**Short Communication**

**Chaos of Vowel /a/ in Japanese Patients with Depression: A Preliminary Study**

Takashi Shimiizu¹, Nobuhiro Furuse², Tomotake Yamazaki¹, Yoichi Ueta³, Taizo Sato¹ and Shoji Nagata¹

¹Department of Mental Health, Institute of Industrial and Ecological Sciences, University of Occupational and Environmental Health, ²Mitsubishi Space Software Co. Ltd., ³Mein Mental Health Institute Inc. and ⁴Department of Physiology 1, School of Medicine, University of Occupational and Environmental Health, Japan

**Key words:** Depression, Largest Lyapunov exponent, Kolmogorov-Sinai entropy, Voice, Chaos

Chaos theory has recently given us a new insight into cardiology,²,³ otolaryngology,⁴ the nervous system,⁵,⁶ and psychiatry.⁶,⁷ Yeragani et al. found the phenomenon whereby chaos of heart rate time series diminished in patients with major depression. Drawing on recent developments in voice analyses, it is a well-known fact that the Japanese vowel /a/ has chaotic characteristics of voice.³,⁷ We have empirically found that patients with depression have a specific vocal sound, that is, a low, discouraged, and sorrowful tone of voice. Furthermore, the screening tests for depression are self-reported and subjective, and depend on respondents’ intention. To investigate the possibility of an objective depression scale, we investigated the relationship between depression and the chaos of vocal sounds.

The orbit instability of time series data, i.e. an electrocardiogram, pulsation wave, and vocal sounds, is expressed as the Lyapunov exponents (λ₁). This represents the degree of extension of the orbit displacement embedded in the reconstructed space. In the λ₁, the largest one is called the largest and first Lyapunov exponent (λ₁). The sum of the positive values in the λ set is called the Kolmogorov-Sinai entropy (KS). Long-term unpredictability of the embedded orbits is quantified using the KS.⁸ The λ₁, refers to the instability of trajectories reconstructed in time series data and the KS refers to their unpredictability. High λ₁ denotes high instability of the trajectories, while high KS denotes high unpredictability.

We hypothesized that there was a difference between the chaotic characteristics of the voices of patients with depression and those of healthy controls. We performed non-linear analyses of the voices of patients with depression and healthy controls in order to verify our hypothesis.

**Materials and Methods**

The participants were twenty-nine male system engineers at Japanese computer companies. Before starting our study, we explained its purpose and contents and obtained their consent. Of them, twelve participants (patients) were diagnosed as having depression or being in a depressive state by a psychiatric doctor. The others were diagnosed as healthy controls by an occupational physician with reference to the Japanese version of the Mini-International Neuropsychiatric Interview.¹⁰ None of the participants showed any otolaryngeal or pharyngeal symptoms such as sore throat, sniveling, sneezing, and hoarseness.

We sent to the participants a set of self-reported questionnaires including questions on age, sleep disturbance, self-perceived fatigue, decrease of intention and memory, and Zung’s Self-rating Depression Scale (SDS).¹² After the questionnaires had been answered, we recorded the vowel /a/, which was naturally spoken by the participants, for two seconds in sound-proofed rooms. We employed the vowel /a/ as a sample of vocal sounds because it has chaotic characteristics and is easy to pronounce. The subjects’ voices were recorded by a digital audio tape recorder (TCD-D100, Sony Corporation, Japan) using an electric condenser microphone (ECM-MS907, Sony Corporation, Japan). After eliminating high frequency background noise, analyses of the non-linear features of the recorded voices focused on the frequency ranges of 0–4 kHz that are relevant to the domain of formant F0–F4.¹³ The λ₁ and KS were calculated for 90 msec time frames (total number of samples=4,000) extracted from the voice time series by using the Sano and Sawada algorithm with some explanatory parameters.

The λ₁ is defined by the following formula:

$$\lambda_1 = \lim_{N \to \infty} \left( \frac{\sum_{i=1}^{N} \sigma_i}{N} \right)$$

Where σ_i (N) is a positive constant matrix required from a Jacobian matrix of voice orbit embedded in the reconstruction space. N takes the convolution product (N-th convolution map) of a Jacobian matrix.

We compared age, λ₁, KS, and SDS of the patients with those of the controls by the Student’s t-test, and also the rates of sleep disturbance, self-perceived fatigue, and decrease of intention and memory of the patients with those of the controls.

After the above study, we investigated the relationship between the λ₁ and KS of vowel /a/ when twelve normal 20- to 50-yr-old male workers spoke it naturally (Natural Normal Speech, NNP) and when the same workers simulating a depressive condition and spoke (Non-Natural Speech, NNS).

