$F2$ was generated from natural productions of ‘‘bid’’ and ‘‘bed’’ using the synthesizer STRAIGHT. These tokens were spliced into a nine-word sentence at different positions that also contained two other test words, one each from pairs ‘‘cot/cut’’ and ‘‘hack/hawk.’’ Listeners were asked to identify the three words they heard in the sentence. Listeners also identified whether ‘‘bid’’ or ‘‘bed’’ was heard when only the isolated tokens were presented. Thresholds to detect a change from ‘‘bid’’ were obtained from psychometric functions fit to the data. Thresholds were similar for the sentence and word-only tasks. Overall, thresholds in both classification tasks were worse than those from the 2AFC tasks. Results will be discussed in terms of the relation between these discrimination thresholds, vowel identification, and vowel spaces. [Work supported by NIHDCD-02229.]

2pSC4. Speech enhancement based on modified phase-opponency detectors. Om D. Deshmukh and Carol Y. Espy-Wilson (Dept. of Elec. and Computer Eng. and Inst. for Systems Res., Univ. of Maryland, College Park, MD 20742)

A speech enhancement algorithm based on a neural model was presented by Deshmukh *et al.*, [149th meeting of the Acoustical Society America, 2005]. The algorithm consists of a bank of Modified Phase Opponency (MPO) filter pairs tuned to different center frequencies. This algorithm is able to enhance salient spectral features in speech signals even at low signal-to-noise ratios. However, the algorithm introduces musical noise and sometimes misses a spectral peak that is close in frequency to a stronger spectral peak. Refinement in the design of the MPO filters was recently made that takes advantage of the falling spectrum of the speech signal in sonorant regions. The modified set of filters leads to better separation of the noise and speech signals, and more accurate enhancement of spectral peaks. The improvements also lead to a significant reduction in musical noise. Continuity algorithms based on the properties of speech signals are used to further reduce the musical noise effect. The efficiency of the proposed method in enhancing the speech signal when the level of the background noise is fluctuating will be demonstrated. The performance of the improved speech enhancement method will be compared with various spectral subtraction-based methods. [Work supported by NSF BCS0236707.]

2pSC5. Knowledge-based formant tracking with confidence measure using dynamic programming. Sandeep Manocha and Carol Y. Espy-Wilson (Inst. for Systems Res. and Elec. and Computer Eng. Dept., Univ. of Maryland, College Park, MD 20742, smanocha@umd.edu)

In this study, we report on refinements to a formant tracking technique originally reported in [Xia *et al.*, ICSLP (2000)]. The formant tracker operates in two phases. First, it finds optimal formant track estimates in oral sonorant regions by imposing frequency continuity constraints using dynamic programming. Second, post-processing is performed to make the estimates more robust and accurate, and to extend formant tracks in nasal and obstruent regions. In recent work, we have improved on our initial estimates of the formants by combining the outputs from a 12th-order LPC analysis and a 16th-order LPC analysis. Additionally, we have added a confidence measure for each formant track in each frame. The confidence measure is based on formant continuity, competing formants, short-time energy, context information, and formant information over the entire utterance. The experiments show that most of the tracking errors are associated with a low confidence value, while the correct formants have high confidence values. The performance of the algorithm in the sonorant regions was tested using randomly selected male and female speech from the TIMIT database. [Work supported by NSF Grant No. BCS0236707.]

2pSC6. Investigating an optimum suppression rate of steady-state portions of speech that improves intelligibility the most as a pre-processing approach in reverberant environments. Nao Hodoshima and Takayuki Arai (Dept. of Elec. and Electron. Eng., Sophia Univ. 7-1 Kioi-cho, Chiyoda-ku, Tokyo 102-8554 Japan, n-hodosh@sophia.ac.jp)

Steady-state suppression has been proposed as a pre-processing approach to improve speech intelligibility in reverberant environments [Arai *et al.*, Acoust. Sci. Technol. **23**, 229–232 (2002)]. The goal of this work is to find the suppression rate that improves intelligibility the most. This is done by exploring the relationship between the suppression rates of steady-state portions of speech signals in a reverberant environment and syllable identification. This needs to be investigated because the optimum suppression rate might depend on the amount of overlap-masking, which in turn is determined by the reverberation. A syllable identification test was conducted with 21 normal-hearing listeners in two reverberant environments and six suppression rates. Results show a significant improvement in intelligibility for a 40% and a 50% suppression rate for a reverberation time of 1.1 s and a 60% suppression rate for a reverberation time of 1.3 s. Findings confirm that intelligibility is affected by both suppression rate and reverberation time. Results also show that an optimal suppression rate might not be derived from reverberation time alone. The spectral balance of the room response should also be taken into consideration. [Research supported by JSPS (16203041 and 176911).]

2pSC7. The relationship between the articulation index spectrogram and off-diagional confusion matrix performance-intensity functions. Bryce Lobdell and Jont Allen (Beckman Inst., MC 251, 405 North Mathews, Urbana, IL 61801)

Confusion matrices (CM), similar to those of Miller Nicely 1955, as a function of the speech-to-noise ratio (SNR), have been collected for 16 consonants and 4 vowels, from 18 talkers and 14 listeners. The goal is to quantify perceptual features (denoted events) in the acoustic waveforms, from the articulation index (AI) spectrogram, and the off-diagonal ‘‘PI-functions’’ $P_{h,s}(\mathrm{SNR})$, indices $0 \leqslant s,h \leqslant 64$ ($h=\mathrm{hear} \neq s=\mathrm{spoken}$). Near ‘‘threshold’’ SNRs (i.e., where the off-diagional PI functions show confusion slightly greater than 1/64), the AI spectrograms show interesting transitions. The AI spectrogram is closely related to an information-theory measure called the channel capacity (and to jnd’s). Examples of events in the ‘‘CCgrams’’ will be given for a variety of speech sounds at threshold SNRs. A cool sound-morphing demonstration will be presented, where the dominant confusion depends on SNR.

2pSC8. Speech intelligibility criteria for PA system qualification in large aircraft. Gopal Mathur (The Boeing Co., MC: H013-B308, Huntington Beach, CA 92647), Manolis Tsangarakis, David Lotts, Kenneth Barry (The Boeing Co., MC: C052-0066, Long Beach, CA 90807), and Naval Agarwal (The Boeing Co., MC: 6M6-31, Renton, WA)

The military aircraft test and evaluation community uses the Modified Rhyme Test (MRT) test for evaluating speech intelligibility of a communication system in an aircraft. An MRT test is conducted for the flight member/mission areas specified in the MIL/ANSI/ISO specifications. These standards specify that the MRT shall be conducted in a noise field that simulates the actual aircraft noise at the member’s position. Since the method is based on the perception of words by listeners, there are no limitations in respect of the characteristics of the sound system or those of the environment. Modern communication systems, however, incorporate several new electronic designs, e.g., automatic gain control, analog-to-digital (ADC) conversions which can introduce non-linearity in the system. This paper will present results of investigations conducted in a large military aircraft where such features in conjunction with high background noise levels and moving speaker were considered and evaluated. Investigations were focused on ‘‘relative’’ measurements between different system configurations. The MRT was found to be the most reliable method for final verification of PA systems if performed as specified in the ANSI