Session 5aSC

Speech Communication: Voice, Tone, Prosody, and Affective and Emotional Qualities of Speech (Poster Session)

Tanya Flores, Chair
Languages and Literature, University of Utah, 255 S Central Campus Dr., LNCO 1400, Salt Lake City, UT 84109

All posters will be on display from 8:00 a.m. to 12:00 noon. To allow contributors in this session to see the other posters, authors of odd-numbered papers will be at their posters from 8:00 a.m. to 10:00 a.m. and authors of even-numbered papers will be at their posters from 10:00 a.m. to 12:00 noon.

Contributed Papers

5aSC1. Effects of perception and vocabulary training of Mandarin tones for native speakers of Japanese: Pre-, post-, and retention test comparison. Shuyi Yang (Graduate School of Intercultural Studies, Kobe Univ., Hyōgo-ken, Kōbe-shi, Nada-ku, Tsurukabuto, 1 Chome−2−1, Kobe 6570815, Japan, syuiyang@gmail.com) and Reiko Akahane-Yamada (ATR, Soraku-gun, Kyoto, Japan)

Native speakers of Japanese were trained to perceive four Mandarin tones, and to semantically and phonetically distinguish Chinese monosyllabic word contrasting in tones. Various tests were administered not only before and after the training period but also forty days after the completion of training. Participants were divided into three groups. First group received perception training first and vocabulary training later. Second group received vocabulary training first and perception training later. The last group, which was the control group, received no training and participated only in testing sessions. The result showed that the accuracy in perception-related-tests improved by perception training, and the accuracy in vocabulary-related-tests improved by vocabulary training. In addition, the effect of training had retained even 40 days after the completion of training. More importantly, the group which received perception training first showed significantly larger improvement from pre-test to post-test than in the group which received vocabulary training first. Note that two groups received the same amount of equivalent trainings in total. Preset results demonstrated the existence of the order effect in foreign language learning. [Work supported by JSPS KAKENHI 23242032.]

5aSC2. What ranges of two-mass model parameters should be used in subject-specific and population-based modeling studies? Douglas Cook (Eng., New York Univ. Abu Dhabi, PO Box 903, New York, NY 10276, prof.laji@gmail.com)

Two-mass models have been used in voice research for over 40 years. It is therefore both surprising and somewhat troubling that there is no firm consensus regarding the values of model parameters that should be used to represent human phonation. A knowledge of the parameter ranges that can (or should) be used is essential for scientifically valid studies involving population-based or subject-specific modeling. In this study, four techniques were used to examine the ranges of two-mass model parameter values that produce behavior representative of human phonation. The first approach involved a review of values that have been used in previous modeling studies. The second approach utilized unrestricted Monte Carlo sampling to examine which ranges can be used to simulate human phonation. The third approach also utilized Monte Carlo sampling, but parameters were restricted based on physical features of the vocal folds. Finally, a reduction or order technique was developed that allows the determination of two-mass model parameters from the physical features of human vocal folds. Finally, results from each of the four methods were compared and contrasted to provide a better understanding of parameter ranges for two-mass models.

5aSC3. A measurement study on voice instabilities during the register transition. Yasufumi Uezu and Tokihiko Kaburagi (Kyushu Univ., 4-9-1, Shiobaru, Minami-ku, Fukuoka-shi, Fukuoka 815-8540, Japan, 3DS14006W@s.kyushu-u.ac.jp)

When one of the dominant harmonics, the fundamental frequency and its harmonic components, is close to the first formant frequency, the effect of the source-filter interaction can induce voice register transition, in which the vocal-fold vibration becomes unstable and the pitch jumps drastically. In this study, we investigated the relationship between the dominant harmonics and the first formant frequency in the modal-falsetto transition to clarify the effect of source-filter interaction. While five subjects performed rising glissandi with /a/ and /i/ vowels, we simultaneously measured their vocal-fold vibration by using electroglottography and the acoustic response from the vocal tract by using the external acoustic excitation method. We analysed temporal patterns of the fundamental and the first formant frequencies in the transition section. We found that the fundamental frequency was regularly in the vicinity of the first formant frequency for the /i/ vowel. Additionally, for the /a/ vowel, it was occasionally observed that the second or third harmonic component was close to the first formant frequency in the transition. Our results indicate that the source-filter interaction is a common factor of the modal-falsetto transition for the participants.

5aSC4. Rhythm segment constitution showing regular periodicity. Shizuka Nakamura (Graduate School of Informatics, Kyoto Univ., 36-1 Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan, shizuka@sap.ist.i.kyoto-u.ac.jp)

To verify the possibility of regular periodicity of English rhythm, each sentence was divided into respective rhythm segments and the properties of its durations were analyzed. Rhythm segment (RhySeg) was defined as a segment including one syllable with a primary/secondary stress to which an adjacent unstressed syllable(s). The following locations of a stressed syllable in RhySeg were compared: forward, semi-forward, middle, semi-back, and back. To reflect the perceptual effect to the RhySeg constitution, the following factors to equally compress all of the unstressed syllables were compared: 0.1-1.0 at an interval of 0.1. To find the RhySeg constitution showing regular periodicity, not only the degree of concentration of the distribution, but the degree of closeness between RhySeg with a secondary stress and 1/2 of that with a primary stress, whose engagement on regular periodicity was indicated in previous studies, was applied as a criterion. Comparative experiments showed the best when the stressed syllable located semi-forward or forward, and the factor was 0.7. Furthermore, the result of harmonic analysis and resynthesis applied to the time function of the average duration of syllables in a sentence indicated periodicity is consisted of the combination of the fundamental and its second harmonic components.