

Automatic Scoring for Prosodic Proficiency of English Sentences Spoken by Japanese Based on Utterance Comparison

Yoichi YAMASHITA[†], Member, Keisuke KATO^{††}, and Kazunori NOZAWA[†], Nonmembers

SUMMARY This paper describes techniques of scoring prosodic proficiency of English sentences spoken by Japanese. The multiple regression model predicts the prosodic proficiency using new prosodic measures based on the characteristics of Japanese novice learners of English. Prosodic measures are calculated by comparing prosodic parameters, such as F_0 , power and duration, of learner's and native speaker's speech. The new measures include the approximation error of the fitting line and the comparison result of prosodic parameters for a limited segment of the word boundary rather than the whole utterance. This paper reveals that the introduction of the new measures improved the correlation by 0.1 between the teachers' and automatic scores.

key words: prosody, comparison unit, fitting error, multiple regression model, CALL

1. Introduction

Two reasons that are given as to why English conversation is difficult for many Japanese to deal with are that reading and writing are still the main subjects of English education at school, and that Japanese is linguistically much different from English. To overcome such circumstances and support the self-training of English speaking, many studies address CALL systems enhanced by speech information processing[1]–[7].

Japanese students have less opportunities to learn prosody generation than segmental pronunciation, thus it is very difficult for some Japanese people to acquire the skill of speaking English with natural prosody. This paper proposes a new method of automatic scoring for prosodic proficiency of English sentences spoken by Japanese learners focusing on the characteristics of Japanese learners.

Our proposed method supposes that the contents of the sentence are fixed and that the same sentence uttered by a native speaker is available. In this situation, the comparison between the native speaker's and learner's utterances is one of the promising ways to automatically score the learner's utterance. Natural prosody itself is not unique for a given sentence.

Comparison-based scoring methods involve a risk of giving a low score for an utterance with a natural prosodic pattern if the prosodic pattern of the utterance happens to be different from a reference utterance by a native speaker. However, it is very efficient for training learners (especially beginners) to mimic a teacher's utterance. In this paper, utterance comparison between native speaker's and learner's utterances is used to automatically score the prosodic proficiency of the learners' utterances.

In general, students tend to make a prosodic expression in the same manner as they use in their native language, when they speak in a foreign language. Some differences between native speaker's and learner's utterances are dependent on the native language of the learner. This paper describes an automatic scoring method for prosodic proficiency of English sentences based on the characteristics of Japanese learners.

2. Methods of Automatic Scoring for Prosodic Proficiency

2.1 A Conventional Method

A simple method of automatic scoring of learners' utterances based on the utterance comparison has been evolved. This method evaluates the difference between two prosodic patterns frame by frame for the entire sentence, using the linear expansion or the DP (Dynamic Programming) matching to obtain the matching information of words or phonemes[8]. When characteristics of the learners are known in advance, the performance of the automatic scoring can be improved by incorporating the characteristics into scoring methods.

2.2 Characteristics of Japanese Learners of English

Sugito(1998) described the difference of F_0 patterns between a native speaker and a Japanese novice learner in uttering word clusters, schematically shown in Fig. 1[9]. Native speakers smoothly utter a word sequence, which should be a cluster, with a smooth F_0 pattern. On the other hand, a valley is found at the word boundary of an F_0 pattern by Japanese novice learners. Her analysis results suggest that the F_0 pattern in the word

Manuscript received January 1, 2004.

Manuscript revised January 1, 2004.

Final manuscript received January 1, 2004.

[†]College of Information Science and Engineering, Ritsumeikan University

^{††}Graduate School of Science and Engineering, Ritsumeikan University

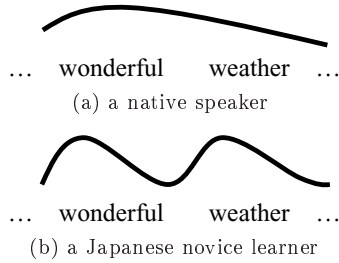


Fig. 1 Typical F_0 patterns at a word boundary.

