

## GRBAS Evaluation of Running Speech and Sustained Phonations

T. Sakata <sup>1)</sup>, N. Kubota <sup>2)</sup>, H.Yonekawa <sup>3)</sup>, S. Imaizumi <sup>1)</sup> and S. Niimi <sup>1)</sup>

- 1) Reserch Institute of Logopedics and Phoniatrics, University of Tokyo
- 2) Department of otorhinolaryngology , Ushiku Aiwa Hospital
- 3) Department of otorhinolaryngology , Nadogaya Hospital

### 1. Introduction

When we assess pathological voices in a clinical situation, we use GRBAS scales which have been recommended by the Japan Society of Logopedics and Phoniatrics.<sup>1)</sup> Since the scaling of the GRBAS is performed psychoacoustically, there could be intra- and inter-examiner variation.<sup>2)3)</sup> Clinically, this variation is small enough to permit practical evaluation of the patient's voice.

In practice, it is recommended to evaluate the quality of patient's voice during an interview. This means that the evaluation should be done based on running speech. On the other hand, recently, acoustic analysis of voices have been utilized using a computer system.<sup>4)-8)</sup> However, in our own clinical experience, sometimes we have found discrepancies between psychoacoustic evaluations and evaluations by acoustic analysis. We assume that the one of the factors which contributes to this discrepancy might be the different properties of the speech materials to be examined, that is, running speech or sustained vowels. This paper is concerned with the differences between read speech and sustained vowels in terms of our assumption.

### 2. Method

The speech materials (sustained vowels and read speech) were taken from 29 patients with various diseases : 4 cases of vocal cord polyps (VCP), 3 renke's edema (RE), 4 recurrent nerve paralysis (RNP), 4 vocal cord nodules (VCN), 3 laryngitis, 3 sulcus vocalis , 3 laryngeal cancer (KKK), 4 normals without pathology as controls. They uttered the 5 Japanese vowels and same sentences from " Jakku to mamenoki " .

These voice samples were digitized through a 12-bit A/D converter at a sampling rate of 20 KHz (read speech) and 5 KHz (sustained vowels) and stored on a disk controlled by a computer. Acoustically regular segments of one second in duration were extracted from the sustained vowels. These voice samples (vowels and speech) were recorded on a listening DAT tape in random order.

Eight experienced clinicians served as the listeners in this study. They were asked to evaluate the speech materials using the GRBAS scale which consists of five factors : Overall Grade of Hoarseness, Roughness, Breathiness, Asthenicity, and Strained quality. Each factor is scaled from 0 to 3, meaning none , slightly, certainly and strongly, respectively. The listeners were asked to evaluate the speech materials once a week. The listening trials were conducted five times. The speech materials used in this study are listed below. (Table.1)

Group	Sex		Totals
	F	M	
VCP	2	2	4
RE	2	1	3
RNP	2	2	4
VCN	4	0	4
Laryngitis	3	1	4
Sulcus	2	1	3
KKK	0	3	3
Hlth	2	2	4
Totals	17	12	29

Table.1  
age : 17~66 , mean 44

### 3. Results

The standard errors of evaluation in a listener were within 0.6 and mostly consistent. The correlations of the evaluations on sustained vowels and read speech were fairly high. The correlation coefficients for " G ", " R ", " B " and " S " were 0.68 , 0.50 , 0.63, 0.44, respectively. The rank correlations of all the factors were statistically significant. We analyzed all of the scores by ANOVA.

The evaluation of the sustained vowels and read speech showed differences in terms of the different voice samples. For factor " G ", seven samples out of twenty-nine were marked higher on all the scales in the scales in the sustained vowels than in the read speech ; (24%) , twelve samples showed no significance (42%) , ten samples were marked higher in the read speech than in the sustained vowels ( 34%). A similar tendency was observed for each of the other factor. ( Fig.1)

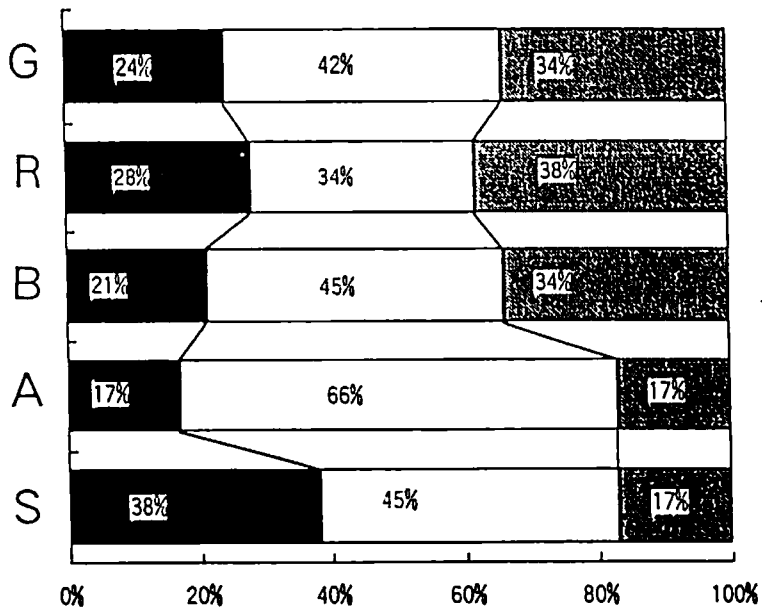


Fig.1 Incidence of samples which received a different evaluation between sustained vowels and read speech. ■ : higher in sustained vowels than in read speech. □ : no significance ▨ : higher in read speech than in sustained vowels.( N=29)

These results indicate that there is no apparent bias of evaluation for sustained vowels and read speech. Next, we examined the samples with a relatively high degree of agreement among eight listeners within one rank.

