



NUANCE

The experience speaks for itself™

**Call Center Speech Application Services:
The Good, The Bad, and the Ugly**

Ann Thyme-Gobbel
Cathy Pearl



You know your job has arrived when...

It's a 'Seinfeld' skit →



← It's a 'Saturday Night Live' skit



What makes these funny?

Presentation Outline

- Speech Recognition 101
- Professional Services in Network Speech
- Challenges of Creating a Speech App
- Case Study
- Typical Speech Project Steps
 - * VUI Design Process
- Q&A



3

Speech Recognition

- Different areas and usage
 - Embedded: Cars, toys
 - Dictation: Dragon, medical transcription
 - Mobile: Cell phones
 - Network: On the phone
 - Automated Tasks
 - Call Routers



4

How does it work?

1. Caller speaks into a telephone: "Traffic info, please."
 2. The computer processes this speech signal (waveform)
 - a. Compares to grammar ('traffic info', 'traffic report', 'traffic updates please')
 - b. Determines which one is most likely
 - c. Translates into a single result → [traffic]
- Only items in grammar will be recognized
 - 'Fillers' ok ("please", "um", "uh")



5

Our jobs: Nuance Professional Services

- Functional Organization
 - **Dialog Designers/Usability Engineers**
 - Speech Technologists
 - System Architects
 - Project Managers
- Organizational changes over time
- The dual role of the speech PS person
 - End-user advocate
 - Help customer reach (business) goal



6

Why do customers want speech rec?

- ROI: Return on Investment
 - Save money by reducing number of agents
 - Automate tasks that usually require agent
 - Get caller to the right agent to save re-routing costs
- Improved customer service
- Use cutting edge technology
 - Touch-tone is brittle
- Keeping up with the Jones'
 - Moving away from “press or say 1”
- *But is speech always the right solution?*



7

Speech is not the only answer

- + Providing an added benefit
- + No operator
- + No need to bother agent
- + Power users
- + Wide alphanumeric data (names)
- + User can't use hands
- + User needs access away from computer
- Agent is needed for task
- App can't handle the request
- User needs to feel reassured
- Task is “complicated”
- Task has high cost of error and/or non-expert user base
- User doesn't want to be overheard



8

Speech Challenges

- GUI < > VUI
 - “We’ve already designed it for the web, it’s basically done”
 - Speech is fleeting
 - Cognitive load
 - “Legal says we must say this.”
 - “Marketing says we need to tell callers this.”
 - “Add one more option to that menu.”
 - Navigation
- DTMF < > VUI
 - “Our callers are used to the old system; they need to do it the same way”
 - Misrecognition



9

Case Study: 511

- Local service for traffic info and public transit
- Replaced existing touch-tone system
 - Added new functionality
- Needed to enable DTMF everywhere
- No agent available
- Sample call



10

Typical Project Steps

1. Gather requirements
2. Create the design
3. Write the code
4. Test
5. Tune



11

Gather Requirements: Sample Business Requirements



- Minimize hold time or call length
- Free up agents
- Burn air minutes
- High level of security
- Very high accuracy
- Market differentiation
- Revenue generation
- Work well for mobile hands-free



12

Gather Requirements: Sample User Requirements



- Frequency of use
- Landline vs. Cellular
- Accents, dialects, languages
- Minimize cost
- Ease of use
- Confidentiality
- Confidence



13

Gather Requirements: Getting to Know the User - What Matters for Speech?

- Method: Listen to real calls made to the Call Center
 - **Goals:** What people need/want to do.
 - E.g., Pay bill by phone *and* get confirmation number
 - **Terminology:** How they ask for it
 - Sometimes they use business jargon: *"friends & family plan"*
 - Sometimes they don't: *"Did you get my last payment?"* vs *"Check recent account activity"*
 - **Knowledge:** What they know and don't know:
 - Their phone number? SSN? Spouse's SSN? PIN?
 - **Props:** What they have when they call:
 - Last bill? Credit card? Member ID card?
 - **Environment:** What's around them:
 - Using cell phone? in car?
 - In noisy home or office?



