Preliminary approach to the method of accent and intonation correction for learners of Japanese*

Anna Sheldeshova and Sumio Ohno (Tokyo University of Technology)

1 Introduction

Shadowing as a popular tool for self-studying of pronunciation and intonation has already proven its effectiveness in different languages. It is supposed mostly for learners of advanced levels, since the beginners have not enough knowledge to determine and self-reflect their mistakes. The research aims to find a method to help learners of beginner level to practice shadowing so it can serve a basis for development of an automatic feedback system for language learners.

The present study provides the results of an experiment of the research on an automatic analysis and feedback method on improvement speaking skills for learners of Japanese during shadowing exercises. The evaluation of shadowing utterances provided by Japanese native speakers shows weakness and incorrectness of the utterances in mora duration and accent control, as well as intonation pattern in the whole utterance. We suggest preliminary findings and criteria for automatic evaluation and feedback on the learners' utterances.

2 Speech Material

The speech material was used from ATR503 sentences speech database [1], which is a collection of well phonetically balanced utterances spoken by professional Japanese announcers.

We chose 20 sentences and their recordings as the 'reference utterances' for shadowing exercises: 10 utterances of one male native speaker and 10 utterances of one female native speaker. Each of the set consisted of one and the same list of utterances, pronounced at normal speed (natural reading) and at slow speed (phrase-segmented reading).

Six persons living in Japan for 1-3 years took part in the experiment: 3 men and 3 women. Their levels of Japanese language proficiency were as follows:

Learner 1 – JLPT N5-4, Russian as first language (FL), man

Learner 2 – JLPT N2, Chinese as FL, man

Learner 3 – JLPT N3, Russian as FL, man

Learner 4 – JLPT N2, Belorussian as FL, woman

Learner 5 – JLPT N3, Russian as FL, woman

Learner 6 – JLPT N5-4, Russian as FL, woman

The learners were asked to practice shadowing of the selected reference utterances, 5 times per each reference utterance. Audacity was used for the learners' shadowing recordings on the R-channel of the speech data while reference utterance was on the L-channel, in order to keep the relative timing between reference and learner's utterance.

The recorded utterances were labeled with phonemes using Julius phoneme segmentation kit [2]. In addition, Praat was used to obtain fundamental frequency contour of the utterances.

3 Evaluation by Native Speakers

Evaluation was received via a questionnaire form from 13 Japanese native speakers, 2 of them were university media club members who practiced public speaking skills. All the participants were in

their 21-24 y.o. They gave their answers to the following questions on each of the utterances:

1. Is there any mistake in pronunciation of sounds in the words?
2. Is there a mistake in pitch accent?
3. Is there a mistake in intonation pattern?
4. Is there anything unnatural in the whole utterance or in a part of it?

4 Results

We conducted a comparative analysis of the questionnaire answers with the acoustic features of the recorded utterances. Reference contours and learners' ones were plotted on the same time axis to detect common mistakes that were significant to the native speakers.

We can divide them into two groups. The first relates to the slow speed speech:

- flatness of the intonation of the whole utterance;
- mistakes in word accent (pitch);
- mistakes in pitch at particles.

Figure 1 represents fragments from the waveforms and fundamental frequency contours of two learners' utterances (lower) and those of their reference utterances (upper).
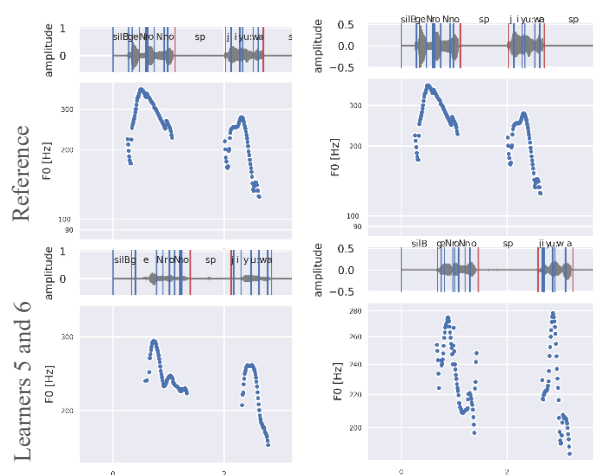

Fig. 1 Learners' flatness of tone
in utterances at slow speed

This is an example of the learners' problem in ranging the tone, so the pitch accent is almost imperceptible making the whole utterance sound monotonous. It can be recognized automatically, hence it can be commented by an automatic feedback program, too.

In the utterance at normal speed, the following mistakes were mentioned in the answers:

- mistakes in word accent (pitch);
- mistakes in mora duration.

There were less commentaries on the intonation patterns in pronouncing at normal speed compared to the slow speed utterances. However, the omission and mispronouncing of words in order to keep the reference pace were common.

Nevertheless, the flatness of the pitch still remains, as it is shown in Fig. 2.
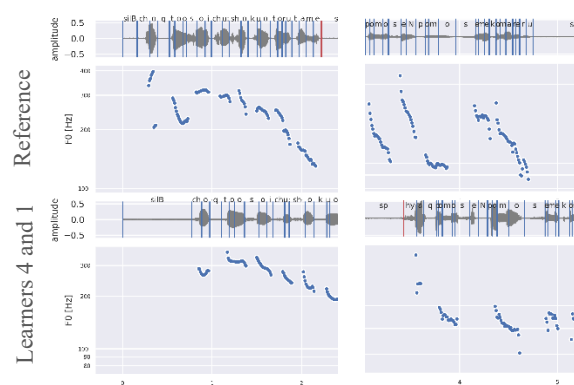

Fig. 2 Learners' results at shadowing
of the utterances at normal speed

5 Future Work

The future work will relate to defining the rules for automatic extraction of mistakes in a 'shadowed' utterance in order to show and explain the significant mistakes to a learner.

References

1. ATR 503 Sentences Database. Speech Resources Consortium (NII SRC).
   Source link: http://research.nii.ac.jp/src/en/ATR503.html
2. Speech Segmentation Toolkit using Julius.
   Source link: https://julius.osdn.jp/index.php?q=ouyoukit.html