Vlingo Mobile Search: Making It Simple

Mike Phillips
CTO & Co-founder
Mobile Interfaces and Web Services

- Open speech interfaces now possible
- Broad range of Applications and Services
  - Information: search, weather, sports
  - Entertainment: music, video
  - Communication: email, SMS, IM, social networks
  - Productivity: calendars, voice dialing
- Can we combine these in a single user experience?
Usability of Speech Applications

Diagram showing a relationship between Usability and Flexibility. The diagram includes a point labeled 'Humans' which appears to be outside the typical usability range, indicated by the question mark '???'.
Usability of Information Retrieval Applications

Flexibility

Humans

???
Successful Information Retrieval Application
How Does This Apply To Voice Dialog Systems?

Simple Model for Users

No Boundaries
Simple Model For Users

- Push “speak” button, say whatever you want
- System tries to do right thing:
  - If Idle, take user to appropriate application/action
  - If in a text field, fill that text field
- Let user know what is recognized and what is being done
- Give them an easy way to change it
- Give them an easy way to redo it
Technology For Unconstrained Speech Input

- New Technology
- Automatic Adaptation
- No up-front user training

- Acoustic Models
  (what speech sounds like)

- Pronunciations
  (what words sound like)

- Vocabulary
  (what words people say)

- Language Model
  (how words go together)

- Speech Recognition Engine

- Data Network

- Usage Data

- Mobile Device
How to Handle Action

• Determine appropriate action
  • General: dialing, search, messaging, phone control
  • Specific: navigation, find a business, book a hotel

• Route input to application
  • Applications are domain experts
  • Local or Web-based applications

• Need to maintain consistent user experience
Technology For Application Routing

New Technology
Automatic Adaptation
No up-front user training

Intent Models
- Acoustic Models (what speech sounds like)
- Pronunciations (what words sound like)
- Vocabulary (what words people say)
- Language Model (how words go together)

Intent Engine

Speech Recognition Engine

Data Network

Usage Data

Say business or category, followed by city and state
or "Order in San Francisco"
Technology for Actions

Local application
Local application

Web Service
Web Service

...
Market Adoption

• Technology is now capable of handling broad range of mobile services

• Market adoption dependent on
  – High performance, reliable, low cost mobile data networks
  – Capable open devices
  – Successful implementations

• Need consistent high quality user experience