

Session 5aSCc**Speech Communication: General Topics in Speech Communication IV (Poster Session)**

Linda Polka, Cochair
McGill University

Christine Mooshammer, Cochair
Haskins Lab and MIT Research Lab of Electronics

Suzanne Boyce, Cochair
Department of Communication Sciences and Disorders, University of Cincinnati

All posters will be on display from 10:20 a.m. to 12:00 a.m. To allow contributors an opportunity to see other posters, contributors of odd-numbered papers will be at their posters from 10:20 a.m. to 11:10 a.m. and contributors of even-numbered papers will be at their posters from 11:10 a.m. to 12:00 p.m.

Contributed Papers

5aSCc1. Call early in the evening on a spring day. Maxine Eskenazi (Language Technologies Institute Carnegie Mellon University, 4619 Newell Simon Hall, 5000 Forbes Ave, Pittsburgh, PA 15213, USA, max@cs.cmu.edu), Antoine Raux (Language Technologies Institute Carnegie Mellon University, 4619 Newell Simon Hall, 5000 Forbes Ave, Pittsburgh, PA 15213, USA, antoine@cs.cmu.edu)

The CMU Let's Go Spoken Dialogue System has been used daily for about three years to answer calls to the Pittsburgh Port Authority for bus information in the evening and on weekends. This has resulted in a database of over 50 000 spoken dialogues as of January 2008, one of the largest publicly available sets of this type of data. While retraining the system with part of this data, it became apparent that there are times of the day, of the week and of the year when the average number of successful calls is significantly higher. We will present evidence, using these three measures of time (hour, day of week, month of year) and criteria such as signal-to-noise ratio, estimated success rate, number of turns per dialogue, number of non-understandings per dialogue, and barge-in rate to detect the regular, predictable appearance of high and low success rates and to suggest methods for palliating this effect in order to increase overall dialogue success rates.

5aSCc2. Duration modeling for English letters embedded in Chinese speech. Wen-Hsing Lai (National Kaohsiung First University of Science and Technology, No. 2, Jhuoyue Rd., Nanzih District, 811 Kaohsiung, Taiwan, lwh@ccms.nkfust.edu.tw)

A review of existing multilingual TTS (text-to-speech) systems shows that the secondary language inserted into the primary language sounds more like isolated individual words in an alien language environment and not congruous with the primary language's prosody. Since the letter-by-letter spelling of English words or acronyms appears in Chinese speech quite often, a duration modeling approach for English letters embedded in Chinese speech is proposed to make the English congruous with the primary language's tempo. It takes several major factors as additive factors and estimates all model parameters by an EM (expectation-maximization) algorithm. Experimental results showed that the standard deviation of the duration from the test set was greatly reduced from 59.82 to 9.37 ms by the duration modeling while eliminating effects from factors. The root mean squared error between the original and estimated durations was 9.35 ms for the open tests. Experimental results have confirmed its effectiveness on isolating several main fac-

tors that seriously affects the duration. Moreover, the estimated value of the factors agreed well to our prior linguistic knowledge. Besides, the hidden state labels produced by the EM algorithm were linguistically meaningful.

5aSCc3. Modeling perception of breathy voice quality in vowels using data obtained in an adaptive matching task. Rahul Shrivastav (University of Florida, Dauer Hall, P.O. Box 117420, Dept. of Communication Sciences and Disorders, Gainesville, FL 32611, USA, rahul@csd.ufl.edu), Arturo Camacho (University of Florida, Dauer Hall, P.O. Box 117420, Dept. of Communication Sciences and Disorders, Gainesville, FL 32611, USA, arturocl@yahoo.com), Sona A. Patel (University of Florida, Dauer Hall, P.O. Box 117420, Dept. of Communication Sciences and Disorders, Gainesville, FL 32611, USA, sona09@ufl.edu), David A. Eddins (University of Rochester, Department of Otolaryngology, Rochester, NY 14618, USA, David_Eddins@URMC.Rochester.edu)