1) Samples with relatively high agreement in scores for read speech

For the factor " G ", there were sixteen samples with relatively high accordance in scores for read speech. The mean scores in read speech were higher than in the sustained vowels for all listeners and were statistically significant ( significance level 5%). (Fig.2)

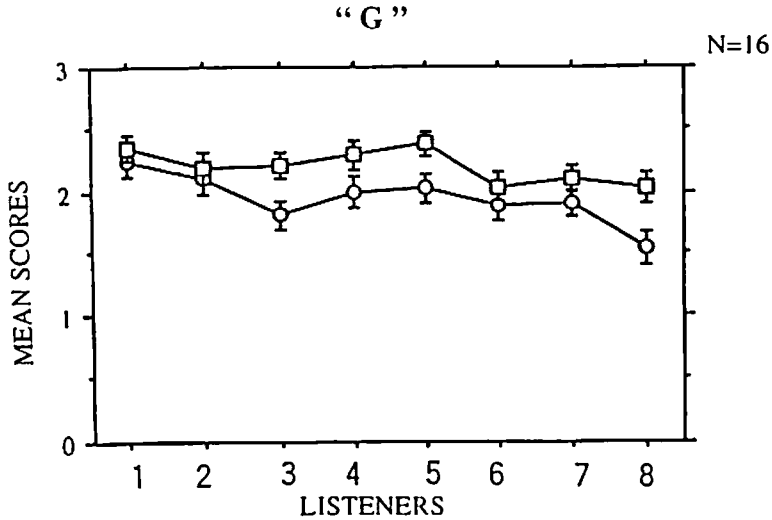


Fig.2 Results of ANOVA for the factor " G ".  
 -○- : sustained vowels -□- : read speech

For the factor " R ", there were nine samples with relatively high accordance. For the factor " B ", there were thirteen samples. Neither factor was statistically significant. (Fig.3)

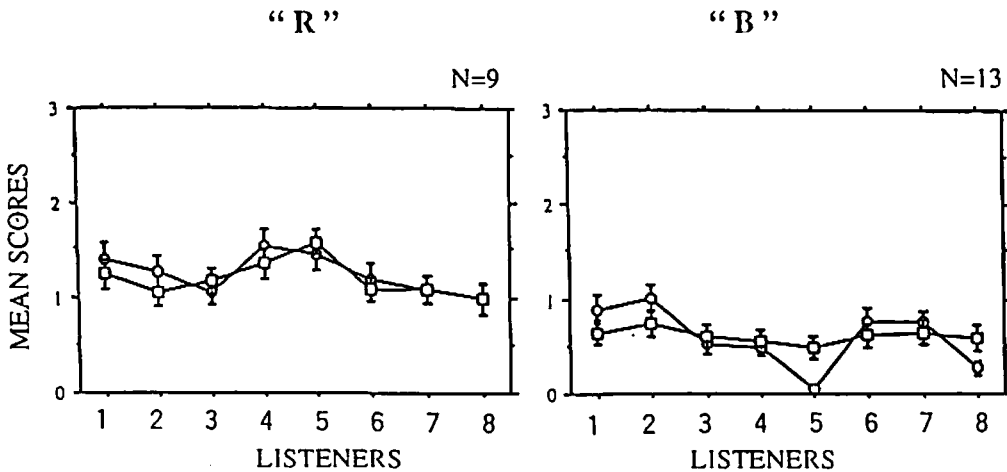


Fig.3 Results of ANOVA for the factors " R " and " B ".  
 -○- : sustained vowels -□- : read speech

For the factor " A " , there were twenty -four samples with relatively high accordance in scores for read speech. For the factor " S " , there were twenty samples. The mean scores for the sustained vowels were higher than read speech for both factors and were statistically significant. ( Fig.4)

