]	ù [ú]	* [+]	ř [ŕ]
MA	11,3		14,4	14,8	11,3	12,1	12,0	11,5
OJ	9,2	7,3	11,6	11,5	9,2	9,0	10,1	10,3
SM	9,4	7,7	11,5	9,8	9,8	10,8	9,5	11,6
DA	9,0		11,2	12,3	9,9	9,1		9,0
average	9,8		12,2	12,1	10,0	10,2	10,2	10,6

Average vowel duration in cs for four informants:

monosyll.
$$\hat{\mathbf{v}}:$$
 19,9 monosyll. $\hat{\mathbf{v}}:$ 19,8 monosyll. $\hat{\mathbf{v}}:$ 19,9 monosyll. $\hat{\mathbf{v}$

Significance testing is not necessary to show that the differences between monosyllables and disyllables are not significant.

The influence of the following consonant on vowel length is seen in the following table:

ľа	b 1	e	4
-	THE REAL PROPERTY.	-	-

ACCUMENTATION OF THE OWNER OF THE OWNER OF THE OWNER OF THE OWNER	nasal/liquid	unvoiced fricat	voiced ive	unvoiced	voiced sive
long vowel	20,4	20,2	20,2	18,6	19,7
short vowel	11,0	12,5	11,9	9,9	10,6

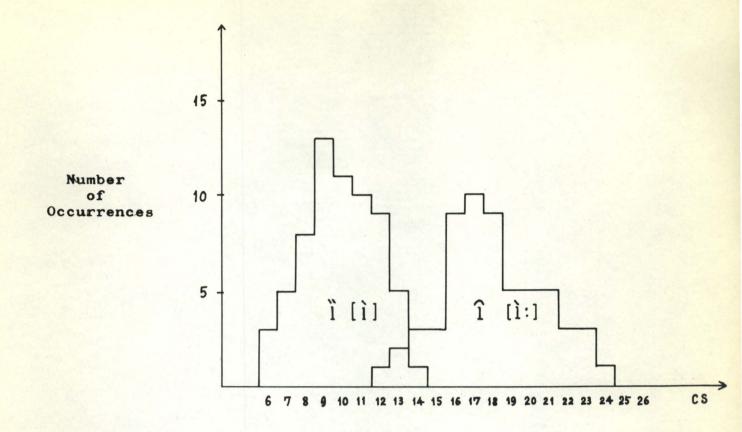
As seen in table 3 there is no significant difference between monosyllables and disyllables; nor is there a significant difference between rising and falling wordtones in this respect. This result disagrees with the statement in Miletic's Osnovi fonetike.

The following consonant influences vowel length, vowels followed by plosives being somewhat shorter than vowels followed by fricatives, nasals or liquids, and vowels followed by unvoiced plosives being shorter than vowels followed by voiced plosives.

The material is clearly divided into long and short vowels. The overlapping is minimal. A diagram of the dispersion of the vowel <u>i</u> is given as an example (Fig. 1.).

Altogether the durations of the individual vowels are as follows:

^{*} Monosyllables and disyllables are taken together since the difference in duration is not significant



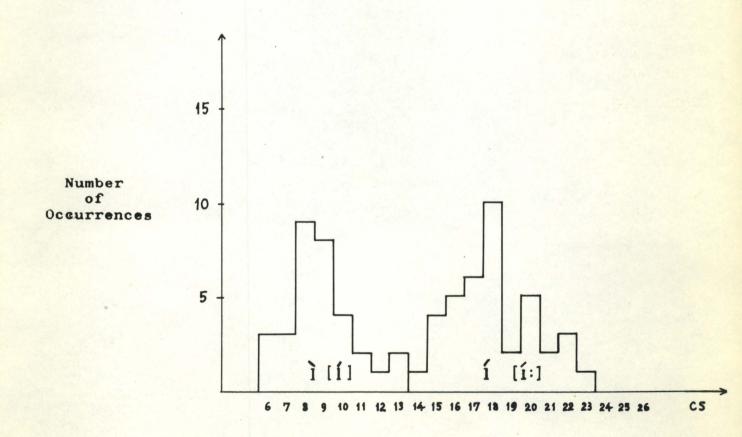


Fig. 1

Table 5. Vowel durations.

long vowels:

i: 18,2 r: 18,7 u: 19,2

e: 20,3 o: 20,4

a: 22,1

short vowels:

i 9,4 r 10,3 u 10,4

e 12,4 o 11,9

a 13,0

- which was to be expected, the closed vowels being physically shorter than the more open ones.

Table 6.

Average fundamental frequencies (beginning, peak, and end)-of the syllabic sounds. The distance of the fundamental frequency peak (if present) from the beginning of the syllabic sound is given as a percentage of the total duration.