A computational model of breathy voice quality has been developed based on perceptual judgments of synthetic /a/ vowels [Shrivastav, Camacho, and Patel, *JASA* **120**(5), 3248]. This model is based on the ratio of the loudness of aspiration noise ("noise loudness;" NL) to the loudness of the periodic components of the vowel when masked by the aspiration noise ("partial loudness;" PL) [Shrivastav and Sapienza, *JASA*, **114**(1), 2218-2224 (2005)]. Results showed that the model accounted for a large amount of variance in perceptual ratings of training data (R-square=0.92) but not for testing data (R-square=0.59). It was hypothesized that this reduction was partly related to the use of a magnitude estimation task to obtain perceptual judgments, since magnitude estimates are biased by the range of the quantity measured and are prone to inconsistencies in how subjects assign numbers to items [Poulton, 1989; Guilford, 1954]. To minimize such biases, a matching task may be used to obtain ratio level estimates of breathiness [Patel, Shrivastav and Eddins, *JASA*, **119**(5), 3340 (2006)]. The present study describes a model to predict breathy voice quality derived from perceptual judgments obtained using a matching task.

5aSCc4. The effect of training in noise on foreign language consonant acquisition. M. Luisa Garcia Lecumberri (University of the Basque Country, Paseo de la Universidad 5, Facultad de Filologia, 01006 Vitoria, Spain, garcia.lecumberri@ehu.es), Martin Cooke (Sheffield University, Computer Science Department, Regent Court, 211 Portobello St., S1 4DP Sheffield, UK, m.cooke@dcs.shef.ac.uk)

Formal exposure to second language sounds normally takes place in clean, laboratory conditions, but at issue is the transfer of such learning to everyday situations. Categories learned in non-natural settings may be fragile, raising the question as to whether learning in noise leads to greater category robustness. The current study compared two groups of Spanish learners of English who were trained in either quiet or noise backgrounds on a 24 consonant discrimination task. Learners' performance was measured in pre- and post-tests and monitored over the course of nine weeks' training. Both groups showed continual improvement during training with similar overall gains of 6-7 percentage points, suggesting that training in adverse conditions is as effective as in quiet. Tests involving consonant identification in quiet and noise revealed no significant differences in pre-post improvement between the two groups. However, voiceless obstruents benefited more from training in noise while voiced obstruents experienced more improvement when trained in a quiet background regardless of the testing condition (quiet vs noise). The noise-trained disadvantage is consistent with masking of voicing in noise, while the quiet-trained deficit for voiceless obstruents may derive from overgeneralisation during the acquisition of new contrasts.

5aSCc5. Analyze effects of the flow on the vocalic reduction and the coarticulation in sequences CV of pharyngeal Arabic. Leila Falek (USTHB, Faculté d'électronique et d'Informatique, 16111 Algiers, Algeria, lilalcppts@yahoo.fr), Othmane Bouferroum (USTHB, Faculté d'électronique et d'Informatique, 16111 Algiers, Algeria, lilalcppts@yahoo.fr), Amar Djeradi (USTHB, Faculté d'électronique et d'Informatique, 16111 Algiers, Algeria, lilalcppts@yahoo.fr)

The degree of coarticulation and the vocalic reduction (RV) are indices related to a good engine control (Gay, 1978). Fowler (1998) explains why locus equation (LE) is used to characterize, at the same time, the place of articulation and the degree of coarticulation between consonants and vowels: a strong slope ($m=1$) indicates a maximum coarticulation between consonants and vowels (i.e., minimal resistance of the coarticulation), while a weak slope ($m=0$) indicates absence of coarticulation between consonants and vowels (maximum resistance of the coarticulation). The bond between the degree of coarticulation and the RV can be explained according to the linear relation between $F2_{onset}$ and $F2_{milieu}$: the modifications of values of $F2_{milieu}$ will affect those of $F2_{onset}$ and consequently those of the slopes of. In this study, the analysis of the vocalic reduction and slopes of the equations of locus, carried out on CV (extracts starting from sentences) in standard Arabic pronounced by speakers having different mother tongues (near to Arabic standard and very far away from standard Arabic), and at speed of variable elocution, revealed a vocalic reduction and a variation of the slope of the locus equation, specific to each speaker, who seems to be related to his mother tongue. El Tamimi (2006) carried out a similar study with normal flow, in dialectical Arabic and in French, with normal flow, an influence of the mother tongue showed on the vocalic reduction and the slope of the equation of locus.