14

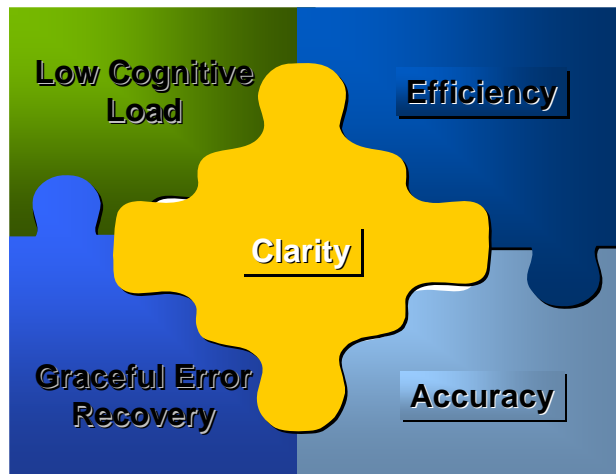
Create the Design

- Purpose:
 - Create detailed definition of dialog flow
 - Based on requirements, call flow, usability
 - Every possible path
 - Everything a system can say to a caller (“prompts”)
 - Handling errors (when things go wrong..)
- Deliverable:
 - Dialog Design Specification (input to development)



15

Design: Core Principles of VUI Design



16

Design: Minimize cognitive load

- Menus
 - 4 – 5 options per menu
 - Allow barge-in
 - Prompt types / style
 - Short phrases for options
 - Parallel structure for options
 - Allow multiple pieces of info



17

Design: Provide easy exits

- In prompts
 - ...When you're done, feel free to hang up.
 - ...Or to speak with someone, say 'representative'...
 - ...If you don't know the loan number, say 'I don't know it'.
- In grammars
 - Universal commands: help, repeat, operator, cancel, start over, go back...
 - "I wanna talk to a real person!"



18

Design: Prevent errors

- Clear prompt → caller knows what to say → lower “out of grammar” → smaller grammar needed → higher recognition success

Probably Not:

“What would you like to change?”

Instead:

“Which phone number do you want to change--
home, office, cell, or fax?”



19

Design: Provide easy error recovery

- Errors can occur even when the user does everything right, i.e., the recognizer can misrecognize what the user said.
- Anticipate in design
 - Confirmation strategies
 - Multi-tier
 - Backoff strategies
 - Universal commands
 - More explicit error handling prompts



20

Design: Provide easy error recovery

- Take the blame

System: "Sorry, I didn't get that..."

System: "I heard..."

- Focus on telling the user what to do, not what went wrong
- Use escalating information on subsequent errors
- Anticipate errors and consider *why* the user might be making an error or not responding
- Use one-step correction to capture not just the answer, but the correction too

System: "That was 312-555-5221. Is that correct?"

Caller: "No, it's 312-555-9221."



21

Test: Usability Testing & WOZ

- Challenges: Performance & Coding effort
 - A working prototype requires adequate grammars and recognition performance
 - Grammar development and tuning are not yet done when we want to usability-test a prototype
- Solution: Wizard of Oz (WOZ) prototypes
 - Task scenarios
 - The human Wizard simulates the application logic *and* the speech recognition engine
 - The Wizard plays out pre-recorded prompts over the phone, in response to what the Participant says



22

Tune

- Analyze and optimize performance on REAL USER DATA
 - Transcribed data
- Typical Deliverables
 - Tuning report
 - Modified grammars
 - Recommendations



23

Top 10 questions / requests from customers

1. How well does this stuff *really* work?
2. Male or female voice?
3. Where can I put my marketing messages?
4. How much do we have to educate the caller that they're talking to a machine?
5. Should we have touchtone back-up everywhere?
6. How do I know how well my app is working?
7. Why do you have to usability test it if you're such experts?
8. Can we use our employees in usability tests?
9. And what about The Holy Grail of agent access?
10. Why can't that prompt say...?
(a.k.a "I speak English – I can design this!")



24

Q&A

Questions?



25



NUANCE

The experience speaks for itself™

Nuance Professional Services is hiring!

Contact: annt@nuance.com