Monosyll./^/ [V:] with unvoiced with voiced initial consonant initial consonant

RE-THEN STATEMENT	TIITUTAL						
	1st syll- able	%	2nd syll- able	lst syll= able	%	2nd s	yllable
MA	237-259-143	30		212-270-155	33		
OJ	222-289-149	24		213-262-154	26		
SM	148-165- 93	17		112 - 128 - 80	35		
DA	165-191-124	37		147-178-117	30		
Disy	√11./^/ [v̀:]						
MA	215-249-169	38	159- 140	198-239-169	48	164-	-142
OJ	221-282-191	31	173128	218-273-207	50	171-	-122
SM	145-171- 90	29	90 81	121-145-101	43	92-	- 82
DA	177-192-132	34	125117	173-198-130	37	128-	-118
Mond	syll./\\/ [v]						
MA	218-246-190	45		193-232-214	46		
OJ	223-296-232	41		218-288-254	50		
SM	134-158-122	46		113113			
DA	170-188-153	51		146-183-139	60		

^{(%:} Place of peak in percent.)

Table 6. - continued.

Disy	711./\\/ ini	C	ith nsor	unvoic nant	ed	with	voiced	init	ial co	nsonant
	1st s	y11- b1e	%	2nd sy ab1		1st s	y11-	%	2nd s	yllable
MA	205-2	35-208	66	175-	-136	191-2	42-235	74	175-	-138
OJ	221-2	91-258	55	182-	-129	211-2	69-259	71	174-	-117
SM	145-1	81-143	46	99-	- 90	132-	-144		103-	- 89
DA	176-1	98 -1 62	53	132-	-113	164-1	82-173	61	149-	-116
	1.1	[Ý:]	(on1	y disy	llabic)				
MA	185-	-204		223-	-187	177-	-210		235-	- 198
OJ	194-	-260		260-	-171	182-	-236		248-	-162
SM	120-	-143		139-	- 96	103-	-140		137-	- 86
DA	140-	-167		174-	- 136	131-	-164		175-	-139
	/ 1	[Ý] (only	disyl	labic)					
MA	189-	-191		220-	-173	183-	-206		228-	-190
OJ	199-	-248		261-	-162	191-	-237		255-	-178
SM	129-	-126		148-	-105	113-	-128		141-	-105
DA	147-	-159		182-	-142	140-	-163		191-	-146

^{(%:} Place of peak in percent)

Description of frequency

Although the averages presented here show a difference in the stressed vowels between short "falling" and "rising" a large part of the examples show that the tone of the stressed vowel is not always relevant for the opposition. (Cf. illustration Fig. 4b of Zenī [3 eni:]*) and Zeni [3 eni]). The opposition cannot be regarded as an opposition between falling and rising tone in the stressed syllable. On the other hand it is evident and common for all four informants that the opposition between ''' ['] and '' ['] is established by the relation between the tone of the stressed syllable and that of the following syllable. The tone in the syllable after the ictus in words with '' ['] starts as high or even higher than the end of the preceding (stressed) syllable, while it is considerably lower in words with '" ['].

The movement of the tone is continued in voiced consonant after short vowel.

The tone in the stressed syllables in words with $/^{\circ}/[v:]$ and /'/[v:] is in itself sufficient to establish the opposition.

Table 7

Intensities in dB. - Averages for all four informants

Word tone		tone	1	. syll	able wel)	place of peak	2	2. syllable (vowel)			
			Beg.	Peak	End	per- centage	Beg.	Peak	End	per- centage	
nono.	-/^/	[v̂:]	41	50	34	29					
syll.		[v]	41	50	38	45					
di-	/^/	[v:]	41	51	41	39	41	-	36	-	
sy11.	1.4	[v]	45	51	41	52	43	47	38	28	
	1.1	[Ý:]	41	50	44	61	47	48	38	34	
	/ 1	[Ý]	41	49	43	58	46	49	38	33	

^{*)} or as vowel length after stressed syllable is not relevant: 2èni [zɛni].

Description of intensity.

The intensity usually rises suddenly in the beginning of the vowel. It may have its peak in the beginning and may then fall towards the end, or it may exhibit a break after the sudden rise, rise slightly towards the end and then suddenly fall just before the end of the vowel. When the peak is situated in the first half of the vowel the intensity is falling, and when it is in the second half the intensity is rising.

Although the table shows falling intensity with // [':] (less evident with / '/ [']) and rising with / '/ [':] and / '/ ['], the deviation in the individual occurrences is so large that one may speak of a tendency only.

