

Yesterday's Enterprise – VUI Design Lessons from Yesterday's IVRs Applied to Today's Conversational IAs

Ann Thymé-Gobbel,
D+M Group (Denon Marantz)
ann.thymegobbel@dmh-global.com

Sondra Ahlén,
SAVIC Voice Interface Consulting
sondra@savicvoice.com

Those Babies and the Bathwater

“IVRs didn’t understand you – and they’d tell you that over and over”

“Everyone hated IVRs – they were all long trees keeping the caller from agents”

“Viable voice solutions started just a few years ago with Siri & Alexa”

“Nobody understood conversational dialog till now”

“Reco is a solved problem”

“What’s an IVR? Isn’t that just ‘Press or say 1’? Why should I care?”



“Some of this ‘story’ is motivated by a desire to differentiate the current offerings”

“For 15+ years, IVRs have used voice to conduct dialogs with users that were deemed (by the users) as satisfactory and sometimes even preferable to live agents.”

“Sure, many are, but how many shitty websites have you been to? Does that mean they're all bad?” “You just remember the bad.”

“Look at the good ones and **learn** so you don’t make mistakes on a new platform that actually have solutions.”

“**Show** them the highly sophisticated statistical systems, as well as all the fun multislot stuff”

“Well-designed IVRs got users what they wanted quickly without trees”

-- IVR experienced VUI designers

plus ça change, plus c'est la même chose

What is EASIER today?

- ✓ * ASR accuracy; broader coverage 'for free'
- ✓ Less need for targeted grammars & tuning
- ✓ NLU toolkits & other tools available
- ✓ Explaining what VUI design is!

What is DIFFERENT?

- Tasks are simpler, less multi-turn needed (so far)
- Users are driving the interaction, happier
- Content omnipresence / omnichannel
- Multiple users in a room
- Quality TTS availability

What is HARDER today?

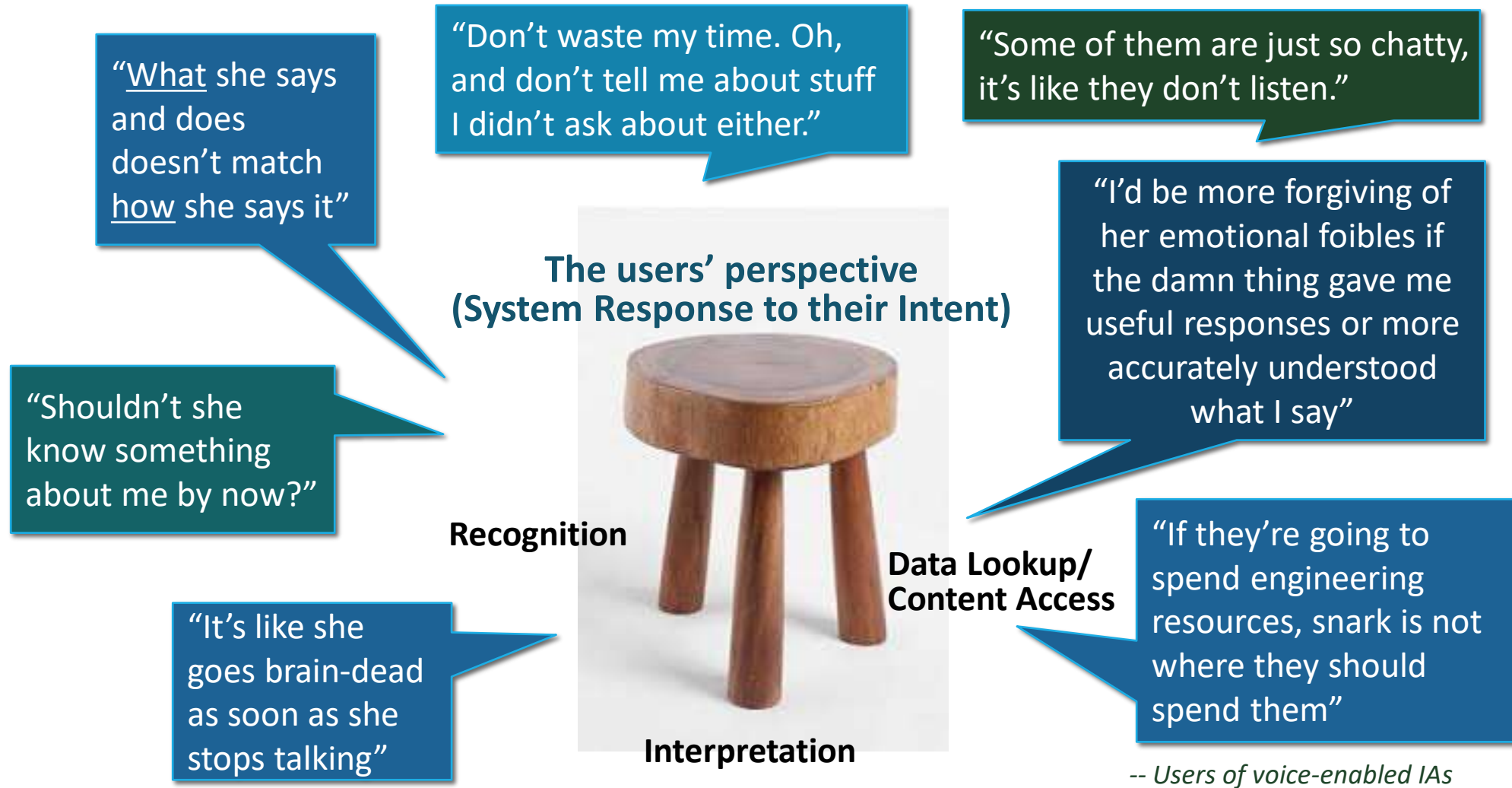
- * Expectations are higher
- 'Guts' of ASR not as available, precise control is difficult & impacts dialog design
- Multi-turn dialog with multi-slot
- Nothing is 'contained', broad content coverage
- Bad PR can make the rounds quickly