5aSCc6. A comparison between two cases of voicing neutralization: Final devoicing and voicing assimilation in obstruent clusters. Rebeka Campos-Astorkiza (Ohio State University, Dept. Spanish & Portuguese, 298 Hagerty Hall, 1775 College Rd, Columbus, OH 43210, USA, camposastorkiza.1@osu.edu)

This study compares two processes that result in voicing neutralization, namely final obstruent devoicing and regressive voicing assimilation in obstruent clusters in Lithuanian. The goal is to find out whether both neutralizing environments present similar acoustic patterns. Acoustic data was collected from native speakers of Lithuanian in order to analyze /k/ and /g/ in word final position (final devoicing) and in the word medial clusters /gs/, /ks/, /kz/ and /gz/ (voicing assimilation). Three main acoustic cues to obstruent voicing were measured: obstruent closure duration, voicing during closure and duration of preceding vowel. The statistical results show that, although both final devoicing and voicing assimilation result in incomplete voicing neutralization, they use the acoustic cues differently. In the case of voicing assimilation, there are statistically significant differences in closure duration, voicing during closure and preceding vowel duration between the stops in /ks, gz/ and their assimilated counterparts in /gs, kz/. As for final devoicing, underlyingly voiceless obstruents and devoiced obstruents show differences in their closure duration and voicing during closure but not in

their preceding vowel duration. These results suggest that final devoicing and voicing assimilation have different acoustic realizations in Lithuanian, supporting their analysis as two different processes.

5aSCc7. Effectiveness of prosodic features in the perception of mandarin utterance-final particles. Bin Li (Department of Chinese, Translation & Linguistics, City University of Hong Kong, 83 Tat Chee Ave. Kowloon Tong, 220 Hong Kong, China, binli2@cityu.edu.hk), Jinping Zhu (Tsinghua University, 104 Main Academic building, 100084 Beijing, China, yx@abcd.edu.cn)

Utterance-final particles (UFPs) in Mandarin are regarded as discourse markers, whose "elusive" meanings and functions have been investigated in different perspectives such as pragmatics and semantics. Our study focuses on two most frequently used UFPs: NE and A, both of which can occur in identical utterances. The difference lies in that NE implies a shared common ground between the speaker and the hearer, whereas such commonness lacks for A. A preliminary acoustic analysis found that prosodic features of an utterance were more salient when a UFP was not used. The current study aims at the effectiveness of these prosodic features in perceiving these particles. Native speakers of Mandarin are asked to listen to sentences, which may originally contain a UFP that is removed. Their tasks are to decide whether a UFP, and which one, is needed in a given sentence. It is hypothesized that without textual cues, speakers will rely on prosodic features in their judgment; therefore, they will perform better in sentences where UFPs are not used originally and where prosodic cues are more salient.

5aSCc8. The effect of learning on listening to ultra-Fast speech. Takuya Nishimoto (Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, 113-8656 Tokyo, Japan, nishi@hil.t.u-tokyo.ac.jp), Yukika Kariya (Department of Communication, Tokyo Woman's Christian University, 2-6-1 Zenpukuzi, Suginami-ku, 167-8585 Tokyo, Japan, g04c043@cis.twcu.ac.jp), Takayuki Watanabe (Department of Communication, Tokyo Woman's Christian University, 2-6-1 Zenpukuzi, Suginami-ku, 167-8585 Tokyo, Japan, nabe@cis.twcu.ac.jp)

We investigated the intelligibility of ultrafast speech which may be used for screen reader for persons with visual disability. The subjects were 35 women who are university students and are not visually/hearing impaired. They were divided into four groups and they listened to 150 words with the speed of approximately 20 morae/s. The vocabulary contained the tasks of high and low familiarity words, and the orders of tasks were different by the groups. Four morae Japanese words from the FW03 database were used as the vocabulary of the recall test. As a result, significant learning effect was observed in cases where the subject listened the high familiarity words in succession. This indicates that the learning effect to the ultrafast speech is promoted when the mental lexical access is easy. We also investigated the mental workloads of the listening task using the NASA-TLX method. As the results, significantly high workload scores were observed at the listening of low familiarity words. The results also suggested that the mental workload decreases when the subject was aware that the mental lexical access was difficult.