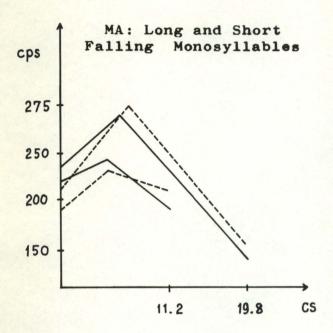
On the other hand there seems to be a difference in the relation between the syllables. In words with / [':] and / ' / ['] the intensity of the second syllable is lower than the intensity of the first syllable, while the second syllable in words with / ' / [':] and / ' / ['] has (almost) the same intensity as the first syllable. Regarded in this way it is correct to underline the parallelism between tone and intensity.

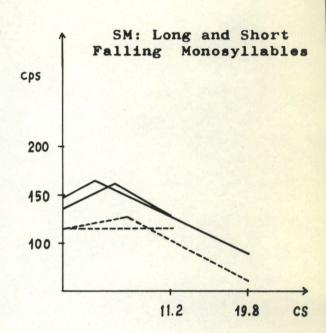
Conclusion

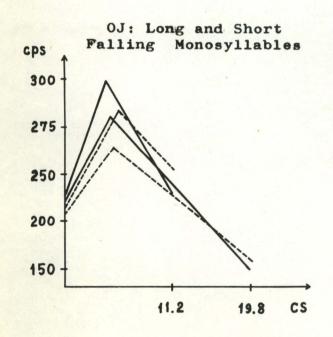
On the basis of the present material I cannot share the traditional conception of the Serbocroatian accent system. In short vowels it seems impossible to locate the distinctive characteristics of word tone within the syllable traditionally said to carry the accent. In long vowels, however, the usual assumption seems to be true (cf. Miletić).

Whether intensity plays any role for the identification is difficult to tell. Experiments with synthetic speech where the different parameters of speech signal can be arbitrarily changed, might throw light on this question.

_____ Unvoiced Initial Consonant







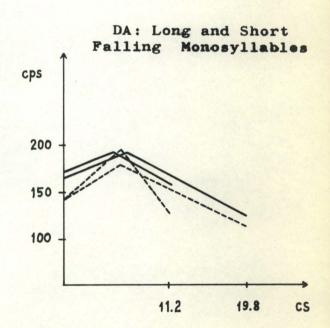
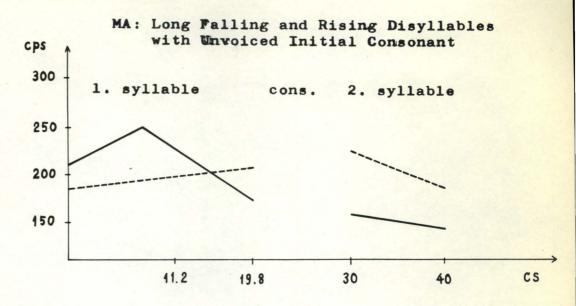
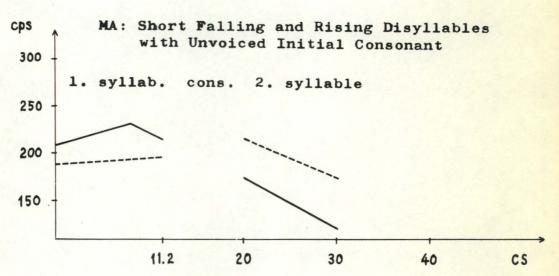


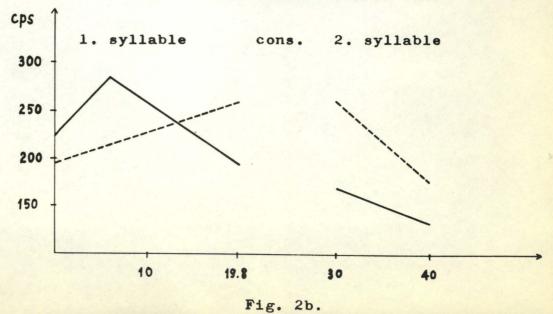
Fig. 2a.
Fundamental Frequency.

