Computerized one-third octave band analysis in real time of the structural relationship vowel consonant in a monosyllabic word

SVETLANA ČIČEVIĆ
JOVAN DAVIDOVIĆ
SMILJKA VASIĆ
VLADISLAVA KNAFLIĆ
ĐURĐE NAUMOVIĆ

Some vowels and consonants also were more influenced by the location in the word, its length and context of a verbal message than the others (Lieberman, Michaels 1962; Bloothooft, Plomp 1985). It refers, among others, to the vowel /E/ and consonant /P/, at least in Serbian language. This is due to changeable spectral position and frequency range of vowel /E/ formants, particularly that of F2. In the case of consonant /P/ it is attributed to its poor frequency content and low energy level. In addition, the acoustic structure of the vowel /E/ is very sensitive, even under normal condition, while the duration of voice /P/ is the shortest among consonants. Thus, the purpose of the present study was to provide additional insight into acoustic structural relationship between vowel-consonant combination in the monosyllabic word. For this reason the acoustic structure variation in monosyllable EP has been investigated under normal condition and during anticipation stress.

METHOD AND MATERIALS

Eleven experienced phoneticians of both sexes (aged 28-62 years) participated in the study as subject. They were seated quietly in a sound proof chamber and spoke, monosyllabic word EP two times successively, directly into the Brel and Kjaer analyzer via dynamic microphone. This procedure was repeated five times during the year, when subjects were in various moods, pronouncing the monosyllable EP, with normal or higher intensity.

The speech material was analysed by using real time Digital Frequency Analyzer Type 2131 (Brel and Kjaer, Denmark). The Analyzer was connected via interface to the Hewlett Packard Desk Top Calculator 9825B and HP 9872A Plotter.

RESULTS

One-third octave band analysis in real time of monosyllable EP, spoken under normal circumstances, reveals the particular structural relationship of these speech sounds. From the basic numerical data of the analysis, it is obvious that characteristic integral frequency-intensity-time of the vowel /E/ is predominant over plosive /P/. Usually the vowel /E/ spectral range is 50 Hz- 6.3 kHz, whereas that of plosive /P/ varies between 50 Hz and 3.15 kHz and may be as much as 10 kHz. In this case the overall EP sound pressure level has been 82.02 dB or 159.22 W/m2 if expressed in microwatts per square meter, while plosive /P/ reached 57.34 dB

or 0.54 W/m2. In other words, plosive /P/ makes up 0.31% of the overall EP energy level in spite of relatively long /P/ duration. The EP total speaking time was 528 ms, whereby the vowel /E/ constitutes 58.33% and the remaining time belongs to the plosive /P/. The occlusion period was about 44 ms although its level may sometimes exceedes the /P/ intensity.

At rest, the formant structure of both phonemes, within monosyllable EP, uttered at normal speaking rate, has been well defined and quantified. For instance, the sound pressure level of the vowel /E/ first formant was 78.36 dB, corresponding to the value of 68.54 W/m2, and F1 of the plosive /P/ accounts for 0.95 of the vowel /E/ first formant. This ratio for the /P/ second formant was 0.58%.

The acoustic structure of the monosyllable EP was markedly affected by the anticipation stress. Its frequency range extended up to 20 kHz. The first formant is located at 500 Hz and F2 at 3000 Hz (Fig. 2.). In comparison with the spacing presented in Fig.1., F1 – F2, shown in Fig. 2., has been greater amounting to 500 Hz – 3 kHz, instead of 630 Hz – 2 kHz. Despite such spacing the sound pressure level of F1 presented in Fig. 2. is higher than that shown in Fig.1. by more than 5 times.

Maximal instantaneous intensity was achieved within the first formant 132 ms after the beginning of the articulation. In the second formant the peak was reached 44 ms before than in F1 and had lower level. In contrast, during the anticipation stress the peak in F1 and F2 was formed 220 and 176 ms, respectively after the beginning of utterance. At the same time, fundamental frequency of the voice in the course of monosyllable EP production was 200 Hz in respect to resting condition when it was 160 Hz.

