hearing-aid noise-reduction algorithms. [Work supported by NSF BCS0236707 (OD,CYE) and NIH DC-001641 (MCA,LHC).]

4aSC5. A new method of extracting the filter characteristics of the nasal cavity using homorganic nasal-stop sequences. Hansang Park (Dept. of English Education, Hongik Univ., 72-1 Sangsu-dong, Mapo-gu, Seoul, Korea)

This study attempts to derive the filter characteristics of the nasal cavity of individual speakers. Since the only difference between a nasal and a homorganic voiced stop, such as [mb] and [nd], is whether the passage to the nasal cavity is open or not, the subtraction of the LPC spectrum of the voiced stop from that of the preceding nasal leads to the filter characteristics of the nasal cavity of an individual speaker regardless of place of articulation. The results showed that the spectral differences between samples of 20 ms taken from the steady states of the nasal and the following voiced stop were close to constant regardless of place of articulation, representing characteristic poles and zeroes, and that the spectral differences varied with speakers. This study is significant in that it provides a new method of extracting the filter characteristics of the nasal cavity, and that the spectral difference between a nasal and a homorganic voiced stop can be used as a parameter of the filter characteristics of the nasal cavity of individual speakers.

4aSC6. Synthesizing speech acoustics from head and face motion. Adriano V. Barbosa, Hani C. Yehia (CEFALA/PPGEE, Universidade Federal de Minas Gerais, Av. Antonio Carlos, 6627, Belo Horizonte, MG, 31270-010, Brazil, adriano.vilela@gmx.net), Andreas Daffertshofer (Vrija Universitat, Amsterdam, The Netherlands), and Eric Vatikiotis-Bateson (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z1)

This work outlines a quantitative analysis of the relation between speech acoustics and the face and head motions that occur simultaneously [A. V. Barbosa, Ph.D. thesis, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, 2004]. 2-D motion data is obtained by means of a video camera. An algorithm has been developed for tracking markers on the speaker's face from the acquired video sequence [A. V. Barbosa, E. Vatikiotis-Bateson, and A. Daffertshofer, in Proceedings of the 8th ICSLP Interspeech 2004, Korea, 2004]. The motion domain is represented by the 2-D marker trajectories, whereas line spectrum pairs (LSP) coefficients and the fundamental frequency $F0$ are used to represent the speech acoustics domain. Mathematical models are trained to estimate the acoustic parameters (LSPs $+$ $F0$) from the motion parameters (2-D marker positions). The estimated acoustic parameters are then used to synthesize the acoustic speech signal. Cross-domain analysis for undecomposed (i.e., full head $+$ face) and decomposed (i.e., separated head and face) normalized 2-D motions is performed. Syntheses from each method using intelligibility tests and qualitative comparison of the original and synthesized utterances are being evaluated.

4aSC7. ArtiSynth designing a modular 3D articulatory speech synthesizer. Florian Vogt, Oliver Guenther, Allan Hannam, Kees van den Doel (Univ. of British Columbia, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4, fvogt@ece.ubc.ca), John Lloyd, Leah Vilhan, Rahul Chander, Justin Lam, Charles Wilson, Kalev Tait, Donald Derrick, Ian Wilson, Carol Jaeger, Bryan Gick, Eric Vatikiotis-Bateson, and Sidney Fels (Univ. of British Columbia, Vancouver, BC, Canada V6T 1Z4)

ArtiSynth is a modular, component-based system for performing dynamic 3D simulations of the human vocal tract and face. It provides a test bed for research in areas such as speech synthesis, linguistics, medicine, and dentistry. ArtiSynths framework enables researchers to construct, refine, and exchange models of all parts of the vocal tract and surrounding structures. ArtiSynth introduces a probe concept to unify input and output data flow, which allows control of and access to models with time varying data series. ArtiSynth supports interconnected heterogeneous models, such as rigid body, mass-spring, and parametric, using a point-set connection method, called markers, for constraint satisfaction. Using ArtiSynth, we created a muscle-driven rigid body jaw model, a parametric principle component tongue model from MRI images, a parametric lip model, and mass-spring face tissue model. We combined them in various ways. Data from medical imaging (MRI, CT, and ultrasound) and other technologies such as optical tracking can be used to drive ArtiSynth models. We are currently developing an acoustical rendering framework supporting source-filter models and other advanced methods. The system incorporates a powerful scripting interface as well as an easy-to-use graphical interface. [Work supported by NSERC Canada and ATR Japan.]

4aSC8. Design of a 6 degree of freedom anthropomorphic robotic jaw. Edgar Flores and Sidney Fels (Dept. of Elec. & Comput. Eng., UBC, 2356 Main Mall, Vancouver, BC, Canada V6T 1Z4)

We have created a 6 DOF robotic jaw capable of producing, in real-time, the complex set of motions described by the human jaw during speech or mastication. The jaw is designed to fit within a larger robotic human figure such as the head, neck and torso of the 25 DOF Infanoid. [Kozima, Hideki: Infanoid: A Babybot that Explores the Social Environment, K. Dautenhahn *et al.* (eds.), Socially Intelligent Agents: Creating Relationships with Computers and Robots, Kluwer Academic Publishers, pp. 157–164, 2002]. The produced mechanical prototype has been designed to accommodate a prosthesis mandible with dentures. The mechanism could fit within the skull of the average man; where it would occupy less than 1/3 of the skull cavity. Two TMJs (temporomandibular joints) support the prosthesis, where each is driven by a 3 DOF parallel manipulator. In order to combine the motion of both manipulators each TMJ is capable of 3 DOF. The system is controlled via a USB port using software that models the human skull including collision detection mechanisms. The jaw allows for linear control, zero-backlash, and up to three times exaggerated mobility ranges making it also suitable for speech research, facial gesture affect research and dentistry applications.

4aSC9. Effects of subglottal acoustics on phonation onset. Juergen Neubauer, Zhaoyan Zhang, and David Berry (UCLA School of Medicine, 31-24 Rehabilitation Ctr., 1000 Veteran Ave., Los Angeles, CA 90095, zyzhang@ucla.edu)

The effect of subglottal acoustic loading on the vocal fold vibration was investigated using a self-oscillating mechanical model of the folds. Although the influence of the supraglottal tract on vocal fold vibration has received more attention than the subglottal system, the influence of the subglottal system on vocal fold vibration is also potentially significant, and merits investigation. In this study, the subglottal system consisted of a uniform tube connected to an expansion chamber on the flow supply end (e.g., a pseudo-lung). The length of the subglottal tube was varied systematically over a relatively large range in order to investigate the influence of subglottal acoustics on vocal fold vibration. Phonation onset and offset pressures were measured in the subglottal tube as a function of tube length. Over the range of investigation, the fundamental frequency of phonation was found to be negatively correlated with the subglottal tube length. However, both phonation onset and offset pressure were positively correlated with subglottal tube length, with the onset pressure increasing faster than the offset pressure. This hysteresis effect vanished and the two pressures merged at a small value of the subglottal tube length, indicating a change in the onset behavior from a subcritical Hopf bifurcation to a supercritical Hopf bifurcation (a codimension-2 bifurcation point). In addition, phonation did not exist below a critical subglottal tube length.