Toward Mobile and Adaptive Conversational Interfaces:
Modeling Speech Convergence with Animated Personas

Sharon Oviatt

During the past decade, rapid advances in spoken language technology, natural language processing, dialogue modeling, multimodal interfaces, animated character design, and mobile applications all have stimulated interest in a new class of conversational interfaces. Such systems are being designed to support users' performance in a variety of task applications (commercial, medical, educational, in-vehicle), and many have been designed with animated characters that aim to facilitate user performance. However, the development of robust systems that process conversational speech is a challenging problem, largely because users' spoken language can be extremely variable. In this talk, I'll describe research in our lab that has identified a new source of variability in users' spoken language to computers. Basically, people spontaneously and rapidly adapt the basic acoustic-prosodic features of their speech signal to the text-to-speech output they hear from a computer partner. These speech adaptations are delivered dynamically, since users will quickly readapt their speech when communicating with a different computer voice. They also are flexibly bi-directional—for example, users will increase their own speech amplitude and rate when conversing with a computer partner that has louder and faster text-to-speech (TTS) output, and will decrease these features when the TTS is quieter and slower. In fact, an analysis of speakers' amplitude, durational features, and dialogue response latencies confirmed that these adaptations can be substantial in magnitude (10-50%), with the largest adaptations involving utterance pause structure and amplitude. This research underscores the need for new speech and multimodal systems that can adapt to users and their communication context. It also emphasizes the importance of auditory interface design for next-generation mobile systems. Implications are discussed for designing future conversational interfaces that are more reliable, well synchronized, and supportive of user performance.

Brief Biography:
Sharon Oviatt is a Distinguished Scientist at Adapx and President of Incaa Designs. Her research focuses on human-centered design and cognitive modeling, communication technologies, spoken language, pen-based and multimodal interfaces, and mobile and educational interfaces. Examples of recent work involve the development of novel design concepts for educational interfaces (science and math), robust interfaces for real-world mobile and field environments, adaptive conversational interfaces with animated software characters, digital paper and pen-based interfaces for individuals and groups, and lifespan interfaces for children through the elderly. She has published over 120 scientific articles in a wide range of venues, including work featured in recent special issues of *Communications of the ACM*, *Human Computer Interaction*, *Transactions on Human Computer Interaction*, *IEEE Multimedia*, *Proceedings of IEEE*, and *IEEE Transactions on Neural Networks*. She was General Chair of the *International Conference on Multimodal Interfaces (ICMI)* in *2003, and is Founding Chair* of *ICMI's* Advisory Board. In 2000, she received an NSF Special Extension for Creativity Award for pioneering work on mobile multimodal interfaces.