The Journal of the Acoustical Society of America

Vol. 143, No. 3, Pt. 2 of 2, March 2018 www.acousticalsociety.org

175th Meeting
Acoustical Society of America

Hyatt Regency Minneapolis
Minneapolis, Minnesota
7–11 May 2018

Table of Contents on p. A5

Published by the Acoustical Society of America through AIP Publishing LLC

obtained from a set of different acoustic representations, in particular through the STRF. We observed that distances computed from spectrotemporal modulation representations provide the best correlation with the perceptual results across the seven timbre spaces. Finally, we highlighted the parts of the representations contributing the most to the correlation suggesting new insights into the underlying perceptual metrics. [Supported by Canada Research Chair, NSERC (RGPIN-2015-05208, RGPAS-478121-15), (RGPIN-262808-2012), and EU MSCf (Project MIM, H2020-MSCA-IF-2014, GA no. 659).]

1pPPb2. Neural coding of perceptual temporal asymmetry for sounds with rising vs. falling intensity envelopes in non-attentive and attentive listening conditions. Bing Cheng (English Dept. & Inst. for Lang., Cognition and Brain Sci., Xi'an Jiaotong Univ., 28 Xianning St. West, School of Foreign Studies, Xi'an, Shaanxi 710049, China, bch@mail.xjtu.edu.cn), Yang Zhang (Speech-Language-Hearing Sci., Univ. of Minnesota, Minneapolis, MN), Keita Tanaka (School of Sci. and Eng., Tokyo Denki Univ., Hatoyama, Saitama, Japan), Robert S. Schlauch (Speech-Language-Hearing Sci., Univ. of Minnesota, Minneapolis, MN), and Toshiaki Imada (Inst. for Learning and Brain Sci., Univ. of Washington, Seattle, WA)

Behavioral studies have shown that ramped sounds are judged to be louder and longer than damped sounds. Here we employed magnetoencephalography (MEG) to examine cortical processing of the perceptual temporal asymmetry. The participants were 6 normal-hearing right-handed male adults. The synthesized stimuli included three kinds, pure tone, piano note, and broadband noise with time-reversed intensity envelopes for the ramped vs. damped comparison. Each stimulus was 200 ms. In the non-attentive condition, subjects were instructed to watch a silent movie and ignore the randomly presented tones. In the attentive condition, listeners were required to judge whether the first or the second of paired sounds was longer. The stimuli were presented at 50 dB SL. The behavioral results replicated previous temporal asymmetry findings for all three types of stimuli. There were significant effects of stimulus type and stimulus order in the MEG ON and OFF responses. Despite the stimulus type effect, the ramped tonal stimuli consistently elicited smaller and later N1m responses than the damped controls in both ignore and attend conditions. These data supports distinct cortical mechanisms for coding the stimulus type information and subjective duration information of the ramped vs. damped auditory stimuli.

1pPPb3. Optimal frequency filtering of auscultation sounds. Lukasz Nowak (Inst. of Fundamental Technolog. Res., Polish Acad. of Sci., ul. Pawinskiego 5B, Warszawa 02-106, Poland, lnowak@ippt.pan.pl)

Most of the auscultation sounds do not reveal any significant single-frequency components, and their acoustic energy is concentrated in the low frequency region—up to about 100 Hz, falling even below the threshold of hearing. Such character is determined not only by the vibroacoustic behavior of sources, but mostly by high damping introduced by the sound transmission path through tissues underlying the skin surface. The contained diagnostic information is very subtle, and thus it can be easily masked by internal or external noise sources. Not all of those corrupting signals can be efficiently blocked, hence frequency filtering is the most obvious solution for improving the diagnostic capabilities. Many various filtering strategies and techniques were developed and implemented in both acoustic and electronic stethoscopes, however they are based primarily on (not always correct) intuition and subjective evaluation, without implementation of any accurate and objective measurement means or optimization algorithms. The present study introduces various signal to noise ratio (SNR) measures, applicable for different examination cases. Frequency filtering optimization strategies, for maximizing the values of the introduced coefficients for different heart and lung auscultation sounds, are presented.

