

Welcome

To Advance through Presentation
Use Page Up and Page Down Keys



99 | Worldwide
Developers
Conference



From “Point-and-Click” to “Ask-and-Tell”

Kim Silverman, PhD

Manager,
Spoken Language Group

Speech Is a First-Class Member of the Operating System

- On all CDs since 8.5
- APIs solid and supported
- APIs are committed for Carbon



Recent Releases

- 8.5: Back into the OS
- 8.6:
 - Work properly with Sherlock
 - Runs on all G3s (44.1 kHz)
 - iMac built-in microphone



Sonata

- New Developer features
 - More accuracy, better user experience
- Speech Recognition in AppleScript
- Opportunity for developers to make faces



Strategy

- Remember-and-Type
- Point-and-Click
- Ask-and-Tell
- Current focus for speech recognition:
 - Command and Control





99

Worldwide
Developers
Conference

Robustness to Noise

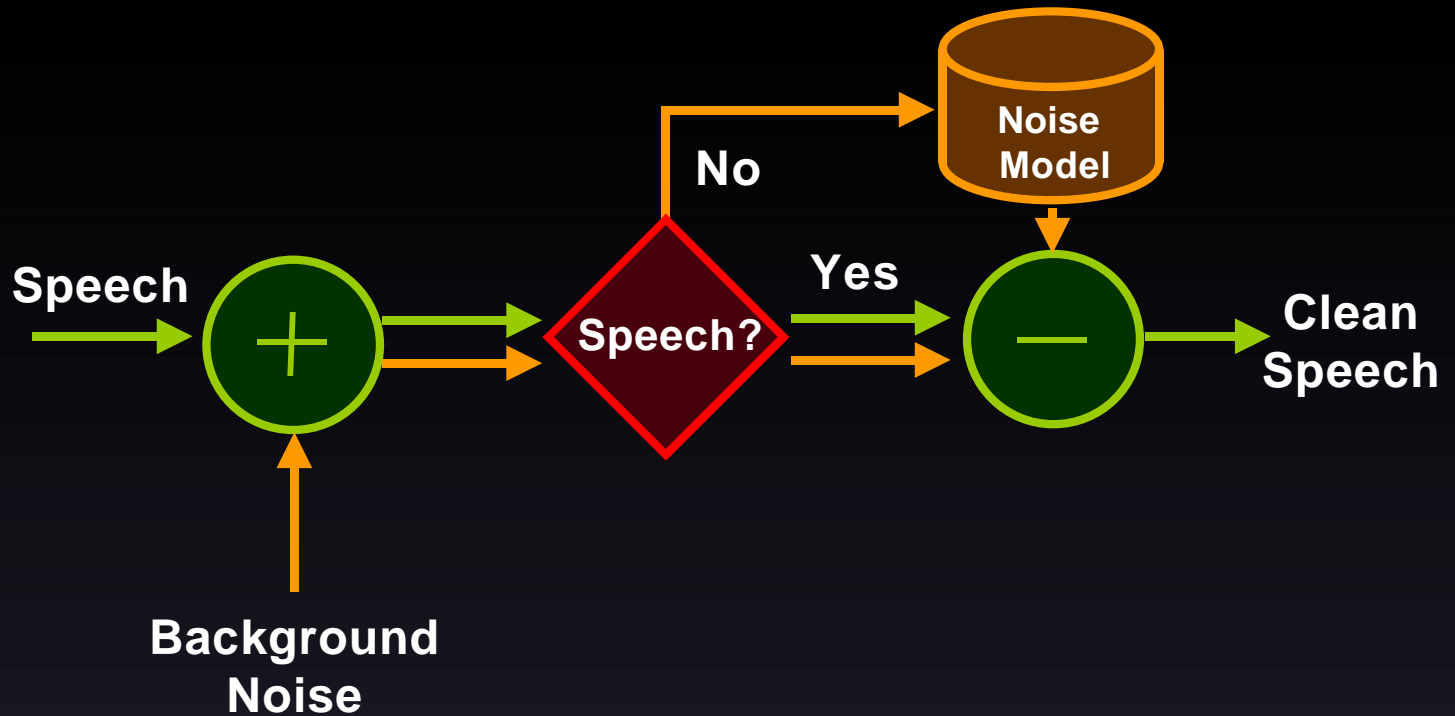
Devang Naik, PhD
Senior Scientist

Factors

- Background noise
- Reverberation
- Fan noise
- User movement



Block Diagram





What's New in English Speech Recognition

Thomas Bonura, PhD
Senior Scientist

PlainTalk Speech Recognition

- New features of SpeakableItems
- The “Speech Listener”—brings SR to AppleScript
- Making effective use of speech recognition