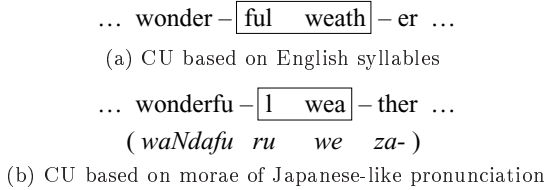


Fig. 2 Definition of the comparison unit.

boundary reflects the skill level of speaking English.

2.3 A Proposed Method

This paper proposes a method of the automatic scoring focusing on segments where the characteristics of Japanese English tend to appear, rather than using all segments of the sentence. The segments to be used for utterance comparison, such as words, word boundaries, the entire utterance, and so on, are called a *comparison unit* (CU).

We introduces the word boundary as one of the comparison units based on the findings mentioned in 2.2, and investigates two types of definitions of the word boundary part as the comparison unit for F_0 patterns as follows:

- The final and initial English syllables of the preceding and following word of a word boundary respectively. (See in Fig. 2(a))
- The final and initial morae of the preceding and following word of a word boundary respectively, when these words are spoken with Japanese-like pronunciation. (See in Fig. 2(b))

A proposed method predicts a score of the prosodic proficiency by the multiple regression model using several prosodic measures. The score of the prosodic proficiency, S , is defined by the following equation,

$$S = \sum_{k=1}^K a_k P_k, \quad (1)$$

where P_k is the k -th prosodic measure and K is the number of the prosodic measures. The multiple regression model is obtained by optimizing weighting factors, a_k , using training data.

The prosodic measures are calculated by averaging the difference of prosodic parameters for comparison units. A prosodic measure, P_k , is defined by the following equation,

$$P_k = \frac{1}{I} \sum_{i=1}^I |P_N(i) - P_L(i)|, \quad (2)$$

where I is the number of the comparison units, and $P_N(i)$ and $P_L(i)$ are prosodic parameters of the i -th comparison unit for a native speaker and a learner respectively.

The alignment between two prosodic patterns is not only determined by the prosodic information but also the segmental information in terms of phoneme labels[10]. The utterances by native speakers are manually segmented phoneme by phoneme. The phoneme labels for learners' utterances are determined by an automatic labeling technique using acoustic HMM models of both English and Japanese phonemes and the phoneme transition patterns which are often found for Japanese speakers. The initial and final points of each comparison unit are given by the phoneme labels.

3. Prosodic Measures

3.1 Fundamental Frequency (F_0)

F_0 is analyzed every 5ms with 20ms frames for utterances. F_0 values are expressed in the log scale and are normalized to make the mean 100.0. After unvoiced segments are interpolated to keep F_0 values, the F_0 pattern is smoothed by using regression line fitting.

3.1.1 F_0 Distance

The first prosodic measure is F_0 pattern distance, F_PD . This measure is the average distance between F_0 value sequences by a native speaker and a Japanese learner in the comparison units by removing the difference of F_0 means of two speakers. F_PD is defined by the following equations,

$$F_PD = \frac{1}{I} \sum_{i=1}^I f_pd(i), \quad (3)$$

$$f_pd(i) = \frac{1}{K_i} \sum_{k=1}^{K_i} \left| f_N(i, k) - f_L(i, k) - \left(\overline{f_N(i)} - \overline{f_L(i)} \right) \right|, \quad (4)$$

where I is the number of the comparison units in an utterance, $f_N(i, k)$ and $f_L(i, k)$ are F_0 values of k -th frame in i -th comparison unit of a native speaker and a Japanese learner respectively. $\overline{f_X(i)}$ is the mean of F_0 values in i -th comparison unit as follows,

$$\overline{f_X(i)} = \frac{1}{K_i} \sum_{k=1}^{K_i} f_X(i, k), \quad (5)$$

where X denotes either N or L .