Received July 30, 2004; Accepted Jan 26, 2005

Correspondence to: T. Shimizu, Department of Mental Health, Institute of Industrial and Ecological Sciences, University of Occupational and Environmental Health, 1–1 Iseigaoka, Yahatanishi-ku, Kitakyushu-City, Fukuoka 807-8555, Japan (e-mail: t-shimiz@nifty.com)
Normal Speech, NNNP). We also studied the relationship between the $\lambda_1$ and KS of vowel /a/ when twelve 20- to 50-yr-old depressive patients spoke it naturally (Natural Depressive speech, NDS) and when the same patients simulating cheerfulness spoke (Non-natural Depressive Speech, NNDS) to confirm the reliability of the $\lambda_1$ and KS of vowel /a/.

We calculated the Pearson’s correlation between the $\lambda_1$ and KS of vowel /a/ of NNP and those of NNNP, and between the $\lambda_1$ and KS of vowel /a/ of NDS and those of NNDS. The calculations were performed with SPSS 11.0 J.

Results

Figures 1 and 2 show the distribution between $\lambda_1$, KS, and SDS of the participants.

The mean age of the patients was not significantly different from that of the controls. On the other hand, mean $\lambda_1$ and KS of the patients were significantly higher than those of the controls (Table 1).

The rates of sleep disturbance, self-perceived fatigue, and decrease of intention and memory of the patients were higher than those of the controls (Table 1).

The Pearson’s correlation coefficient between the $\lambda_1$ or KS of vowel /a/ of NNS and those of NNNS was $r=.68$ ($p<.01$) or $r=.74$ ($p<.01$), respectively. On the other hand, The Pearson’s correlation coefficient between the $\lambda_1$ or KS of vowel /a/ of NDS and those of NNDS was $r=.72$ ($p<.01$) or $r=.76$ ($p<.01$), respectively. These results suggest that the $\lambda_1$ and KS have sufficient high test-retest reliability.

Discussion

This is the first study to investigate the relationship

---

Table 1. The characteristic of the patients with depression and healthy controls

<table>
<thead>
<tr>
<th></th>
<th>Patients (n=12)</th>
<th>Controls (n=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>33.3 (6.3)</td>
<td>35.5 (5.5)</td>
</tr>
<tr>
<td>SDS</td>
<td>45.0 (8.5)**</td>
<td>36.2 (7.2)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.25 (0.02)**</td>
<td>0.22 (0.02)</td>
</tr>
<tr>
<td>KS</td>
<td>0.36 (0.04)**</td>
<td>0.31 (0.04)</td>
</tr>
<tr>
<td><strong>Percentage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>50%</td>
<td>0%</td>
</tr>
<tr>
<td>Self-perceived fatigue</td>
<td>50%</td>
<td>0%</td>
</tr>
<tr>
<td>Decrease of intention</td>
<td>50%</td>
<td>0%</td>
</tr>
<tr>
<td>Decrease of memory</td>
<td>42%</td>
<td>0%</td>
</tr>
</tbody>
</table>

1) **: $p<.01$ by the t-test  2) SDS : Zung’s Self-rating Scale  3) $\lambda_1$ : Largest Lyapunov exponent  4) KS : Kolmogonov-Sinai entropy
between depression and the chaos of vocal sounds of Japanese patients with depression. Although the present study had a small sample size, our results are compatible with the empirical finding, that patients with depression frequently have a low, discouraged, or sorrowful tone of voice. Previous studies have shown that depression is related to chaotic pulsation of human capillary vessels, and autonomic nervous function and chaos of heart rate time series. It is well-known that the motion of vocal folds has chaotic characteristics and that patients with depression have decreased chaos of heart rate time series, which suggests that depression decreases vocal function. Our results imply that depression decreases laryngeal vocal function, which control the chaotic motion of vocal folds, and consequently elevate chaos of vocal sounds. We cannot completely deny other reasons for our results, for example, depressive patients tend to lose the use of their vocal folds and to experience frequent thirst, a common side effect of anti-depressive medication. However, we observed a decrease in the $\lambda$ and KS in accordance with the improvement of the depressive status in some patients (data not shown).

Chaos theory can be applied to “homeodynamics” of the human body. Our findings might indicate “homeodynamics” of the vocal function. Further study should investigate the relationship between the chaos feature of voice quality and mental health status by a well-designed method.

The present study had three limitations. First, it was not a non-randomized control study. The control group was a completely randomized group although we matched gender and work of the control group with those of the patient group. Second, we were unable to remove the effect of medicine (80% of the patient group were taking anti-depressive drugs) and intra-personal variation, completely. Also, the study sample size was small. Our method, being non-invasive, has the advantage that it is clinically easy to measure both $\lambda$ and KS as objective indices. We hope that this method will be applied to the care of workers’ mental health in the future, although careful consideration should be given to protecting their privacy.

In conclusion, the $\lambda$, and KS of the patients with depression were significantly higher than those of healthy controls. This finding implies that depression is potentially associated with the chaotic features of vocal sounds. The method of recording vocal sound is non-invasive and easy to measure clinically. Further study should investigate the association between depression and the complexity of vocal sound obtained from various voice protocols by a well-designed method.

References