ANN.Bull.RILP(1990)No.24, 235-241

A QUANTITATIVE EVALUATION OF DYSARTHRIC SPEECH USING ACOUSTIC ANALYSIS

Ghen Ohyama*, Hajime Hirose, and Yoko Watanabe

We have been engaging in a study of the quantitative evaluation of dysarthric speech using speech analysis method. In our previous report[1], we reported the result of an analysis of mimicked dysarthric speech produced by a speech therapist who read test words by mimicking dysarthric speech of a spastic type with three different degrees of severity, i.e. slight, moderate and severe. In this paper, we propose several new parameters for evaluating the dysarthric speech and give the results of the evaluation of speech samples uttered by normal and pathological subjects.

1. Collection of speech samples

1.1 Subjects

The subjects were 4 normal subjects and 10 patients with spastic dysarthria due to pseudobulbar palsy (PBP) secondary to a cerebrovascular accident (PBP group). The normal subjects were university students. Table 1 shows the sex, age and classification of the patients based on clinical evaluation, including a perceptual judgment of their speech. The perceptual judgment was made using our diagnostic method[2], and the patients were classified into three severity groups accordingly.

1.2 Speech samples

Each subject was requested to utter two-mora nonsense words of the form /VCV/, where /V/ was /a/ and /C/ was one of the six plosives /p/, /t/, /k/, /b/, /d/, /d/, /g/. They uttered the test words three times each.

Recordings were made in a quiet room of a hospital for the PBP group, and in a quiet room of a school for the control group. Speech samples were recorded on a digital audio tape recorder (DAT).

2. Method of acoustic analysis

The recorded speech samples were low-pass filtered at the cut-off frequency of 9 kHz and were digitized with a sampling frequency of 20 kHz. The samples were then analyzed using LPC

*Kansai Advances Research Center Communications Research Laboratory Ministry of Posts and Telecommunications

analysis and Cepstrum analysis. The type of window was the Hamming and its length was 20 msec, with a frame period of 10 msec.

In the present study, the following parameters were obtained for every 10msec.

a. power difference
b. power difference in long periods
c. difference in power between vowel segments and consonant segments
d. spectral difference
e. peak value in cepstrum
f. formant frequency

3. Definition and meaning of parameters

3.1 Power difference

Power difference is defined as the difference in power in dB between the current frame and the previous frame. This parameter was used for evaluating the quick power change.

3.2 Power difference in long periods

This parameter is the angle of the least square error line among the values of the power in dB for 5 consecutive frames. This parameter was used for evaluating the power change over the period of 50 msec.

3.3. Difference in power between vowel segments and consonant segments

This parameter was defined as the difference between the averaged power in dB of vowel segments and that of voiceless stops. This parameter was used for evaluating the voicelessness.

3.4. Spectral difference

This parameter was defined as the spectral distance between the current frame and the preceding frame calculated from the LPC cepstrum values. This parameter was used for evaluating the quick change of the spectrum.

3.5 Peak value in cepstrum

The peak value in cepstrum which is dependent on periodicity was obtained. This parameter was used for evaluating the degree of the periodicity of the speech signals in the voiceless consonants.

3.6 Formant frequency

The first and the second formant frequencies were estimated from the LPC coefficients. This parameter was used for evaluating the articulation of the vowels.

4. Results

4.1 Power difference

Tables 2-A, B show the results of the power difference for the speech samples obtained from the control group and the PBP group, respectively. It can be seen that the values are larger in the control group than in the PBP group. The difference is most marked for /p/, /t/ and /k/, while it is not so significant for /g/. In the PBP group, the value is larger for mild cases than for severe cases.

4.2 Power difference in long periods

Tables 3-A, B show values of the power difference in long periods obtained from the control group and the PBP group, respectively. A tendency similar to that of power difference above is seen in this parameter. Namely, the value is larger in the control group than in the PBP group, especially for /p/, /t/ and /k/, while the difference is not so marked for /g/. The value is larger for mild cases than for severe cases in the PBP group.

4.3 Difference in power between and consonant portions

Tables 4-A, B show the values of the difference in power between vowel and consonant segments for the two groups of control and PBP, respectively. In these tables two values are presented for each case. Namely, the upper value indicates the power difference between the preceding vowel and consonant; whereas the lower value indicates that between the following vowel and consonant. In general, the difference value is larger in the control group. The value is especially small in subjects 6 and 8 in the PBP group, in which /apa/ sounded like /aba/.