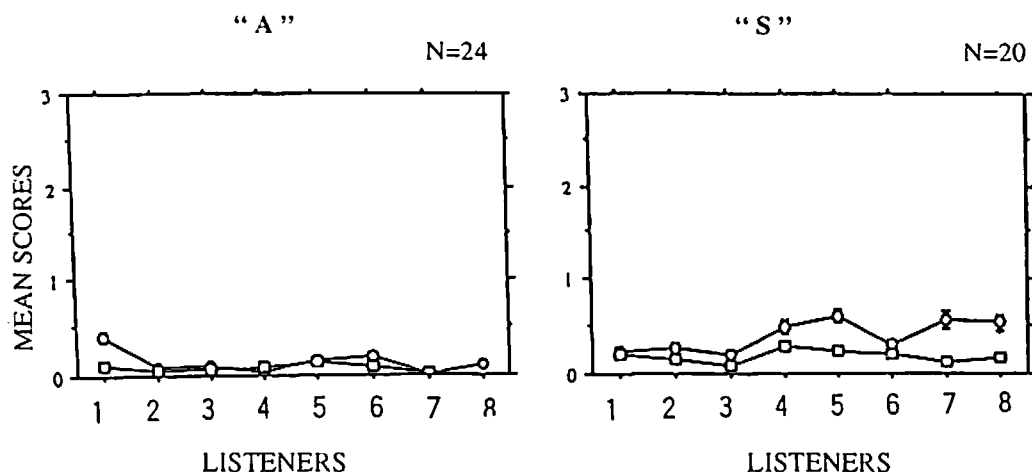


Fig.4 Results of ANOVA for the factors " A " and " S ".  
 —○— : sustained vowels    —□— : read speech

## 2) Sample with relatively high accordance in scores for sustained vowels

Similarly , we selected and analyzed with relatively high accordance in scores for sustained vowels. However, none of the factors was statistically significant.

## 3) Sample with relatively high accordance in sustained vowels and read speech

Next, we selected samples with relatively high accordance in scores for sustained vowels and read speech. For the factor " G " , the mean scores in read speech were higher than in sustained vowels. On the other hand, for both the factors " A " and " B " , the mean scores in sustained vowels were higher than in read speech and were statistically significant. These results are in accord with the results of the samples with a relatively high accordance in read speech.

## 4. Discussion

From clinical experience, we have the impression that there are some discrepancies between psychoacoustic evaluation and the acoustic analysis of hoarse voice. One of the causative factors in this discrepancy could be the nature of the examination materials. For psychoacoustic evaluation, we use running speech from interviews.

On the other hand, a sustained vowel is used for acoustic analyses.

It is obvious that different voice samples give different psychoacoustic impressions to listeners. In previous studies, it was reported that various laryngeal diseases may cause a pathological change in voice quality, and each abnormal voice may give a different perceptual impression to different listeners.<sup>9)</sup>

In this study, we tried to clarify the difference in psychoacoustic evaluation for different speech samples from each speaker. In other words, do read speech and sustained vowels give us the same GRBAS score or not ? If not, which samples gives us a clearer psychoacoustic impressions ?

We can see from Fig.2 to 4 that the " G " score was higher for read speech than for sustained vowels. On the other hand, the scores for factors " A " and " S " were higher for sustained vowels than for read speech.

Why are there differences in psychoacoustic evaluation between read speech and sustained vowels ? Running speech has special features which do not exist in sustained vowels, that is, taking breath, the onset and offset of voicing, stress, and so forth. It is reasonable to assume that these features have an influence on the evaluation of hoarse voice. We conclude that the features mentioned above made the " G " score higher for read speech than for sustained vowels. On the other hand, since the sustained vowels used in this study have an abrupt onset and offset, it is possible that this psychoacoustic evaluation causes a high " S " score for sustained vowels.

In future, further study on the evaluation of read speech may be desirable.

## 5. Conclusion

- 1) We showed a difference in the psychoacoustic evaluation of sustained vowels and read speech produced by different voices.
- 2) The factor " G " tended to be evaluated in read speech higher than in sustained vowels, and the factors " A " and " S " tended to be evaluated in sustained vowels higher than in read speech.

## 6. References

- 1) S.Hiki and S.Imaizumi : Psychoacoustic examination. in Clinical Examination of Voice. Jpn.Soc. Logopedics Phoniatrics, Eds.( Ishiyaki Shuppan, Tokyo .1979) , Chap.7, 181-209 (in Japanese)
- 2) H.Abe, H.Yonekawa , F.Ohara and S.Imaizumi : Reproducibility of Hoarse Voice Psychoacoustic Evaluation. Jpn. J.Logop.Phoniatr., 27: 168-177 (1986)
- 3) S.Imaizumi. : Acoustic measures of pathological voice qualities -Roughness- . Ann.Bull.RILP No.19, 179-190 (1985)
- 4) Y.Koike : Application of some acoustic measures for the evaluation of laryngeal dysfunction. Stud.Phonol., 7 : 17-23. (1973)
- 5) E.Yumoto : Quantitative measurements of the degree of hoarseness. Practica Otologica 76:9 2151-2168 (1983)
- 6) H.Kazuya : Voice evaluation by acoustic analysis. Jpn. J.Logop.Phoniatr., 29 : 194-199 (1988)
- 7) L.Eskenzi, D.G.Childers and D.M.Hicks : Acoustic correlates of vocal quality. J.Speech and Hearing Reserch, 33, 298-306 (1990)
- 8) J.Laver, S.Hiller, and J.M.Beck : Acoustic Waveform Perturbations and voice disorders. J Voice Vol.6 No.2 115-126 (1992)
- 9) H.Muta, T.Baer, H.Fukuda, S.Saito. : Pitch-synchronous analysis of hoarseness in running speech. Ann.Bull.RILP No.22, 85-96 (1988)