5aSCc9. Speech segmentation in monolingual and bilingual infant learners of Canadian English and Canadian French. Linda Polka (McGill University, School of Communication Sciences and Disorders, Beatty Hall, 1266 Pine Avenue West, Montreal, QC H3G 1A8, Canada, linda.polka@mcgill.ca), Jennifer Proulx (McGill University, School of Communication Sciences and Disorders, Beatty Hall, 1266 Pine Avenue West, Montreal, QC H3G 1A8, Canada, Jennifer.proulx@mail.mcgill.ca), Megha Sundara (UCLA Department of Linguistics, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA, megha.sundara@humnet.ucla.edu)

Speech segmentation skills emerging in infancy are influenced by the infant's native language. English infants favor a stress-based strategy (Jusczyk et al., 1999) whereas French infants favor a syllable-based strategy (Nazzi et al., 2006). Cross-linguistic findings from our lab, based on between group comparisons, show that monolingual 8-month-olds learning either Canadian English or Canadian French segment bisyllabic words in their native lan-

guage but not in a rhythmically-different non-native language. Canadian French infants fail to segment Canadian English; Canadian English infants fail to segment Canadian French. Bilingual infants (exposed to both languages) were tested in each language on separate days. Although they appear to segment in both languages, attrition was high in the second test session. To assess cross-language segmentation more directly we tested monolingual and bilingual 8-month-olds using a task designed to assess segmentation in both languages in the same infant within a single test session. Findings confirm that monolingual 8-month-olds fail to segment bisyllabic words in a rhythmically-different non-native language. Preliminary data show that bilingual 8-month-olds segment only in the language that is favored in their language input. Thus, in early stages of speech processing all infants appear to develop speech segmentation strategies that are optimal for one language.

5aScC10. Clear speech intelligibility and accentedness ratings for native and non-native talkers and listeners. Rajka Smiljanic (Northwestern University, Department of Linguistics, 2016 Sheridan Road, Evanston, IL 60208, USA, rajka@northwestern.edu), Ann Bradlow (Northwestern University, Department of Linguistics, 2016 Sheridan Road, Evanston, IL 60208, USA, abradlow@northwestern.edu)

This study investigated how native language background (L1) interacts with speaking style in determining levels of speech intelligibility. In four experiments, we explored whether native and non-native hyper-articulation clear speech strategies provide similar intelligibility benefits for native and high proficiency non-native listeners. The sentence-in-noise perception results revealed that native speech was preferred over non-native speech by both listener groups even when non-native talkers and listeners shared the same L1. Clear speech was shown to be beneficial for both the native and fluent non-native listeners. However, non-native clear speech enhanced intelligibility less than native clear speech, supporting the hypothesis that clear speech production strategies involve enhancement of language-specific phonological contrasts. In order to assess the relationship between objective intelligibility measures and subjective accentedness ratings, we obtained accentedness ratings of native and non-native conversational and clear speech by native and non-native listeners. The results showed that objective intelligibility and subjective accentedness were independent. Overall, these results provide strong evidence that clear speech involves language-specific modifications. Nevertheless, native and high proficiency non-native clear speech modifications are generally helpful for both native and high proficiency non-native listeners even when the objective intelligibility and subjective accentedness levels diverged for various listener and talker groups.

5aScC11. Vietnamese monophthong vowel production by native speakers and American adult learners. Matthew Winn (University of Maryland College Park, Department of Hearing & Speech Sciences, 0100 Lefrak Hall, College Park, MD 20742, USA, mwinn@hesp.umd.edu), Allison Blodgett (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, ablodgett@casl.umd.edu), Jessica Bauman (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, jbauman@hesp.umd.edu), Anita Bowles (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, abowles@casl.umd.edu), Lykara Charters (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, lykarac@mac.com), Anton Rytting (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, crytting@casl.umd.edu), Jessica Shamoo (University of Maryland College Park, Center for Advanced Study of Language, 7005 52nd Ave, College Park, MD 20742, USA, jshamoo@casl.umd.edu)

