

Building Robust Speech Applications Cost- Effectively

Ye-Yi Wang

Microsoft Research, Redmond WA, USA

December 10, 2007

Bring Speech into the Mainstream

- Pieraccini's chasm
- *Better, Faster, Cheaper*
- **Better** – more robust, more accurate
- **Faster** – rapid development and tuning
- **Cheaper** – less dependency on expertise and data

Spoken Language Understanding

Form Filling (ATIS)

Show flights from Seattle to Boston on December 23

ShowFlights

From=Seattle

To=Boston

Date=12/23

Call Routing (AT&T HMIHY)

My outlook does not show any new messages for about 5 hours

Networking

OS

Email

Hardware

Voice Search (Directory Assistance)

I need the number of Big 5

Sears Roebuck & Company

• 2200 148th Ave
Redmond, WA
• (425) 644-6526

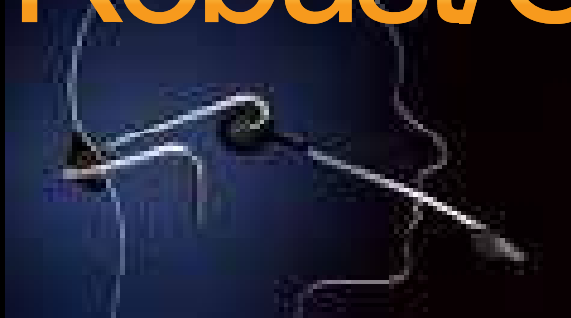
Calabria Ristorante Italiano

• 132 Lake St S,
Kirkland, WA
• (425) 822-7350

Big 5 Sporting Goods

• 4315 University
Way Seattle WA
• (206) 547-2445

Robust/Cost-effective Technology



	LM	SLU
Form filling/ directed dialog	Grammar Controls Creating grammar from high-level description.	
Form filling/ mixed initiative	Unified LM LM adaptation Combination of data-driven and engineered knowledge reduced data requirement & improved robustness	HMM-CFG model, SGStudio, CRFs
Call routing	Self-adapting LM Making errors more consistent	Model training w/o transcription
Voice search	Variation LM/ Efficient LM switching	Enhanced VSM More robust to linguistic variants & ASR errors.

Robust/Cost-effective Technology

Dialog Systems

Dialog Design

Dialog Controls.
Creating customized dialogs from common controls.

Tuning/ Feedback loop

Dialog Flow Analysis
Finding troublesome dialog flows via data analysis.

Discovery of Uncovered Semantics.
Finding frequent common inputs from users that a dialog system failed to cover

Better, Faster, Cheaper