

ACOUSTICS2008/2482
**Spatial rendering of audio-visual synthetic speech use for
immersive environments**

Markus Noisternig, Brian Katz and Christophe D'Alessandro
LIMSI-CNRS, B.P. 133, 91403 Orsay, France

Synthetic speech is usually delivered as a mono audio signal. In this project, audio-visual speech synthesis is attributed to a virtual agent moving in a virtual 3-dimensional scene. More realistic acoustic rendering is achieved by taking into account the position of the agent in the scene, the acoustics of the room depicted in the scene, and the orientation of the virtual character's head relative. 3D phoneme dependant radiation patterns have been measured for two speakers and a singer. These data are integrated into a Text-To-Speech system using a phoneme to directivity pattern transcription module which also includes a phoneme to viseme model for the agent. In addition to the effects related to agent's head orientation for the direct sound, a room acoustics model allows for realistic rendering of the room effect as well as the apparent distance as depicted in the virtual scene. Real-time synthesis is implemented in a 3D audio rendering system.