This study provides new native speaker data regarding the vowel space and duration contrasts of Vietnamese monophthongs. These data add to existing empirical accounts and raise questions about descriptions aimed at adult language learners. In addition, specific acoustic measures target difficulties that English-speaking adults encounter when producing Vietnamese vowels. These measures include the rounding distinction for back vowels, height distinction for central vowels, and duration distinction between short

and long vowels. Results show that adult learners exhibit greater variability and overall less accurate production than native speakers for the high-back unrounded vowel *u*, with slightly better performance for the mid-back unrounded vowel *o*, perhaps because it is similar to English / *ʌ* /. Additionally, these learners demonstrate varying levels of success with regard to the production of central vowel height but generally fail to distinguish the two vowel duration categories exhibited by native speakers.

5aScC12. The role of segmental and intonational cues in dialect discrimination. Chad Vicens (UCLA Department of Linguistics, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA, cvicens@humnet.ucla.edu), Megha Sundara (UCLA Department of Linguistics, 3125 Campbell Hall, Los Angeles, CA 90095-1543, USA, megha.sundara@humnet.ucla.edu)

Research indicates that adult listeners are able to use intonation to discriminate between two languages when one of the languages is familiar (Ramus and Mehler, 1999; Pijper, 1983). In this paper, we test adults to determine whether they use segmental or intonational cues to distinguish their native dialect from a foreign one. In three experiments, American English listeners were asked to categorize American and Australian English sentences when (a) segmental and supra-segmental cues are available, (b) sentences are re-synthesized with flat intonation, leaving only segmental cues, (c) segmental information is stripped away, leaving only intonation. Results will be discussed in the context of infant research demonstrating that five month olds are able to distinguish different dialects of the same language (Nazzi, Jusczyk, and Johnson 2000).

5aScC13. Effects of syllable structure on reaction times in a delayed naming task. Christine Mooshammer (Haskins Lab and MIT Research Lab of Electronics, 300 George Street Suite 900, New Haven, CT 06511, USA, tine@haskins.yale.edu), Louis Goldstein (Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA, goldstein@haskins.yale.edu), Mark Tiede (Haskins Lab and MIT Research Lab of Electronics, 300 George Street Suite 900, New Haven, CT 06511, USA, tiede@haskins.yale.edu), Hosung Nam (Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA, nam@haskins.yale.edu), Man Gao (Haskins Laboratories, 300 George St., Suite 900, New Haven, CT 06511, USA, gao@haskins.yale.edu)

Syllable complexity has been found to affect the time the speaker needs for planning and initiating utterance production. Shorter latencies for complex onsets (CCV) as compared to simple onsets (CV) have been explained by effects of segment-specific biomechanical constraints at the level of motor execution, and by neighborhood density at the planning level. Within the framework of Articulatory Phonology, shorter planning latencies for CV syllables (compared to VC) have been attributed to quicker stabilization for tighter gestural coupling hypothesized for in-phase coupling of the onset consonant and release with the vowel. We attempted to test both onset complexity (C vs CC) and coda complexity (open vs. closed syllables) within a single experiment, so that we could evaluate the relative magnitudes of these effects and uncover potential interactions. To do so, American English monosyllabic words varying in syllable structure were presented using a delayed naming paradigm. The results replicated both effects, showing additive effects of roughly comparable magnitudes. Whether these effects are due to planning or motor execution will be tested by further simple naming and picture naming experiments.

5aScC14. Breathy and whispery voicing in White Hmong. Sean Fulop (California State Univ. Fresno, 5245 N. Backer Ave., Linguistics PB92, Fresno, CA 93705, USA, sfulop@csufresno.edu), Chris Golston (California State Univ. Fresno, 5245 N. Backer Ave., Linguistics PB92, Fresno, CA 93705, USA, chrisg@csufresno.edu)

