Human Factors and Speech Technology

Reflections on 20 Years of User Interface Design for Speech-Enabled Applications
Daryle Gardner-Bonneau, Ph.D.
Principal, Bonneau and Associates
and
Adjunct Associate Professor
Speech Pathology and Audiology
Western Michigan University

JDNBonneau@cs.com
Speech Technology in 1983

- Few commercial applications
- Lots of research and demonstrations of the technology’s potential
- Discrete speech recognition, mostly small-vocabulary speaker-dependent systems, with onerous training protocols for users
- Synthetic speech that was reasonably intelligible, but sounded very robotic
The Applied Voice/Input Output Society (AVIOS)

- Established in the United States in the early 1980s
  - to educate potential end users and developers about the potential of speech technology for application in the real world
  - more a Board of Directors that planned an annual conference than an actual society
  - established what is now the *International Journal of Speech Technology*
  - still has no counterpart anywhere else in the world
The First U.S. “Killer” Commercial Application – IVR Systems

- Intelligibility of synthetic speech was good enough for commercialization
- Telecommunications industry and businesses saw great potential for customer “self-service” with interactive voice response systems that used touch-tone input and synthetic speech output
- Applications became ubiquitous in the late 1980s
IVR System Challenges

- Many poorly designed applications
- Attempts to “transfer” visually based design concepts to over-the-phone transactions were awkward, at best
- Attempts to standardize on good designs in the early 1990s came too late
- Most of the rest of the world “skipped” this evolutionary phase because of limited touch-tone penetration and other reasons
Frustration in the early 1990s

- Technology didn’t deliver as promised
- Large vocabulary speech recognition not quite there yet
- Speaker independence not yet achieved, except for very small vocabularies
- Synthetic speech highly intelligible now, but still robotic
Major Breakthroughs in the Mid-1990s

- Large vocabulary speaker-dependent and adaptive systems achieved, with much less onerous training requirements
  - first “shrink wrapped” dictation applications (e.g., Dragon Naturally Speaking)
  - still problems with user interface issues, including error correction, which remains a problem
Linguistic research allowed for the development of better synthesizers that sounded like humans

Speaker independent systems improved and allowed for larger vocabularies

Developers became encouraged again

IVR system developers, in particular, now saw the potential for speech enabled IVR applications and big business in customer “self service.”
The “Curse” of Customer Self-Service

- IVR systems – whether speech-enabled or not – are a huge frustration to many users.
- No longer possible to reach a human being over-the-phone for many companies in the U.S.
- The technology works, but a lack of understanding of human-to-human dialogue or conversation has led to very poor applications which don’t meet user goals.
The Curse of Customer “Self-Service” (continued)

- Applications designed by young “power” users are utterly unusable by older people and those with disabilities
  - speech too fast
  - responses required too fast
  - too many options; too much to remember
  - application logic that makes no sense to users
  - no way to get help
A Man Named Paul English Gets Mad

- In late 2005, a man named Paul English became frustrated with being unable to reach a human operator from inside an IVR application
- He “experimented” with the applications of a large number of Fortune 500 companies until he found the “magic” touchtone sequence for reaching a human operator. Then he published the entire list on he WWW!
Renewed Attention to Speech Application User Interfaces

- Paul English got noticed, and guidelines for user interface design for speech-enabled IVR applications started getting attention.
- ANSI/HFES ballots HFES 200.4 – a draft standard on IVR user interface design in 2006; currently being balloted as final draft U.S. national standard.
- HFES 200.4 will serve as input for a new work item proposal to be balloted by ISO TC159 SC4 WG5 – Software user interfaces.
- Will business pay attention? It remains to be seen.
IVR Applications and People with Disabilities

- IVR applications are as problematic for deaf people as Windows applications were, at first, for blind people.
- It has been a challenge to design IVR applications to work with U.S. text-telephone devices (TTYs or TTDs).
- ANSI/HFES 200.4 contains accessibility-related guidance for IVR application design.
Improving IVR Accessibility and Usability – Guidance

- Provide context-sensitive help
- Use consistent language and terminology and avoid jargon
- Use prosodic devices in prompts, announcements, and option lists to help users process information more easily
- Minimize keypad input
- Allow users to replay prompts and announcements
- Don’t place unnecessary time limits on user responses
Designing for the Needs of Older Adults

- Adults over the age of 65 represent the fastest growing population segment in many industrialized countries.
- Significant attention is being given to needs of older adults with respect to technology user interface design.
- Older adults are the most highly variable population segment in terms of most measures of human performance.
IVR Applications and Older Adults

- General “cognitive slowing” that occurs with aging necessitates applications that allow more time for processing information and responding to prompts.
- Memory requirements should be minimized in IVR applications, as short-term memory declines, generally, with age.
- Applications need to be tested with older users!
Standards Development Efforts

- There are many standards development efforts related to designing for accessibility.
- There are so many efforts, in fact, that ISO/IEC JTC1 established a Special Working Group on Accessibility (SWG A) to catalogue all of them.
- ANSI/HFES 200.4 is the only current standard related specifically to the accessibility of speech systems, but this will likely change with proposed ISO work in TC159 SC4 WG5.
Some other general accessibility standards may be helpful in designing speech enabled applications that serve the needs of older users and people with disabilities.

These include:

- ISO 9241 171 – Software accessibility
- ISO TR 22411
- ISO Guide 71
- Outputs from ISO/IEC JTC1 SWG- A (especially lists of design needs for people with disabilities)
Thank You!!

Thank you for your kind attention and the opportunity to participate in this conference!