Falling Rising

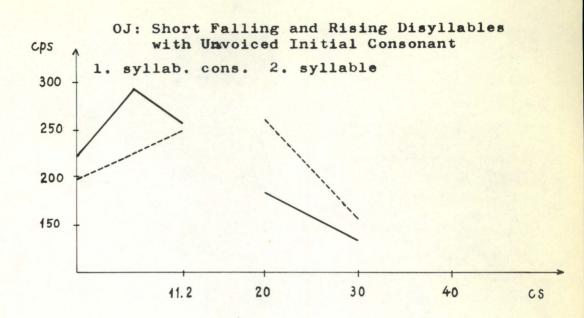


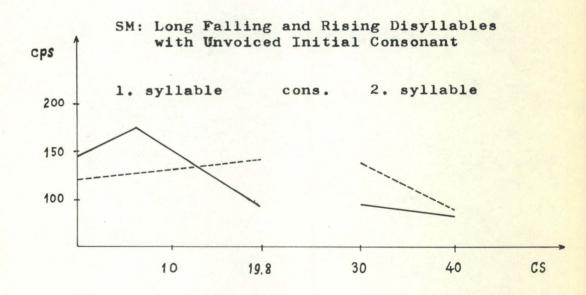


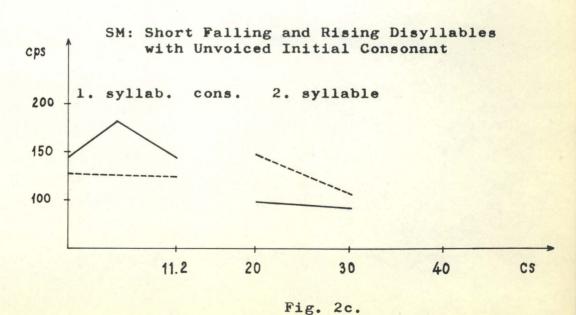
OJ: Long Falling and Rising Disyllables with Unvoiced Initial Consonant



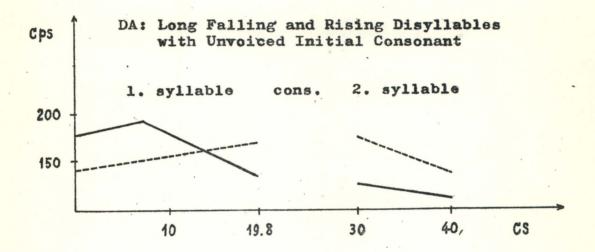
----- Falling
----- Rising

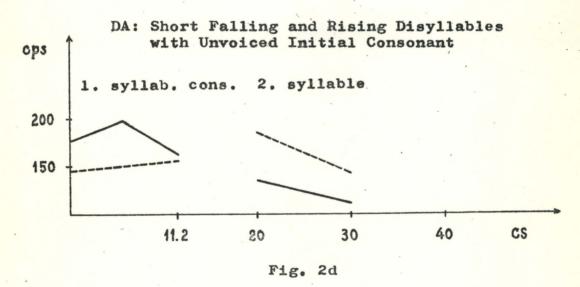






---- Falling





The diagrams above are graphic illustrations of the measured values given in Table 6. The short vowel duration shown here represents the average of all short vowels, i.e. 11.2 cs. The long vowel duration represents the average of all long vowels, i.e. 19.8 cs. The consonants and the second syllables which have not been measured are arbitrarily set to 10 cs each (the consonants in long disyllables to 10.2 cs).

In monosyllables —— symbolizes long and short vowels preceded by an unvoiced initial consonant, while ---- symbolizes vowels preceded by a voiced initial consonant. In disyllables —— symbolizes falling, ----- rising word tone.

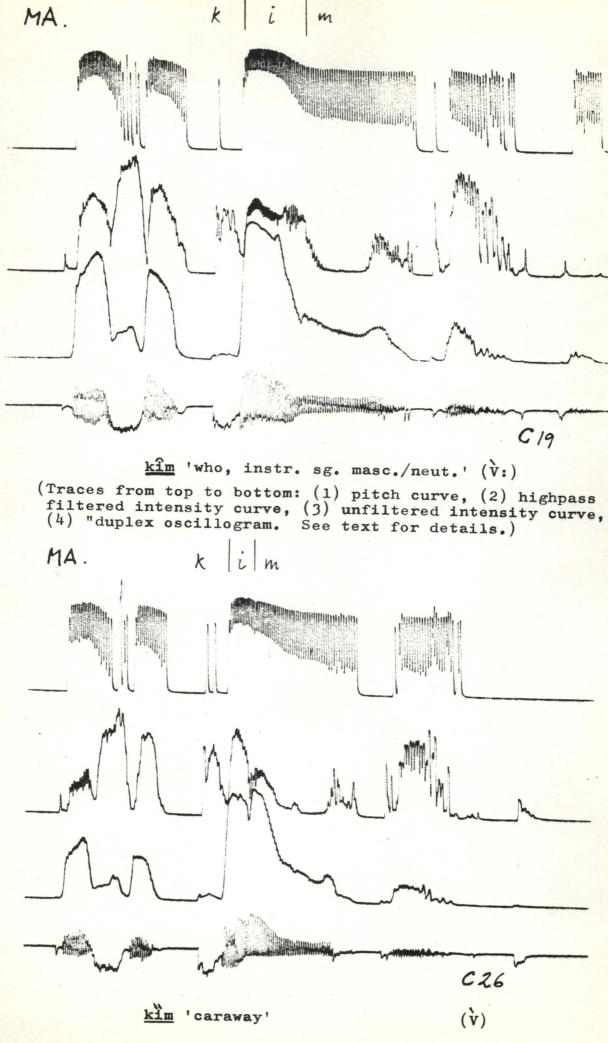


Fig. 3
Mingograms of monosyllables.

