GUI Meets VUI:
Some Possible Guidelines

James A. Larson
VP, Larson Technical Services
jim42@larson-tech.com
Situations Where Display Technology Is Useful

\[
\int e^{4x} \cos 5x \, dx = \text{Re} \left[ \int e^{4x}(\cos 5x + i \sin 5x) \, dx \right] = \text{Re} \left[ \int e^{(4+5i)x} \, dx \right]
\]

\[
= \text{Re} \left[ \frac{1}{4+5i} e^{(4+5i)x} + c \right] = \text{Re} \left[ \frac{1}{4+5i} e^{4x}(\cos 5x + i \sin 5x) + c \right] = \text{Re} \left[ \frac{1}{4-5i}(4 \cos 5x + 5 \sin 5x + 4 \sin 5x - 5 \cos 5x) e^{4x} + c \right]
\]

\[
= \frac{4 \cos 5x + 5 \sin 5x}{41} e^{4x} + c
\]
Situations Where Speech Technology Is Useful
Combine GUI and VUI

• Leverage advantages of GUI and VUI
  – GUI
    • Extend user memory
    • Look and manipulate
    • Personalize display arrangement
  – Speech
    • Annotate, narrate, and dictate
    • Present errors and notification
    • Formulate complex queries and requests

• Enable user to switch between GUI and VUI easily and naturally
11 Guidelines for Multimodal User Interfaces

1. Enable Users to Choose Mode
2. Make Presentation Reactive
3. Minimize Cognitive Processing
4. Use Voice Assistance
5. Use Multifactor Speaker Identification and Verification
6. Use Voice in Emergencies
7. Use Voice When Hands, Eyes Busy
8. Leverage of User Emotion
9. Integrate Voice and Gestures
10. Inform User When to Speak
11. Involve Users at Every Stage
1a. Enable User to Choose Input Mode

- Click or speak
- Swipe, touch, or speak
- Type or touch
1b. Enable User to Choose Output Mode

• For privacy,
  – Turn off voice output
  – Route graphical output to a personal display
  – Present result “availability,” but save content for later
2. Make Presentation “Reactive”

Now is the time for all good men to come to the aid of their country
3. Minimize Cognitive Processing
Select Rather than Recall

Possible actions

Objects

Task status
4. Use Voice Assistance

Voice Control of Video

Voice-Over Narration for Help Animation

To add an app to your home screen, touch, and hold “Download”

Stop, replay previous 10 seconds
5. Use Multifactor Speaker Identification and Verification
6. Use Voice in Emergencies
7. Use Voice When Hands or Eyes Busy
8. Leverage User Emotion

• Detect user emotion
  – Emotion detection from word choice
  – Emotion detection from voice properties

• Convey emotion to user
  – Use expressive voices
  – Emotion effects program logic
9. Integrate Voice and Gestures

Put that there.

10. Inform User When to Speak
11. Involve Users at Every Stage

User test!
User test!
User test!
Summary

• Combining GUI and VUI presents challenges
  – Coordinate voice and gesture input
  – Determine whether to speak or display

• Combining GUI and VUI has benefits
  – Leverage advantages of both GUI and VUI
  – User convenience

• Plan and develop carefully
  – User test, test, and test some more
Questions