

# B201

Speech search vs touch sensitive screens: who will win the mobile user interface war?

- B201 Multimodal user interfaces
- Mobile phone have multiple modes of entering and displaying data, all of which can support and be support speech technology. Mixing modes of interaction can effective, but runs the risk of being confusing, and the gives some exmples of the right way to mix speech with other modes.
- Speech search vs touch sensitive screens: who will win the mobile user interface war?
- The wide acceptance of the Apple Ipod and similar mobile devices demonstrate that speech is not required for useful interactive dialogs on mobile devices. Mobile device users have become very proficient at texting and searching on mobile devices using touch input. What are the strength and weaknesses of touch and speech interfaces? What role will speech play in interfaces for mobile devices in the future?

# Speech Search VS Touch Sensitive Screens: Who Will Win the Mobile User Interface War?

Jim Larson

Larson Technical Services

[www.larson-tech.com](http://www.larson-tech.com)

# VUI

- VUI is derived from telephone usage
- Initially customers spoke directly with a telephone operators to place calls
- Touch tone-based IVR enabled customers to be their own telephone operators
- ASR greatly extended functionality of IVR systems
- VoiceXML 2.x promoted system-directed dialog style

# GUI

- GUI is derived from PC usage
- Screen displays information and choices, promoting a user-directed dialog style
- Screen shrinks to fit on small hand-held devices
- Users become skilled “thumbers” as keyboard shrinks to fit on small hand-held devices

# Today GUI Dominates Mobile Devices

- The answering machine GUI
  - Visual lists of messages
  - Select and listen/text/connect
- Most other apps use GUI
- Limited VUI
  - Dial by name
  - Voice commands

# The evolution to multimodal

1. Touchtone input and voice output
2. Speech input and output
3. Speech input and output and visual output
  - Media player
  - Visual menus with select by key or speaking
4. Multimodal modes of input and multimedia output

# Multimodal: Swiss Army Knife of User Interfaces

## **Input**

- Touchtone
- Audio and video capture
- Speech recognition
- Keypad
- Handwriting
- GPS

## **Output**

- Display
- Audio and video replay
- Speech synthesis

# What must we do?

- Multimodal interface design guidelines
  - Don't use speech just because you can
  - Always have a backup to speech
  - Use consistent commands and dialog styles



# What must we do?

- Standard set of APIs across mobile operating systems
  - Audio/Video record and replay
  - Speech Recognition and Synthesis
  - Speaker Identification and Verification

# What must we do?

- Deploy successful multimodal applications that “prime the pump” of user popularity
  - VCR controller
    - Control replay of audio and video files
  - Guides
    - Repair and debugging
  - Command and control
    - Environment, juke box, ...
  - Virtual butlers
    - TV Butler
    - Environmental Butler
    - Family-activity coordinator
  - Synthetic beings: Converse with artificial agents representing real or imaginary people

# Who Will Win the Mobile User Interface War?

- ~~VUI~~
- ~~GUI~~
- Multimodal UI for Electronic Companions