3.1.2 Gradient of F_0 Fitting Line

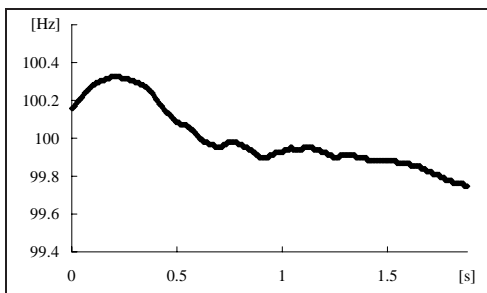
The second measure related to F_0 is F_SD , which is the difference between gradients of F_0 fitting straight lines in the comparison unit for a native speaker and a Japanese learner. F_SD is defined as

$$F_SD = \frac{1}{I} \sum_{i=1}^I \left| \Delta f_N(i) - \Delta f_L(i) \right|, \quad (6)$$

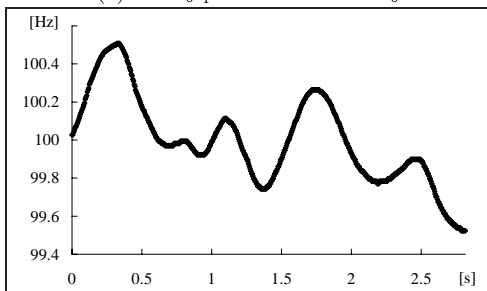
where $\Delta f_N(i)$ and $\Delta f_L(i)$ are the gradient of F_0 fitting lines in i -th comparison unit for a native and a Japanese speaker respectively.

3.1.3 Approximation Error of F_0 Fitting

Unnatural F_0 patterns by Japanese learners are classi-



(a) An F_0 pattern with flat F_0



(b) An F_0 pattern with too large F_0 change of words

Fig. 3 Typical F_0 patterns by Japanese novice learners.

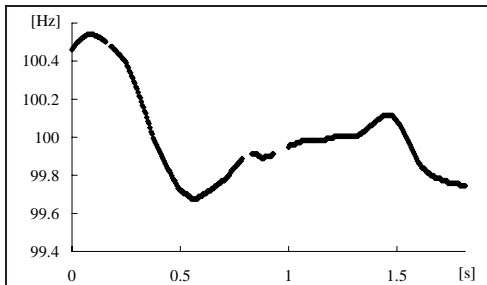


Fig. 4 An F_0 pattern by native speakers.

fied into two categories. One is a flat F_0 pattern which has small dynamic change in F_0 and it is like a straight line. This type of unnaturalness is described in [11] and some experiment results uncovered this characteristic [12], [13]. The other is concatenation of word F_0 patterns where an inherent prosodic pattern of a word appears in a sentence as it is, and it results in the unnaturally large F_0 change in words. Typical examples of two types of unnatural F_0 patterns are shown in Fig. 3 (a) and (b) respectively. The sentence of both F_0 patterns is “That’s from my brother who lives in London”. Figure 4 is an F_0 pattern by a native speaker for the same sentence.

When an F_0 pattern by Japanese novice learners is approximated by a straight or curved line, an approximation error decreases or increases in comparison with a native speaker’s F_0 pattern, for the two cases mentioned above. Therefore, this paper proposes the approximation error of F_0 fitting as a new prosodic measure in order to capture the characteristics of Japanese learners of English.

The measures, F_ND1 and F_ND2 , are the differences of the approximation errors of F_0 fitting which approximates an F_0 trajectory by a straight and 2-order curved line respectively. F_NDn is defined as

$$F_NDn = \frac{1}{I} \sum_{i=1}^I \left| fd_N^n(i) - fd_L^n(i) \right|, \quad (7)$$

where $fd_N^n(i)$ and $fd_L^n(i)$ are the approximation errors of F_0 fitting in i -th comparison unit for a native and a Japanese speaker respectively, and n indicates the order of the fitting line. The approximation error is evaluated in terms of RMS normalized by the number of frames in the comparison unit.

3.2 Power

The characteristic of Japanese novice learners mentioned in 3.1.3 is found for the power contour as well as the F_0 contour. The measures related to the power parameter, P_ND1 and P_ND2 , are the differences of the approximation errors of power contour fitting which approximates a power trajectory by a straight and 2-order curved line respectively, in the same manner as F_NDn . P_NDn is defined as

$$P_ND = \frac{1}{I} \sum_{i=1}^I \left| pd_N^n(i) - pd_L^n(i) \right|, \quad (8)$$

where $pd_N^n(i)$ and $pd_L^n(i)$ are the approximation errors of fitting for the power in i -th comparison unit for a native and a Japanese speaker respectively. The approximation error is evaluated in terms of RMS normalized by the number of frames in the comparison unit.

