Day 1 – September 21: Introduction to the Summer School — Sound Perception in Rooms		
	13:30-14:30	Fundamentals of Human Hearing and Psychoacoustics *by Prof. William Yost, Arizona State University, Tempe, AZ*
		Psychophysics was developed by Gustav Fechner in the 19th century to describe the study of the relationship between sensations and perceptions and the physical parameters of objects in the world. Psychoacoustics is the psychophysical study of hearing and sound, i.e., the study of the relationship between a description of what one hears and the physical properties of sound and sound sources. Loudness, pitch, and the perceived location of sound sources are three of the more prominent and widely used attributes of hearing, and have been extensively studied as psychoacoustic relationships. Subjective measures of loudness and pitch depend on several different physical properties of sound and how it is transformed by the auditory periphery. The spatial location of a sound source can be entirely determined by its sound even though sound has no spatial dimensions. Neural analyzes made in lower brainstem circuits compute interactions between sound and the torso and head and in some cases, probably, reflective surfaces. These computations provide the information that allows for the determination of the location of sound sources. This lecture will review the basic psychoacoustic relationships between perceived loudness, pitch, and sound source location and the physical properties of sound and how those sound properties are transformed by the peripheral auditory system.

Day 2 – September 22: Peripheral Hearing and Auditory Modeling — Auditory Scene Analysis		
	15:00-16:00	Why rapid adaptive processes are essential for listening in realistic auditory environments *by Prof. Shihab Shamma, Univ of Maryland, College Park, MD*
		Humans can attend to one of multiple sounds, and follow it selectively over time. The neural underpinnings of this perceptual feat are the object of extensive investigations. I will review the fundamentals of sound representation in the auditory cortex. I will then explain how source segregation depends primarily on the rapid adaptation in auditory cortical responses that can track the temporally coherent responses induced by simultaneous sources. I will then review recent results in support of these ideas, and then demonstrate algorithms that can segregate sources with no prior information or training and that are inspired by auditory cortical mechanisms. *See attached files: TINS2011_Shamma.pdf and Segregation_Model.pdf at the end of this page.*

Day 4 – September 24: Robot audition – Quality of Experience – System Integration for the Challenge		
Morning	**Robot Audition - Invited Talk**	
	09:00-10:00	Speaker detection, localization and tracking with a microphone array on a mobile robot *by Dr. Ivan Marković, Univ of Zagreb, Croatia*
		Directional data, like bearing (azimuth), elevation or heading, is encountered often in many applications, including mobile robot audition where a speaker bearing is estimated with a microphone array. Working with directional data, especially under uncertainty, imposes a problem on how to represent them in probabilistic frameworks. Commonly this problem is solved by using a Gaussian distribution, which unfortunately does not capture well the non-euclidean properties of directional data. Furthermore, since small errors in the heading can result with great errors in the final location, the need to faithfully model the directional data should not be dismissed lightly. This talk will study the problem of speaker localization and tracking using directional statistics. Directional statistics studies mainly observations which are unit vectors either in the plane or in a three dimensional space. We will present our research results on speaker localization and tracking based on von Mises and von Mises-Fisher distributions and our on-going research based on Kalman filtering on Matrix Lie Groups. Furthermore, in robot audition, robust and reliable voice activity detection in varying acoustical conditions is of great importance for speaker localization and tracking. In this talk we will present our approach to voice activity detection, which use a set of informative distinct features, next to the likelihood ratio of a statistical model-based voice activity detector, in order to enhance the detection performance via a supervised learning approach. For all algorithms we will present simulation and/or experimental results. We will conclude the presentation with a discussion of potential directions for future research.