4.4 Spectral difference

Tables 5-A, B show the values of the spectral difference in the two groups of control and PBP, respectively. The value is larger in the control group than in the PBP group for /p/ and /t/, while the difference is not so marked for /k/, /b/, /d/ and /g/.

4.5 Peak value in cepstrum

Tables 6-A, B show the peak value in cepstrum for the consonant segments of the test words consisting of vowel and voiceless stops. When the value is large, it indicates that the voiceless segment is replaced by periodical signals. For example, the value was large for /t/ and /k/ in subject 5 of the PBP group, for which the samples /ata/ and /aka/ sounded like /ada/

and /aga/.

4.6 Formant frequency

Table 7 shows the first and second formant frequencies of the vowel portion in the PBP group. It can be seen that the values are out of the standard range for the vowel /a/. The recorded vowels here sounded like /e/ or /u/ instead.

6. Concluding remarks

In this paper, we proposed several acoustic parameters for a quantitative evaluation of dysarthric speech. It can be suggested that the parameters of power difference and the spectral difference are most useful for estimating the mobility of the articulators in speech production. The peak value of cepstrum and the difference in power between vowel and consonant segments can be used for evaluating the voicelessness of the consonantal segment. The formant frequency values are also useful for evaluating correct vowel articulation.

References

1) Hirose, H., Ohyama, G., Konno, K., "A Preliminary Study of the Acoustic Evaluation of Dysarthric Speech ", Ann. Bull. RILP No.22, pp.97-103 (1988)
2) Fukusako, Y., Monoi, H., Tatumi, I. F., Kumai, K., Hijikata, T. and Hirose, H.."Analysis of characteristics of dysarthric speech based on auditory impressions. Jpn. J. Logop. Phoniatr. No.24, pp.149-164, (1983)

Table 1. Selected patients and thier degree of severity.

subject	sex	age	degree of severity	articulation score		abnormality	intelligibility
				v	c		
sub 1	M	72	slight	0	0	0	1
sub 2	M	64	slight	0	1	1	1
sub 3	M	58	slight	0	1	1	1
sub 4	M	72	slight	0	1	1	1
sub 5	M	74	moderate	2	2	2	2
sub 6	M	63	moderate	0	1	1	2
sub 7	M	?	moderate	0	2	2	2
sub 8	M	73	severe	2	3	3	3
sub 9	F	?	severe	1	2	3	3
sub 10	M	38	severe	3	4	4	5

Table 2-A. Power difference in dB in the control group.

	P	T	K	B	D	G
SUB1	42.5	43.2	39.3	10.1	10.3	1.2
SUB2	30.3	35.4	31.2	25.3	14.1	2.7
SUB3	34.7	36.0	33.8	16.6	14.7	4.0
SUB4	21.7	15.6	16.4	14.1	16.5	9.1

Table 2-B. Power difference in dB in the PBP group.

	P	T	K	B	D	G
SUB 1	23.4	23.3	19.9	32.4	13.6	3.4
SUB 2	15.1	10.6	11.4	5.1	4.6	1.6
SUB 3	25.1	19.1	18.0	6.4	6.0	1.6
SUB 4	20.8	10.2	21.5	16.0	12.7	9.2
SUB 5	9.0	3.4	4.0	2.8	2.4	8.9
SUB 6	14.7	15.5	10.7	13.7	15.7	3.3
SUB 7	19.6	28.3	30.5	15.1	19.1	28.2
SUB 8	3.8	6.5	1.6	2.5	3.4	7.0
SUB 9	11.2	10.5	11.7	7.4	5.3	3.7
SUB10	3.9	1.4	13.1	3.2	1.3	1.2

Table 3-A. Power difference in long periods in dB in the control group.

	P	T	K	B	D	G
SUB 1	14.0	13.1	12.1	2.4	2.1	1.1
SUB 2	13.9	13.1	10.7	5.5	4.0	0.6
SUB 3	12.1	10.9	9.8	4.9	4.4	0.8

Table 3-B. Power difference in long periods in dB in the PBP group.

	P	T	K	B	D	G
SUB 1	8.96	8.09	5.71	5.52	4.06	0.68
SUB 2	6.79	6.10	6.05	2.28	2.00	0.40
SUB 3	6.92	6.68	5.56	1.90	1.45	0.66
SUB 4	6.60	4.58	4.47	6.11	3.53	0.62
SUB 5	1.11	1.26	0.46	1.24	0.77	1.51
SUB 6	6.43	5.59	4.22	6.04	5.12	1.66
SUB 7	9.09	7.90	8.70	7.46	5.20	5.54
SUB 8	0.52	0.72	0.73	0.69	0.78	0.53
SUB 9	6.93	5.53	4.28	1.83	2.75	2.03
SUB10	0.78	1.15	0.42	0.92	5.12	0.32

Table 4-A. Difference in power in dB between vwoel and consonant segments in the control group.