3.3 Duration

3.3.1 Utterance Duration

Japanese novice learners tend to utter sentences with longer duration because they often insert unnecessary vowels in a consonant sequence and some pauses in a halting utterance. The total duration of the utterance can be an important factor of scoring of the utterance. The first measure related to the duration is the total utterance duration, T_D_{all} , which is defined as

$$T_D_{all} = |T_N - T_L|, \quad (9)$$

where T_N and T_L are the total duration times of the native speaker's and learner's utterance respectively.

3.3.2 Duration of Words

If the utterance duration of the learner is almost the same as that of the native speaker, the inappropriate rhythm decreases the teachers' evaluation score. We use two measures related to the word duration, T_D_{word} and T_RD_{word} , which are defined as

$$T_D_{word} = \frac{1}{J} \sum_{j=1}^J |t_N(j) - t_L(j)|, \text{ and} \quad (10)$$

$$T_RD_{word} = \frac{1}{J} \sum_{j=1}^J \left| \frac{t_N(j)}{T_N} - \frac{t_L(j)}{T_L} \right|, \quad (11)$$

where $t_N(j)$ and $t_L(j)$ are the durations of j -th word and J is the number of the words in the utterance. T_RD_{word} is calculated based on the relative duration of words which is normalized by the total utterance duration while T_D_{word} is calculated based on the duration without normalization.

3.3.3 Duration of Pauses

The pauses are very important to a prosodic impression of the utterance in addition to the word duration. Two measures related to the pause are investigated. T_D_{pause} and T_RD_{pause} are defined as follows,

$$T_D_{pau} = \frac{1}{J-1} \sum_{j=1}^{J-1} |pau_N(j) - pau_L(j)|, \quad (12)$$

and

$$T_RD_{pau} = \frac{1}{J-1} \sum_{j=1}^{J-1} \left| \frac{pau_N(j)}{T_N} - \frac{pau_L(j)}{T_L} \right|, \quad (13)$$

where $pau_N(j)$ and $pau_L(j)$ are the durations of the pause which immediately follows j -th word. The difference between T_D_{pau} and T_RD_{pau} without and with normalization is the same as the word duration parameters.

4. Evaluation Results

4.1 Speech Data

The method of the automatic scoring of prosodic proficiency is trained and evaluated with the ERJ (English Read by Japanese) corpus[14],[15]. The ERJ corpus provides pronunciation proficiency scores, which are rated by some American teachers of English from segmental, rhythmic, and intonational aspects, for some utterances.

For intonational evaluation, 60 sentences are designed to have various intonation patterns, and 8 Japanese speakers uttered each sentence. A native speaker also uttered the same sentence set. The raters were asked to listen to the native utterances as the reference which are supposed to have the correct intonational patterns of the sentences, and they gave 5-scale rating (1:poor, 5:good) to the utterances by Japanese from the intonational aspect. Most of the utterances were rated by 4 teachers, but some utterances were rated by 2 or 3 teachers. We used 10 sentences among 60 sentences both for training and testing of the method, and the number of the utterances for training and testing was 73 and 75, respectively, because some utterances were removed from the data set due to errors of the automatic labeling process.

Our proposed method automatically scores the learners' utterances based on the comparison with the native speaker's utterances of the same sentence. The native speaker's utterances used by the raters as the reference are also used for the utterances to be compared with the learners' utterances by the automatic scoring.

4.2 Correlation between Teachers' Scores and Prosodic Measures

In order to select prosodic measures for the multiple regression model, correlation coefficients between the teachers' scores and the prosodic measures were calculated. We investigated 5 kinds of CUs including the sentence, the word, the prosodic phrase, and two types of word boundary parts mentioned in 2.3. A Japanese teacher of English identified prosodic phrases in a sentence using a written text of the sentence based on his expertise. Tables 1, 2, and 3 summarize the correlation coefficients for the comparison units. The coefficient -1.0 indicates perfect correlation.