New Features of SpeakableItems

- Application listening contexts
- Built-in commands



Application Listening Contexts

- Automatic language model activation: global and frontmost listening contexts
- Allows for overloading speakable actions
- Manage larger language models



Developer Impact

- 3 steps to plugging into SpeakableItems



Developer Impact

- 3 steps to plugging into SpeakableItems
 - Create a folder in the “Application Speakable Items” Hierarchy



Developer Impact

- 3 steps to plugging into SpeakableItems
 - Create a folder in the “Application Speakable Items” Hierarchy
 - Create an invisible alias to your application in that folder

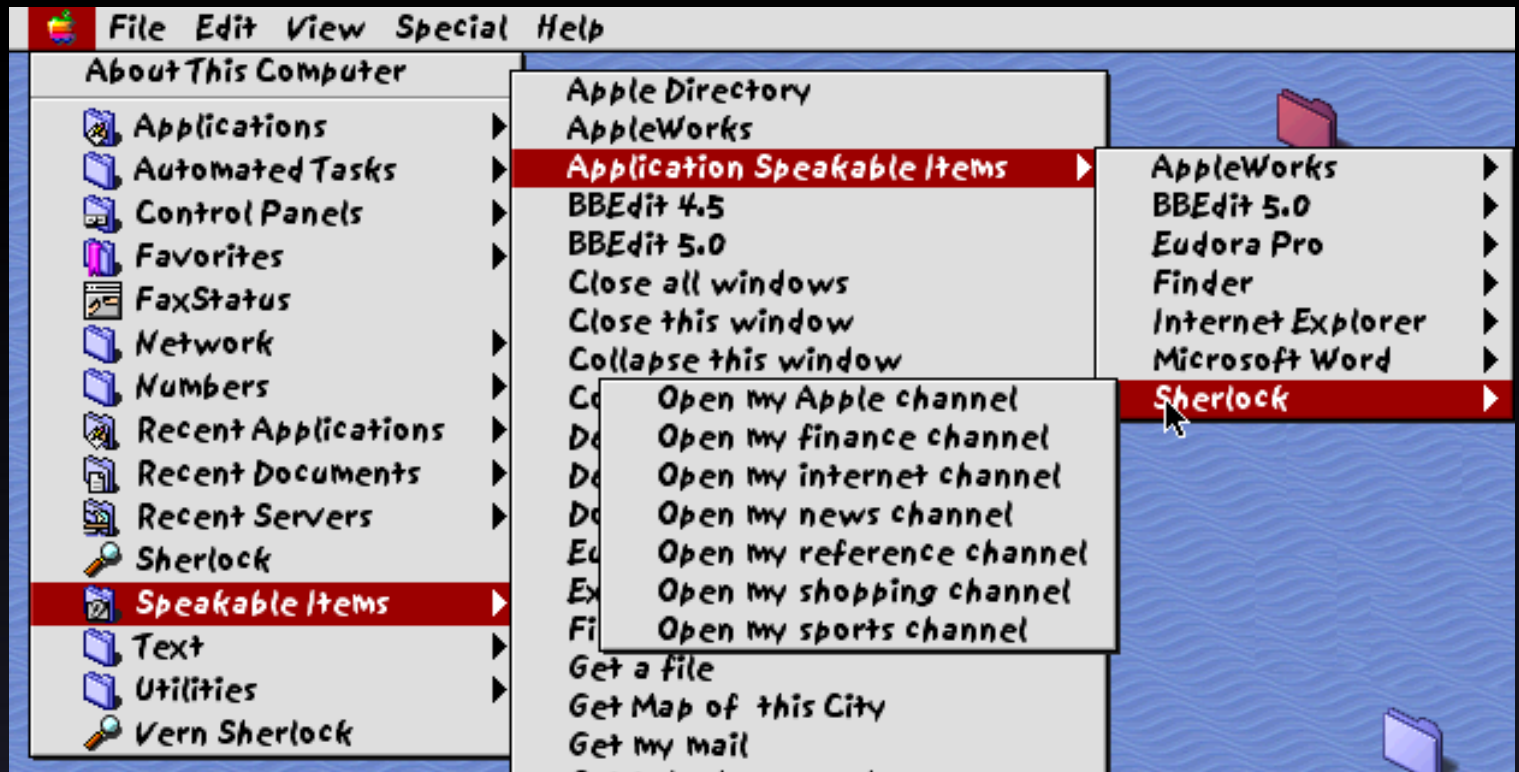


Developer Impact

- 3 steps to plugging into SpeakableItems
 - Create a folder in the “Application Speakable Items” Hierarchy
 - Create an invisible alias to your application in that folder
 - Add your speakable items



For Example



Built-in Commands



Built-in Commands

- “Make this application speakable”
 - Adds an alias to the top level of Speakable Items hierarchy
 - Creates an application specific folder
 - Creates an invisible alias to the application





99 | Worldwide
Developers
Conference

Demo

Speech Recognition in AppleScript

- Access via the “Speech Listener” application
- Easy access to SR from any script
- Speech Listener implements a powerful and useful subset of PlainTalk SR APIs



For Example:

```
try
  tell application "Speech Listener"
    listen for {"Strawberry", "Lime", "Blueberry",
    "Tangerine", "Grape"} with prompt "What is your favorite iMac
    color?"
  end tell
on error
  say "Sorry, did not recognize a color, try again."
end try
```

If the result is "Lime" then ...





99 | Worldwide
Developers
Conference

Demo

Speech Recognition in Your Application

- How do you take advantage of SR?
- Creating a great user experience



Taking Advantage of SR in Your Application

- Do nothing!
- Use AppleScript, Speech Listener and SpeakableItems
- Call the toolbox APIs directly by your application



Say What You Can't See!



Say What You Can't See!

But point to what you can see!





99 | Worldwide
Developers
Conference

Demo

Say What You Can't See

- Take advantage of direct manipulation to refer to objects; use speech to apply commands