	P	T	K
SUB 1	47.5	44.5	45.3
	40.9	42.9	43.0
SUB 2	40.6	43.6	45.1
	32.9	39.6	36.2
SUB 3	41.1	42.5	42.4
	39.5	38.6	40.8
SUB 4	30.3	27.6	26.8
	34.8	34.5	33.2

Table 4-B. Difference in power in dB between vwoel and consonant segments in the PBP group.

	P	T	K
SUB 1	27.4	24.8	11.4
	24.1	25.3	10.4
SUB 2	17.4	15.5	18.4
	15.5	18.5	18.9
SUB 3	19.9	19.3	13.2
	13.1	15.1	4.7
SUB 4	17.9	12.9	12.3
	19.7	10.7	14.1
SUB 5	12.6	1.5	1.7
	13.5	-2.5	2.1
SUB 6	3.8	19.9	13.8
	-2.3	8.3	2.8
SUB 7	18.7	22.4	18.0
	21.7	19.1	25.2
SUB 8	1.3	9.6	3.9
	-1.0	8.8	2.4
SUB 9	19.1	20.9	22.2
	16.4	18.6	21.1
SUB10	3.8	1.3	1.8
	-14	-9.9	5.6

Table 5-A. Spectral difference in dB in the control group.

	P	T	K	B	D	G
SUB1	31.1	26.5	56.7	17.7	28.0	3.9
SUB2	12.8	26.3	23.8	8.4	13.9	2.3
SUB3	23.7	24.2	23.1	7.3	14.4	4.7
SUB4	3.2	8.4	5.4	3.4	5.6	3.0

Table 5-B. Spectral difference in dB in the PBP group.

	P	T	K	B	D	G
SUB 1	8.9	20.2	0.7	1.1	1.0	0.2
SUB 2	13.5	11.2	1.0	0.6	0.8	0.4
SUB 3	14.1	13.3	1.0	0.6	0.7	0.2
SUB 4	6.7	7.8	0.8	0.9	1.0	0.6
SUB 5	2.8	1.0	0.2	0.5	0.5	0.7
SUB 6	4.5	7.4	0.8	0.6	0.9	0.2
SUB 7	10.6	29.6	1.3	0.8	1.2	0.7
SUB 8	0.6	8.7	0.3	0.2	0.4	0.2
SUB 9	1.3	1.8	0.7	0.5	0.4	0.5
SUB10	1.5	0.5	0.8	0.5	0.4	1.0

Table 6-A. Peak value in cepstrum in the control group.

	P	T	K
SUB 1	455.0	477.0	194.1
SUB 2	487.2	287.7	495.1
SUB 3	420.6	280.2	471.0
SUB 4	438.6	461.6	484.3

Table 6-B. Peak value in cepstrum in the PBP group.

	P	T	K
SUB 1	422.7	224.3	477.7
SUB 2	412.8	563.3	529.4
SUB 3	519.4	543.8	501.9
SUB 4	464.1	360.5	606.6
SUB 5	444.3	1088.1	1162.9
SUB 6	319.1	416.3	485.7
SUB 7	381.3	529.6	583.2
SUB 8	472.7	507.3	1157.3
SUB 9	485.0	403.5	489.3
SUB10	1077.5	1162.6	1182.2

Table 7. First and second formant frequenciesx in Hz in the PBP group.

	P		T		K		B		D		G	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
SUB 1	840	1500	700	1510	760	1510	730	1480	680	1560	600	1600
SUB 2	640	1180	410	1340	400	1650	730	1180	650	1330	870	1410
SUB 3	700	1220	540	1450	510	1490	720	1230	670	1410	700	1370
SUB 4	650	1120	730	1270	630	1170	890	1190	670	1220	860	1240
SUB 5	780	1140	780	1470	800	1260	840	1140	770	1490	850	1270
SUB 6	500	920	650	2400	440	1460	410	1060	490	1430	490	1030
SUB 7	740	1130	650	1290	650	1320	720	970	840	1300	570	1370
SUB 8	640	1060	640	1350	370	1180	570	1020	550	1140	650	1170
SUB 9	710	1540	740	1830	650	1320	690	1400	760	1540	720	1760
SUB10	320	1110	410	1260	360	1160	380	1020	300	1350	330	1050