For F_0 measures, F_PD and F_SD , the measure based on the word boundary has larger correlation between teachers' score and the F_0 measures. New F_0 measures F_NDn based on approximation errors of line fitting show the high correlation using a word as a unit of the line fitting. The measure based on the prosodic phrase has smaller correlation because some prosodic

Table 1 Correlation between teachers' score and F_0 Measures for several comparison units (CUs).

| Measure | CU | cor. |
|----------|-------------------------|-------|
| F_PD | sentence | -0.27 |
| | word | -0.36 |
| | prosodic phrase | -0.22 |
| | syllable-based boundary | -0.49 |
| | mora-based boundary | -0.41 |
| F_SD | word | -0.26 |
| | prosodic phrase | -0.16 |
| | syllable-based boundary | -0.36 |
| | mora-based boundary | -0.45 |
| F_ND1 | sentence | -0.14 |
| | word | -0.37 |
| | prosodic phrase | +0.02 |
| | syllable-based boundary | -0.26 |
| F_ND2 | sentence | -0.15 |
| | word | -0.42 |
| | prosodic phrase | -0.02 |
| | syllable-based boundary | -0.40 |

Table 2 Correlation between teachers' score and power measures for several comparison units (CU).

| Measure | CU | cor. |
|----------|-----------------|-------|
| P_ND1 | sentence | -0.18 |
| | prosodic phrase | -0.38 |
| | word | -0.38 |
| P_ND2 | sentence | -0.21 |
| | prosodic phrase | -0.27 |
| | word | -0.27 |

Table 3 Correlation between teachers' score and duration measures for several comparison units (CU).

| Measure | cor. |
|----------------|-------|
| T_D_{all} | -0.25 |
| T_RD_{word} | -0.23 |
| T_D_{word} | -0.44 |
| T_RD_{pau} | -0.02 |
| T_D_{pau} | -0.13 |

phrase boundaries do not match the actual prosodic phrase boundaries in the utterances.

For the duration measures, the difference of non-normalized word duration shows the highest correlation. It suggests that the combination of rhythm and duration is effective to predict the prosodic proficiency score. The pause information does not contribute to the prosodic proficiency.

4.3 Multiple Regression Model

The score of the prosodic proficiency, S , is predicted by the multiple regression model formulated as Eq. (1). Some prosodic measures mentioned in Sect. 3 have high inter-measure correlation with each other. To reduce the number of model parameters and to increase the robustness, the input variables of the multiple regression model consist of the selected prosodic measures which have high correlation to the teacher's score and have low correlation to other input variables. The cor-

Table 4 The correlation between the teachers' score and automatic scoring.

| measure set | closed | open |
|------------------|--------|------|
| set-0 (baseline) | 0.40 | 0.41 |
| set-I (proposed) | 0.69 | 0.51 |

relation between prosodic parameters were investigated to select measures.

We selected 5 measures,

- P_1 : F_PD [syllable-based boundary],
- P_2 : F_ND [word],
- P_3 : P_ND2 [word],
- P_4 : T_D_{all} , and
- P_5 : T_RD_{word}

as the input variables of the multiple regression model, Eq. (1), where $[\]$ indicates the CU. This set of measures is called set-I. The score of the prosodic proficiency, S , is predicted by

$$S = -0.44P_1 - 0.22P_2 - 0.16P_3 - 0.31P_4 - 0.23P_5. \quad (14)$$

The weights of the input variables are obtained by normalizing the mean and the variance of each input variable and the target S to 0 and 1 respectively.

To verify the improvement by the proposed method, a baseline model is also trained with another set of measures (set-0), and it uses 3 measures,

- P'_1 : F_PD [sentence],
- P'_2 : F_SD [word], and
- P'_3 : T_D_{all} .

