People Say the Darnedest Things

Charles Galles
Principal Speech Scientist
Overview

The Human Element
To Speak or not to Speak
People are Talking
Well … I never …
Short vs Long Phrases
The Human Element

**Voice User Interface Design Myth:**

» “I can decide what callers are allowed to say” by
  – Restricting “operator” opt-out
  – Ignoring caller phrasing patterns

**Voice User Interface Design Reality:**

» I have significant influence over what callers will say
  and how the system reacts
  – Hidden menu options
  – Direct and indirect menu prompts
  – Inactive phrase behavior

**Successful VUI Designs must adapt to the Human Element**
Comparing Caller Reactions

*Three open prompt systems compared*

» S1 – Financial industry, customer self-service
  – Caller must authenticate before main menu
  – Directed dialog lead-in, followed by open prompt

» S2 – Multimedia content provider, self-service application
  – Caller is greeted with open prompt main menu
  – Single attempt at main menu, directed dialog following

» S3 – Also a Multimedia provider, self-service application
  – Initial prompt is main menu
  – Multiple dialog states use open prompt grammar
  – Multi-slot implementation recognizes caller intent, product, or both
To Speak or not to Speak

“How may I help you today?” - A prompt that sometimes strikes fear into callers

- System 1: “How can I help you today” (examples on reprompt only)
- System 2: “In a few words, tell me what I can help you with today. <pause> you can say anything from … ” - with 2 examples
- System 3: “How can I help you today. <pause> you can say anything from … “ - with 3 examples including “operator”

SLM Response Dispositions

S3’: S3 system but excluding operator responses
The needs of the many … are few

Call routing applications tend to have unequal category distributions

- 60-70% of utterances in top 5 categories
- Another 20% in next 15 categories
- Remainder split among 40-60 categories
- Data collection requirements:
  - 200 samples @ 0.5%
  - Requires 40000 utterances
  - And 57000 calls

Category Distributions

![Category Distributions Chart](chart.png)
What people want

» High runner categories can be common across domains
» But frequencies of those categories are domain-specific
» High-runner categories on evaluated systems
  – Technical Support (TS)
  – Customer Service (CS)
  – Bill Payment (BP)
  – Account Balance (AB)
  – Account Activity (AA)
  – Account Changes (AC)
Do People Really Say That?

In some cases callers are simply echoing back pieces of the prompt:

- **Sample Main Menu Prompt:**
  - … Please tell me briefly what you're calling about, for example, you could say: "report a payment" or "billing inquiry." Or ask for whatever you need …
  - Nearly 1% of all responses were variants of “ask for whatever I need”!
  - “ask for whatever I need” was one of the top 100 utterances at the main menu prompt

- **Other gems**
  - “what I am calling about”
  - “how can you help me”
  - “why am i calling”
  - “you called me”
Suggestions

*Improvements to the prompt can reduce “parroting”:*

» Design suggestions
  – Include phrases like “in your own words”, “tell me briefly …”
  – These phrases should be used near the end of the prompt
  – Don’t prefix the question with “or”
  – … Please **tell me briefly** what you’re calling about, for example, you could say: "report a payment" or "billing inquiry." **Or ask for** whatever you need …

» Better prompt sample
  – “You can say anything from “pay my bill” to “technical support for widgets”. So in a few words, tell me what I can help you with today.”
  – Parroting effect was greatly reduced (all three systems less than 0.01%)
**Sometimes responses are not for the faint of heart**

» Frustration triggers
- Apology prompts: “I’m sorry I didn’t hear you”, “I’m sorry, I didn’t understand you”
- Long sequences of navigation or caller input
- Some choice selections:
  - you can't hear me you must be deaf …
  - I wanna talk to a @#$@!! operator
  - I want a real live person you moron
  - Oh … here we go again

**Frustration Rates by System**

![](chart.png)

*S4 original S3 application*
The long and short of it

*What should you expect in terms of response length*

» Most responses are between 2 and 4 words

» Some systems have key phrases that are one word
  - Billing
  - Sales
  - Representative

Phrase Length Distribution
Top ten (or so) favorites

You just never know:

- It’s … our tv went kooky
- My dog ate my remote control
- My dog stepped on the programmer and it went pssh and now I can’t …
- you don’t understand nothing because you are not a human being …
- You called me, you tell me
- Mommy’s gonna put in Shrek for you
- Are you some kind of a tape or something