User Interface Design Challenges in Multimodal Interaction

Catherine Zhu
Avaya Professional Services
Self Service Solutions
Multimodal Interaction On Mobile Devices

Voice User Interface (VUI)

Multimodal User Interface (MUI)

Graphical User Interface (GUI)
Multimodal Application Types

Differences in availability of modalities within one interaction:

- VUI-only interaction with GUI providing pictures to supplement (no input to the GUI) or vice versa
- VUI-only interaction followed by a GUI-only interaction
- Interaction where modalities alternate (must use only one modality at some points during the interaction)
- VUI and GUI interaction where both modalities are available at all times
Why would you use it?

- **Flexibility of Interaction**
  - Choose the modality that’s most applicable in the user’s current situation

- **User Control**
  - Allows users to choose how they interact with a system

- **Efficiency**
  - Users are able to use the modality that is the most efficient for the information presented
    - e.g. A visual list versus an auditory list
    - e.g. Saying a credit card number versus typing a credit card number

- **Fun**
  - New technology, New way to interact with a system
Voice User Interface Design Considerations

Which Tasks to Automate
- Repeatable, mundane tasks with large calling populations

Order of Task
- Consistent with user’s mental model
- Consistent with business rules

Which Information Does the Caller Need to Provide
- Login procedures and transactional processes
- Information format
  - Numeric or Alphanumeric Information - Patterned or Not
  - Number of Menu Options

What Stored Information Can the System Use
- Available for dynamic (even personalized) experiences
Voice User Interface Design Considerations

- **Overall Tone and Feel**
  - Voice talent, word choices, and earcons all help to create a specific persona
  - Matches the brand image of the business

- **Prompting**
  - Clearly defining the acceptable responses and constraining open-ended speech when necessary
  - Terminology and verbiage should be appropriate for the user population
  - Terminology and verbiage should be consistent

- **Error Handling Strategy**
  - Typically 3 errors (timeouts or retries) and transfer out

- **Confirmation Strategy**

- **Hints and Landmarks for Universal Commands**
Graphical User Interface Design Considerations

- **Which Tasks to Automate**
  - Repeatable, mundane tasks with large user populations

- **Order of Task**
  - Consistent with user’s mental model
  - Consistent with business rules

- **Which Information Does the Caller Need to Provide**
  - Login procedures and transactional processes
  - Information format
    - Numeric or Alphanumeric Information - Patterned or Not
    - Presentation Format of Options (e.g. drop-down box or free-form text box)

- **What Stored Information Can the System Use**
  - Stored data available for dynamic (even personalized) experiences
Graphical User Interface Design Considerations

- **Overall Look and Feel**
  - Page Layout, Colors used
  - Matches the brand image of the business

- **Text Format and Verbiage Used**
  - Clearly defining the acceptable responses and constraining open-ended input when necessary
  - Terminology and verbiage should be appropriate for the user population
  - Terminology and verbiage should be consistent

- **Error Handling Strategy**
  - Simple and relevant
  - Allow easy reversal of actions
  - Provide relevant visual feedback (e.g. progress bars)

- **Shortcuts and Landmarks**
Multimodal User Interface (MUI)

Voice User Interface (VUI)

Graphical User Interface (GUI)
MUI-Specific Design Challenges

- prompting (in VUI) and text (in GUI)
- navigation
- error handling
- usage environment
- overall modality preference
- modality syncing
- inconsistent user mental models

Multimodal User Interface (MUI)
Overall Modality Preference

- Based on prior experiences and/or general tendencies, users may prefer one modality over the other
  - Some users may have a bias to use verbal or touchtone input
  - Some users may have a bias to use text or touch input
- Some users will interact with the system choosing between modalities equally
- Users may change their mode of interaction depending on the mobile device being used
- Users may change their mode of interaction depending on the task they are trying to accomplish
- The most usable MUIs will take user preferences into account and adapt the interaction to them.
Usage Environment

- Users will change how they interact with the system depending on their surroundings
  - Users in quiet situations will typically prefer speech
  - Users in noisy environments (e.g. the subway) will typically prefer textual or touch input

Because these devices allow users to be mobile, the environment can change during a single interaction and will affect how users choose to interact with the system.

Privacy, Discreetness, and Hands-Busy Situations all affect how users want to interact with the system.

The most usable interfaces will allow users to change their mode of interaction at will as their environment changes.
Prompting and Text

The prompting heard in the VUI and the text shown on the GUI should be consistent.

Prompting and Text can be modified to be more modality neutral

For example:
– Typical VUI prompt: When you hear the one you want, just go ahead and say it.
– MUI prompt: When you find the one you want, just go ahead and say it or select it on your screen.

Prompting can point the user to look back at the screen and vice versa

Text on the GUI might also need to be numbered to match the touchtone options in the VUI
Navigation

- MUI applications typically work best with a guided structure that allows the user to proceed through a series of steps to complete a task.

- Guided structure will typically use directed dialogue in the VUI where users have a limited set of acceptable responses.

- The GUI will also need to be constrained and display the acceptable choices (rather than having free-form text fields).

- Differences between the logical order of steps in a VUI and the logical order of steps in a GUI will need to be considered.

- What navigational commands are active will also need to be considered.
  - Go Back? Main Menu? Start Over?
Error Handling

- Speech Recognition Errors
  - Occur when the system doesn’t recognize something the user said
  - Does 3 Errors and Transfer still apply?
  - Should you even transfer out to a representative?
  - Upon escalating errors, should the system point the user back to the GUI? Still provide touchtone fallback?

- Typical GUI Errors
  - Occur when the user inputs something in the wrong format or when the user fails to input something required
  - How does the speech system handle typical GUI errors?
  - Should the system point the user back to the VUI if multiple GUI errors have occurred?

- Because users have the choice of modality, are users more tolerant of errors?
Modality Syncing

What happens when one modality completely fails?
- Is the other modality aware of the failure?
- Does the failure change the experience in the continuing modality?

What happens when one modality is out-of-sync with the other?
- Long latency times to reload a page
- User barges-in in the middle of a prompt and says their response
- User barges-in in the middle of a prompt and touches or texts their response
- Is the other modality aware of the synchronization problem?

What happens if the other modality recovers in the middle of the interaction after a failure or synchronization problem?
Inconsistent Mental Models

- Users come to VUI and GUI interactions with specific mental models of how to interact with each type of system.

- In VUI interactions, users are typically familiar with the terms “Main Menu” and “Representative” even when not prompted for them.

- In GUI interactions, users are typically familiar with the terms “Home Page” and “Back”.

- When the mental models conflict, the user may be unsure how to interact with the system:
  - Will the user ever say “Home Page” to go back to the Main Menu?
  - Will the user want a “Representative” button on every GUI page?

- A systematic framework for interacting with MUI applications will need to be created.
Conclusions

In designing user-friendly MUls, basic VUI and GUI design principles must be considered.

However, in creating a system that allows users the choice of multiple modalities, new MUI-specific challenges must also be considered including:
- Overall Modality Preference
- Usage Environment
- Prompting in VUI/ Text in GUI
- Navigation
- Error Handling
- Modality Syncing
- Inconsistent User Mental Models

By accounting for these MUI-specific challenges, multimodal applications for mobile devices will become more useful and more usable.