This set eliminates the newly proposed measures X_NDn based on approximation errors of line fitting and the measures for CU of the word boundary. The score S is predicted by

$$S = -0.23P'_1 - 0.14P'_2 - 0.23P'_3. \quad (15)$$

The performance of the models is evaluated on the testing data which is not included in the training data. The correlation coefficients between the teachers' score and automatic scoring by the models are shown in Table 4. We conducted statistical tests to verify that the improvement by the proposed method is significant. The existence of correlation was tested before tests for the increase of the correlation coefficients. A t -test can be used to test whether the correlation between two variables is significant [16], [17]. The null hypothesis that the correlation coefficient is 0 is rejected for all 4 cases in Table 4 at the level of significance $\alpha = 0.01$. In the second test, the difference of correlation coefficients was tested based on Fisher's Z transformation [16]. The null hypothesis is that the correlation coefficient by the proposed method is not larger than that by the baseline method. This hypothesis is also rejected for the closed and open prediction at $\alpha = 0.01$ and 0.25

respectively. These results reveal that the proposed method improves the accuracy of automatic scoring of the prosodic proficiency. However, the improvement for the open prediction is not supported at a small significance level. It implies that it is necessary to increase the training data for a more robust multiple regression model.

5. Conclusion

This paper proposed the introduction of new prosodic measures which describe the characteristics of Japanese novice learners of English into automatic scoring of the prosodic proficiency. The new measures include the approximation error of the fitting line and the comparison result of prosodic parameters for the limited segment of the word boundary rather than the whole utterance. The multiple regression model predicts the prosodic proficiency. The introduction of the new measures improved the correlation between the teachers' score and the automatic score from 0.40 to 0.69 and from 0.41 to 0.51 for the closed and open evaluation respectively. The increase of the training data will be an important issue of the future work.

Acknowledgments

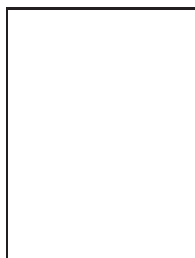
This paper uses the corpus of "English Speech Database Read by Japanese (ERJ)". We would like to thank the researchers who devoted to considerable efforts to construct the ERJ corpus and the project.

References

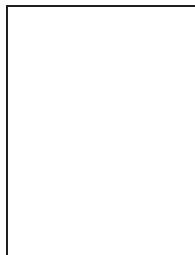
- [1] H. Hamada, S. Miki, and R. Nakatsu, "Automatic Evaluation of English Pronunciation Based on Speech Recognition Techniques", IEICE Trans. Inf. & Syst., vol.E76-D, no.3, pp.351-359, 1993.
- [2] S. Nakagawa, A. A. Reyes, H. Suzuki, and Y. Taniguchi, "An English Conversation CAI system Using Speech Recognition Technology", Trans. of IPSJ, vol.38, no.8, pp.1649-1658, 1997.
- [3] G. Kawai, A. Ishida, and K. Hirose, "Detecting and Correcting Mispronunciation in non-native pronunciation learning using a speech recognizer incorporating bilingual phone models", Journal of ASJ, vol.57, no.9, pp.569-580, 2001.
- [4] N. Minematsu, Y. Fujisawa, and S. Nakagawa, "Automatic Estimation and Visualization of English Pronunciation Habit and Its Application to Prosodic Evaluation of Pronunciation Proficiency", IEICE Trans., vol. J83-D-II, no.2, pp.486-499, 2000.
- [5] S. Kobashikawa, N. Minematsu, K. Horose, and D. Erickson, "Modeling of Stressed Syllables for their Detection in English Sentences to Develop an English Rhythm Learning System", Technical report of IEICE, SP2001-100, pp.99-104, 2001.
- [6] K. Imoto, Y. Tsubota, T. Kawahara, and M. Dantsuji, "Modeling and Automatic Detection of English Sentence Stress for Computer-Assisted English Prosody Learning System", Journal of ASJ, vol.59, no.4, pp.183-191, 2003.
- [7] S. Nakagawa, "Language Learning Systems Using Spoken Language Processing Techniques", Journal of ASJ, vol.59, no.6, pp.337-344, 2003.
- [8] H. Yanagase and Y. Yamashita, "Automatic Scoring of Prosody in English Learners' Speech Based on Utterance Comparison", Technical report of IEICE, SP2000-118, pp.9-16, 2001.
- [9] M. Sugito, "*Nihon-jin no eigo* (English by Japanese speakers)", Izumi-shoin, 1998.
- [10] H. Yanagase and Y. Yamashita: "A Study on Methods of Comparing Two Utterances for Supporting Language Learning", Proc. of ASJ Fall meeting, 3-5-7, pp.83-84, 2000.
- [11] K. Shimizu, "*Eigo onseigaku* (Phonetics of English)", Sokei-shobo, 1995.
- [12] T. Miwa and S. Nakagawa, "Analysis and Comparison of the Prosodic Features for Japanese English and native English", Technical report of IEICE, SP2001-142, pp.51-58, 2002.
- [13] H. Obari, R. Tomiyama, M. Yamamoto and S. Itahashi, "Comparison of Prosodic Features between the Japanese Speakers of English and the Native Speaker", Proc. of ASJ Fall meeting, 2-8-6, pp.255-256, 2003.
- [14] N. Minematsu, Y. Tomiyama, K. Yoshimoto, K. Shimizu, S. Nakagawa, M. Dantsuji, and S. Makino, "Rating of Japanese Students' Utterances of English by Native English Teachers", Proc. of ASJ Fall meeting, 1-6-4, pp.215-216, 2002.
- [15] N. Minematsu, Y. Tomiyama, K. Yoshimoto, K. Shimizu, S. Nakagawa, M. Dantsuji, and S. Makino, "Development of English Speech Database Read by Japanese to Support CALL Research", Proc. of ICA 2004, Kyoto, vol.1, pp.557-560, 2004.
- [16] M. R. Spiegel and L. J. Stephens, "Schaum's Outline of Theory and Problems of Statistics —Third Edition", McGraw-Hill, 1998.
- [17] R. Larson and B. Farber, "Elementary Statistics — Picturing the World", Prentice-Hall, 2000.