- Allows the user to stay focused on content not commands



Spoken Dialog



Spoken Dialog

- Use dialog to assist the user in providing the right information at the correct time



Spoken Dialog HI Guidelines

- Always provide help
- Make the dialogs simple with few choices
- Allow for graceful exits



Consider

A language model for setting up a meeting, e.g.,
“Meet with Kim on Tuesday”



Language Model

“Meet with Kim on Tuesday”

$\langle \text{top} \rangle = \text{meet with } \langle \text{person} \rangle \text{ on } \langle \text{day} \rangle;$

$\langle \text{person} \rangle = \text{tom} \mid \text{kim} \mid \text{kevin};$

$\langle \text{day} \rangle = \text{monday} \mid \text{tuesday};$



But What Happens Here?

User says: “Meet with **Mary** on Tuesday”

Computer: “Sorry, did not understand that”

WRONG!



A Better Interaction

User says: ““Meet with Mary on Tuesday””

Computer: “Sorry, did not understand the name. You can say ...”



Solution

$\langle \text{top} \rangle = \text{meet with } \langle \text{person} \rangle \text{ on } \langle \text{day} \rangle;$

$\langle \text{person} \rangle \{ \text{kSRRejectable} = \text{true} \}$
 $= \text{tom} \mid \text{kim} \mid \text{kevin};$

$\langle \text{day} \rangle \{ \text{kSRRejectable} = \text{true} \} = \text{monday} \mid \text{tuesday};$





99 | Worldwide
Developers
Conference

Demo

In Summary



In Summary

- Speech is a different modality than DM—
“Speak what you can’t see”



In Summary

- Speech is a different modality than DM—
“Speak what you can’t see”
- Speech and direct manipulation can
compliment one another



In Summary

- Speech is a different modality than DM—
“Speak what you can’t see”
- Speech and direct manipulation can
compliment one another
- Use dialog in order to guide and provide
assistance to the user





Speech Recognition Avatars

Matthias Neeracher, PhD
Senior Software Engineer

What Are Avatars?

- Also known as Feedback Characters
- Animations that pop up in a floating window when speech recognition clients are active
- Provide visual feedback during listening and processing phases
- May include visualization of spoken text
- Were always there, but not documented previously, will be documented in forthcoming Technical Note



Why Create Avatars?