The utterance of the monosyllable EP has been prolonged by the anticipation stress. If EP overall articulation period is assumed to be 100 %, the voice /E/ duration may exceed 3/4 of its time, and the remaining time belongs to plosive /P/. In spite of relatively long /P/ duration, its energy contribution to overall sound pressure level was small, being between 0.16 and 1.17 %. In relaxed state it was still lower.

The difference between the word EP spoken in a low and that uttered in intense voice is shown in Fig. 3. In both cases the plosive /P/ structure, in terms of acoustic energy level, duration, peak shift and general structural alteration was more influenced than that of vowel /E/.

DISCUSSION.

The voice /E/ acoustic structure is one of the most sensitive among Serbian vowels and presumably among those of other languages. In addition, its structure change is remarkably affected by the vicinity of certain consonants. One of them is plosive /P/. Their mutual influence, among other, depends upon the plosive /P/ location in the word. When the voice /P/ is located immediately after the vowel /E/, as is the case in monosyllabic word EP, it is considered as the most complete regarding its acoustic structure and duration. Unfortunately, studies on modification of voice /E/ and /P/ in EP combination during stressed state are scarce(Ruiz *et al.* 1990; Scherer, 1981; Streeter *et al.* 1983). Therefore, our results indicates that three-dimensional characteristics are better indicators of the particular vowel-consonant relationship than those previously obtained by two-dimensional measurements.

CONCLUSION.

The stuctural relationship between the vowel /E/ and plosive /P/ in the monosyllabic word EP is changed by the anticipation stress. It caused widening of F1 – F2 spacing, shifted the formant spectral position and energy peak, increased the sound pressure level and prolonged the time of the word utterance.

REFERENCES

Bloothooft, G.; Plomp, R. 1985. Spectral analysis of sung vowels. II. The effect of fundamental frequency on vowel spectra. *The Journal of the Acoustical Society of America. 77.* no 4. 1580-1588.

Lieberman, P.; Michaels, S. B. 1962. Some aspects of fundamental frequency andenvelope amplitude as related to the emotional content of speech. *The Journal of Acoustical Society of America. 33,* 597-603.

Lieberman, P. 19 . *Speech Physiology and Acoustic Fhonetics: An introduction.* New York, London, Macmillan Publishing Co., Inc.

Ruiz, R., Legros, C., Guell, A. 1990. Voice Analysis to Predict the Psychological or Physical State of a Speaker. *Aviation, Space, and Enviromental Medicine, 61,* 266- 271.

Scherer, K. R. 1981. Vocal indicators of stress. In J. K. Darby eds. *Speech evaluaiton in psychiatry.* New York, Grune, Stratton.

Streeter, L. A.; Mc Donald. N. H.; Apple, W.; Krauss, R. M.; Galotti, K. M. 1983. Acoustic and perceptual indicators of emotional stress. *The Journal of Acoustical Society of America.73,* 1354-60.

Figure 1 – At rest produced acoustic structure of the monosyllabic word EP, obtained by one-third octave band analisys in real time