1pPPb4. The effects of frequency fine-tuning in hearing impaired phone recognition. Ali Abavisani and Jont Allen (ECE, Univ. of Illinois at Urbana-Champaign, 405 N Mathews Ave., Rm. 2137, Urbana, IL 61801, aliabavi@illinois.edu)

A key factor on correct phone recognition in Normal Hearing (NH) and Hearing Impaired (HI) listeners, is the intensity of primary cue. One can assess this intensity for a given speech sound, by examining it at various Signal to Noise Ratios (SNR) presented to NH listeners, and detect the threshold in which listeners recognized the token at least 90% correct (SNR_{90}). For each token, we have determined the time-frequency window corresponding to correct recognition, as well as the conflicting cues. Two sets of tokens T_1 and T_2 having same consonant-vowels but different talkers with distinct SNR_{90}s had been presented at flat gain (frequency independent) at listeners' most comfortable level (MCL). We studied the effects of frequency fine-tuning of the primary cue by presenting tokens of same consonant but different vowels with similar SNR_{90}s. Additionally, we investigated the role of changing the intensity of primary cue on HI phone recognition, by presenting tokens from both sets T_1 and T_2. This presentation discusses how the frequency an/or intensity of the primary cue changes the confusion pattern of phone recognition for HI listeners, given a flat gain condition at MCL. We will also explore the effect of these changes on conflicting cues.

1pPPb5. An oscillatory template pitch model. David A. Dahlbom and Jonas Braasch (Architecture, Rensselaer Polytechnic Inst., 110 8th St., Troy, NY 12180, dahlbd@rpi.edu)

Traditional approaches to explaining missing fundamental and pitch-shift phenomena have relied on either the fitting of harmonic templates to spectral information (generally assumed to correspond to a neural excitation pattern), or by performing an autocorrelation-based calculation on temporal information. In this paper, an alternative approach relying on the dynamics of phase-locking oscillators is proposed. Rather than applying some decoding procedure, such as spectral analysis or autocorrelation, a pitch estimate is given by the state (firing frequency) of an appropriately excited oscillator in a template structure operating on peripheral information. This approach is shown to reproduce many of the classical pitch shift-phenomena. In addition to a direct implementation in terms of phase-locking oscillators, a simplified model consisting of arrays of adaptive templates, operating directly on timing information, is also presented. This latter approach offers a straightforward way to tune the model in accordance with existing psychoacoustical data and suggests an approach to modeling pitch strength and multiple f0 detection. The overall goal is to suggest the advantages of a modeling approach that relies on arrays of tuned, adaptive elements which are physiologically-inspired, though not intended to be detailed representations of precise physiological components. [Work was supported by NSF BCS-1539276.]

1pPPb6. Pitch synchronous speech analysis for the assessment of subjects with Parkinson's disease. Sai Bharadwaj Appakaya and Ravi Sankar (iCONS Res. Lab, Dept. of Elec. Eng., Univ. of South Florida, 4202 E Fowler Ave. ENB 381, Tampa, FL 33620, saibharadwaj@mail.usf.edu)

Analysis of speech samples from subjects with Parkinson's Disease (PD) is a field of growing research interest. Studies in this field predominantly include pitch-based features with the intuition that PD affects the movement of musculature involved in speech production. These features are usually extracted on a segment of fixed length that traverses over the entire speech sample. This methodology, however, gives the net estimate of the feature over each segment and cannot account for the fast variations that transpire before and after a vocal fold closure. In this paper, we present a pilot study with the focus on in-depth pitch synchronous analysis that can fill-in the gap by capturing the said variations. Speech samples from 22 patients are preprocessed and used for analysis. Data that can pin point the precise indices of the vocal fold closures for each cycle are extracted in preprocessing and features that can capture the swift variations are extracted from every cycle and used for analysis. The results show a good margin in features between PD and healthy speech that reinforced the intuition.