Yoichi YAMASHITA was born in Ehime, Japan, in 1959. He received his B.E., M.E. and D.Eng. degrees from Osaka University, in 1982, 1984 and 1993 respectively. He was a Technical Official (1984-1992), a Research Associate (1993-1994), and an Assistant Professor (1994-1997) at the Institute of Scientific and Industrial Research of Osaka University. In 1997, he joined Ritsumeikan University as an Associate Professor in the College of Science and Engineering. He was promoted to Professor in 2001, and is currently a Professor in the College of Information Science and Engineering since 2004. His research interests include speech understanding, speech synthesis and spoken dialog processing. Dr. Yamashita is a member of the Acoustical Society of Japan, Japanese Society for Artificial Intelligence, Information Processing Society of Japan, International Speech Communication Association, and IEEE.



Keisuke KATO was born in Kyoto, Japan, in 1978. He received his B.E degree from Ritsumeikan University in 2002. Since 2002, he has been a student in the Master's Program in the Graduate School of Science and Engineering, Ritsumeikan University. His research interests are automatic scoring for language learner's speech. He is a member of the Acoustical Society of Japan.



Kazunori NOZAWA was born in Utsunomiya, Japan, in 1951. He received his B.Edu. from Utsunomiya University, Japan in 1974 and M.A. in Education from the University of Kansas, U.S.A. in 1978 respectively. He was an Instructor at Nagoya International College (1978-1981), an Assistant (1981-1985), a Lecturer (1985-1989), an Associate Professor (1989-1998) at Toyohashi University of Technology. In 1998 he joined Ritsumeikan University as a Professor in the Faculty of Economics. In 2003 he also became a Professor at Graduate School of Language Education and Information Science. In 2004 he was transferred to College of Information Science and Engineering. His main research areas are CALL (Computer Assisted Language Learning) and Intercultural Communication. Professor Nozawa is a member of Council for Improvement of Education through Computers, The Association of International Behavioral Studies, Japan Association of Educational Technology, The Communication Association of Japan, Japan Association for Language Teaching, Japan Association of College English Teachers, Teaching English to Speakers of Other Languages, and European CALL Association. He is a co-editor of CALL-EJ Online since 1999.