Time (ms)	63Hz	80	100	125	160	200	250	315	400	500	630	800	1 kHz	1.25	1.6	2	2.5	3.15	4	5	6.3	8	10	12.5	16kHz	Frequency (Hz)
	44.96	46.23	51.81	63.94	66.99	51.10	65.20	67.60	62.64	76.69	78.40	61.66	55.79	55.80	52.59	72.08	66.04	54.18	56.11	57.01	46.86	44.89	45.48	40.24	35.01	→ 82.02 db
44		33					32																			35.53
88	32	31			32	46	53	51	54	68	63	48	45	45	49	62	55	44	46	46	36	34	34	E		70.47
132	35	39	47	58	56	46	61	60	54	74	74	57	52	52	56	69	62	51	53	54	44	42	42	37	33	78.14
176	36	40	42	59	61	41	59	64	59	70	75	58	51	51	55	67	62	49	51	52	41	44	41	36	33	77.48
220	34	35	40	58	64	44	59	62	56	60	69	53	46	46	49	61	56	43	45	46	35	34	35	32		73.00
264	36	36	47	55	58	39	53	56	50	60	63	47	40	40	43	54	50	37	39	40						67.12
308	39	35	42	49	52	34	47	50	44	53	57	41	34	34	37	48	44	31	34	34	OCCLUSION					60.96
352	35		36	43	45	31	41	43	38	47	50	35			32	42	37									54.33
396	34	34	32	37	40		35	37	36	43	45	37	39	39	38	40	39	36	31	32	34		31			51.37
440	31	33	32	35	37		34	41	39	42	42	32	35	35	35	36	34	31			31			P		49.27
484	31	36	32	32	33		31	35	34	37	37			32	32	32										44.29
528		33	31							32	32															38.07
					Fo			Pr			F_1					F_2										Sound pressure level in dB re 10^{-12} W/m²

Figure 2 – Acoustic structure of the monosyllabic word EP produced under anticipation stress obtained by one-third octave band analisys in real time

Time (ms)	50Hz	63	80	100	125	160	200	250	315	400	500	630	800	1 kHz	1.25	1.6	2	3	3.15	4	5	6.3	8	10	12.5	16	20kHz	Frequency (Hz)
	49.22	50.78	58.21	65.99	68.80	71.80	73.97	69.69	75.50	79.93	85.41	78.38	62.91	62.33	58.11	59.26	72.36	75.63	63.54	61.19	59.76	53.54	57.97	52.90	47.85	45.92	43.01	→ 88.30 db ←
44		34				32																						36.12
88							52	50	58	58	69	54	46	43	41	44	59	56	45	44	41	34	35	34		31	31	70.44
132		32	37	35	43	64	62	50	68	69	75	68	54	53	48	50	66	65	55	50	48	41	44	43	37	35	34	78.11
176	35	31	35	32	39	62	69	47	64	74	70	73	55	56	50	52	66	68	56	53	50	47	51	45	40	38	37	78.99
220	35	35	32	33	38	58	68	46	64	75	75	74	57	57	51	52	65	69	57	54	51	47	53	47	42	39	36	80.47
264	33	31		31	43	65	67	49	71	71	82	70	56	55	52	53	64	69	56	54	52	46	51	46	41	39	35	83.31
308	34			38	57	67	61	66	68	70	78	67	54	52	50	50	62	68	56	55	55	46	50	45	40	37	34	80.20
352	32			54	64	62	62	65	64	68	75	62	50	48	46	46	59	65	52	51	51	43	45	41	37	37	32	77.46
396		31	41	63	61	56	59	61	62	65	70	57	45	43	42	42	55	60	46	45	45	38	40	35	34	34		73.52
440	35	34	53	61	65	53	54	56	56	59	64	50	39	37	36	36	48	54	40	39	39	32	34				E	69.84
484	37	35	55	55	49	46	48	50	50	53	57	44	33	32		31	42	48	34	34	34							62.64
528	39	45	49	50	43	41	42	43	44	45	51	38		43			36	42					OCCLUSION					57.26
572	39	43	44	44	37	35	36	37	37	40	45	32		38				35									P	51.84
616	35	38	38	38	31		34	41	45	38	46	42	47	33	43	44	45	46	38	35	34	33		33				55.00
660	32	33	38	38	34		Fo	35	39	34	42	38	36		38	41	40	40	33	31	31	31	33	33				50.29
704	34	40	37	34	32			31	34	31	37	33	32		33	35	35	35					31					46.81
748	43	43	33	31							32					31		F_2										46.62
792	43	37									F_1																	43.97
																												Sound pressure level in dB re 10^{-12} W/m²

Figure 3 – Strongly and weakly pronounced monosyllabe EP