Two phonatory strategies are acknowledged to involve a superposition of periodic vocal cord vibration with excessive airflow generating turbulent noise. Breathy voice has the vocal folds vibrating along their entire length, but incomplete closure allows a constant airflow. Whispery voice has the vocal folds vibrating modally along an anterior segment, while the arytenoids maintain a posterior hole which allows a constant airflow. In linguistic phonetics, these are standardly regarded as variant strategies of a

single “breathy voice” specification, and have never been noted to coexist in one language in an important way. They do in White Hmong. One of the seven “tones” of White Hmong is breathy. Syllables bearing this tone display breathy voicing (of the first type) on the vowel. White Hmong also has a stop [d^h], which sounds like whispery voicing during the release phase. In this study, the breathy tone is shown to have a higher H1 amplitude relative to H2 (compared to modal voice), while the whispery voiced stop is even more extreme in this measure. Whispery voiced stops are also shown to display less harmonicity than the breathy tone, which in turn has less harmonicity than modal voicing in a similar syllable.

5aScC15. Developing a bilingual communication aid for a Japanese ALS patient using voice conversion technique. Akemi Iida (School of Media Science, Tokyo University of Technology, 1404-1, Katakura-cho, Hachioji, 192-0982 Tokyo, Japan, ake@media.teu.ac.jp), Shimpei Kajima (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, s.kajima@gmail.com), Keichi Yasu (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, k-yasu@sophia.ac.jp), John M. Kominek (Language Technologies Institute, Carnegie Mellon University, 5000 Forbes Ave. West, Pittsburgh, PA 15213, USA, jkominek@cs.cmu.edu), Yasuhiro Aikawa (1-36-2, Shinjuku, Shinjuku-ku, Shinjuku dai-nana-biru 7F, 160-0022 Tokyo, Japan, aikawa@loopedpicture.com), Takayuki Arai (Dept. of Electrical and Electronics Engineering, Sophia University, 7-1 Kiyoi-cho, Chiyoda-ku, 102-8554 Tokyo, Japan, arai@sophia.ac.jp)

A bilingual communication aid for a Japanese amyotrophic lateral sclerosis (ALS) patient has been developed. From our previous research, a corpus-based speech synthesis method was ideal for synthesizing speech with voice quality identifiable as the patient’s own. However, a recording of a large amount of speech, which is a burden for the patient, is required for such system. In this study, a voice conversion technique was applied so that a smaller amount of recording is needed for synthesis. An English speech synthesis system with the patient’s voice was developed using Festival, a corpus-based speech synthesizer with voice conversion technique. Two methods for Japanese speech synthesis were attempted using HTS toolkit. The first used an acoustic model built from all 503 recordings of the patient. The second used an acoustic model built from 503 wavefiles of which voice was converted to the patient’s from a native speaker’s. The latter method requires fewer recordings of the patient’s. The result of the perceptual experiment showed that the voice synthesized with the latter was perceived to have a closer voice quality to the patient’s natural speech. Last, GUI on windows was developed for the patient to synthesize speech by typing in the text.

5aScC16. Individual differences in perception of emotions from nonsense speech. Sona A. Patel (University of Florida, Dauer Hall, P.O. Box 117420, Dept. of Communication Sciences and Disorders, Gainesville, FL 32611, USA, sona09@ufl.edu), Rahul Shrivastav (University of Florida, Dauer Hall, P.O. Box 117420, Dept. of Communication Sciences and Disorders, Gainesville, FL 32611, USA, rahul@csd.ufl.edu)

Identification of the acoustic cues used to perceive emotions in speech is important for a number of applications including rehabilitation, natural speech modeling, and speech synthesis. In a recent experiment, Patel, Shrivastav, Harnsberger, and Shrivastav (2007) found that a four-dimensional solution accounted for 90% of the variance in similarity judgments for 19 emotional categories in nonsense speech. This solution was determined for averaged judgments across twelve listeners. The present study investigated individual differences in the perception of emotions for speech devoid of semantic information but rich in suprasegmental cues. Six male and six female listeners participated in a same-different discrimination test of a set of nonsense sentences produced in nineteen emotional contexts by two actors. Nonsense sentences were used in order to avoid any biases caused by semantics. The perceptual distance between each stimulus pair was computed in terms of d' values for each listener. These distances were submitted to a multidimensional scaling analysis using the INDSCAL

algorithm. The INDSCAL analysis reports the best fitting solution for all listeners as a group, along with the weights assigned to each dimension by individual listeners. The results of this analysis will be presented.