- To support the visual identity of your speech recognition client by giving it a distinct appearance
- To deliver separately for commercial or promotional purposes
- Because it's fun!



Avatar Anatomy

- Avatars are QuickTime components, can be included in your client or put into stand-alone ‘thng’ files which are autorouted to extensions folder
- Consist of
 - Component and Name resources
 - Pictures for expressions
 - A Face resource, an Expression resource, a Frame resource and a Phoneme resource



Component Resources

- Avatar components have type ‘fdbk’ and an unique subtype (use a creator code registered to you)
- Name resources provide a unique name and a description



Face Resource

- Declares the size of the feedback window, the fonts to be used and points to the expression, frame, and phoneme resources



Expression Resource

- Declares a list of states, each consisting of
 - A flag value
 - A flag mask
 - The number of the starting frame for the associated animation



Frame Resource

- Declares a list of animation frames, each consisting of:
 - The picture ID to display
 - The duration of the frame
 - The offset to the next frame

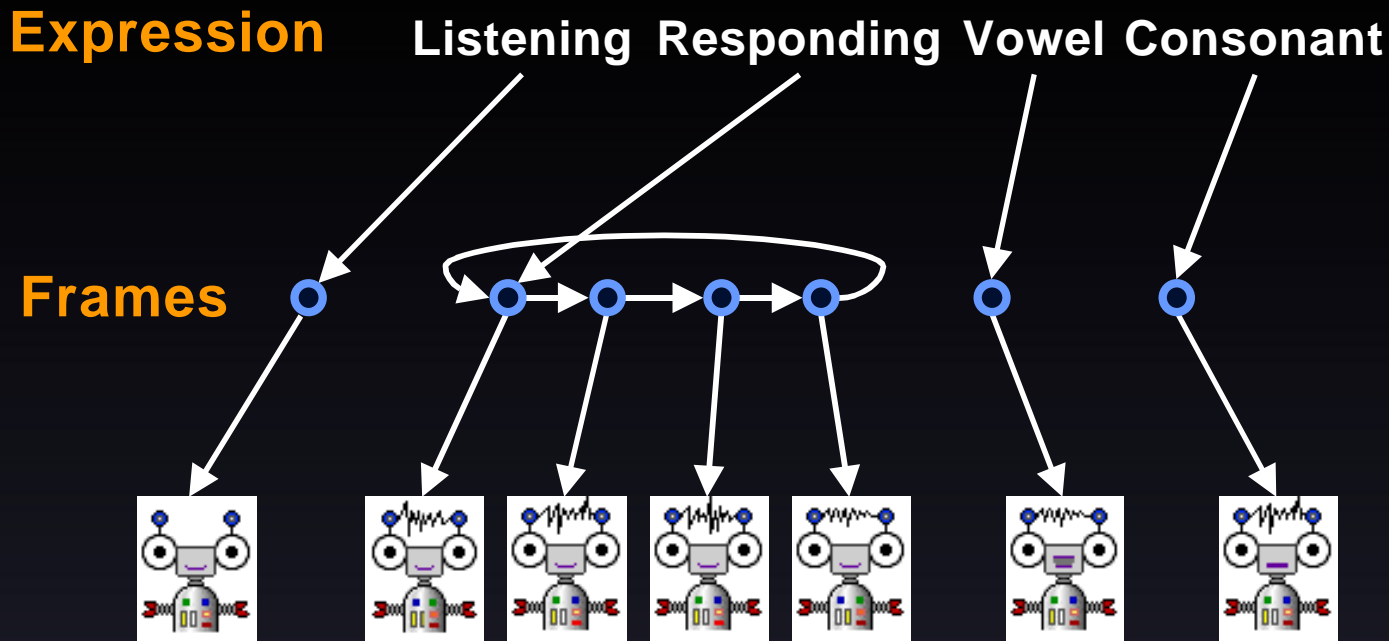


Expression Pictures

- Need to be provided in both B&W and 8 bit color
- Color versions should use standard Apple 8 bit palette; can save further space by removing CLUT from pictures using CLUTless
- All images should be same size



Resulting Animations



Further References

- “A Space-Saving PICT Trick,” develop issue 20
- Avatar Technical Note, to appear





Think different.TM



Welcome

To Advance through Presentation
Use Page Up and Page Down Keys



99 | Worldwide
Developers
Conference