5aScC17. Phonological representation of negation in Sakizaya. Wen-Chi Shen (Grad. Inst. of Linguistics, National Taiwan Univ., 1, Roosevelt Rd. Sec. 4, 106 Taipei, Taiwan, r94142007@ntu.edu.tw), Wen-Yu Chiang (Graduate Institute of Linguistics, Grad. Inst. of Linguistics, National Taiwan Univ., 1, Roosevelt Rd. Sec. 4, 106 Taipei, Taiwan, countryroad0509@hotmail.com)

This paper aims at investigating phonological representation of negation in Sakizaya, an endangered Formosan language in Taiwan. Defying Yaeger-Dror’s “Cognitive prominence principles” (2003), Chiang (2006) examines negators in Saisiyat, a SVO language as English, and finds the sentential subjects are more prominent acoustically rather than negators. Thus, I will adopt Chiang’s modal (2006) and analyze phonological representation of six negators in Sakizaya to see where the pitch accent falls. Besides, if falling on the negators, I will see whether it is because of either cognitive prominence properties or its sentence-initial position? Six informants (three male and three female), from 50-74 years old, participated in the experiment, recorded by DAT, analyzed by Praat and digitized at a 22050-Hz sampling rate. Affirmative and negative sentences are asked in Mandarin while the informants are instructed to translate them into Sakizaya. Besides, they are asked to produce iterative -ay sentences since the negator ca’ay is the most frequently used negator with higher prominent syllable -ay. The results show negators are more prominent in Sakizaya compared to sentential subjects. I conclude it is word order that influences the prominence of pitch contour since Saisiyat is an SVO language and Sakizaya a VSO one.

5aScC18. The effect of vowel duration on formant frequencies - data from Hakka Chinese. Wai-Sum Lee (Phonetics Lab, Dept. of Chinese, Translation and Linguistics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Hong Kong, Hong Kong, w.s.lee@cityu.edu.hk), Eric Zee (Phonetics Lab, Dept. of Chinese, Translation and Linguistics, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon Hong Kong, Hong Kong, eric.zee@cityu.edu.hk)

In the target undershoot model, vowel duration is considered as the main determinant of vowel reduction, resulting in undershoot in formant frequencies relative to the “bull’s-eye formant pattern” (Lindblom, 1963). The model predicts more schwa-like formant frequencies as vowel duration is shortened. In Hakka Chinese, vowel phonemes /i e a o u/ may be realized as long [i: e: a: o: u:] in the CV: syllables and short [i e a o u] in the CVS syllables (S=a stop consonant). Results of a formant frequency analysis of the long and short vowel sets in Hakka Chinese from ten male and ten female speakers show that (i) in all the short vowels there is a displacement of vowel formant frequencies away from the target frequencies associated with the long vowels; and (ii) relative to the target frequencies, F1F2 do not become more schwa-like for all the short vowels, with [u] being the only vowel that undergoes centralization in the F1F2 plane and with [i e o a] displaying a significant rise in F1, [i e] an insignificant decrease in F2, and [o] an insignificant increase in F2. It appears that the Hakka data do not support the target undershoot model.

5aScC19. An experimental and modeling study of anticipatory coarticulation in VCV sequences. Liang Ma (Laboratoire Parole et Langage, 29, Avenue Robert Schuman, 13621 Aix-en-Provence, France, liang.ma@lpl.univ-aix.fr), Pascal Perrier (ICP/GIPSA-lab, INPG, 46 Avenue Félix Viallet, 38031 Grenoble Cedex 01, France, Pascal.Perrier@gipsa-lab.inpg.fr), Jianwu Dang (Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, 923-1292 Ishikawa, Japan, jdang@jaist.ac.jp)

Anticipatory coarticulation within VICV2 sequences is studied for EMMA data in French and in Mandarin Chinese. The corpus was designed consistently for both languages. V1 and V2 were /i, a, u/ and C was either /t/ or /k/. The influences of V2 on tongue position of V1 and of C were analyzed for three French and two Chinese speakers. For French speakers vowel V2 influences the whole sequence VICV2, while its influence is strictly limited to the syllable CV2 for Mandarin speakers. These suggested