What is the SAME?

**Humans are humans. They can say ANYTHING.
You still need the same conversation model / dialog
principles / Gricean maxims / discourse structure...**

There are no cookie cutter solutions, but best practices

The 3-Legged Stool of Voice UI/UX Design



Users don't care what's under the seat – they'll just see that you didn't give them what they wanted

“It’s like she goes brain-dead as soon as she stops talking”

- Conversation = dialog = multiple turns = keep context
- Context allows you to build on & refine earlier dialog
- Types of context
 - **Current dialog turn**, Most recent dialog, Real world, Preferences
- Context stickiness
 - 15-20 seconds may be a good start

“It’s like she goes brain-dead as soon as she stops talking” (2)

IVR Multi-Slot

“Transfer 500 dollars from my joint savings to my premier checking account tomorrow.”

“Transfer from joint savings to premier checking account tomorrow.” → “How much?”

“Transfer 500 dollars to premier checking account tomorrow.” → “From what account?”

Home IA VUI Without Context

“Find me action movies with Sean Connery” → [result]

“Show action movies with Sean Connery I don’t have to pay for” → [result]

“Show action movies with Sean Connery I don’t have to pay for that start within the hour?” → [result]

Home IA VUI With Context

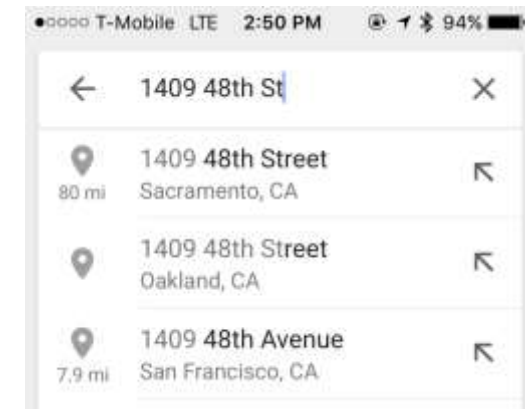
“Find me action movies with Sean Connery” → [result] →

→ “Show only the ones I don’t have to pay for” → [result] →

→ “Which ones start within the hour?” → [result]

“Shouldn’t she know something about me by now?”

- Types of context:
 - Current turn, Most recent dialog, **Real world, Preferences**
- Geo location for best match
- Dynamic options = only relevant choices
- Be proactive, tailor & predict – but don’t be creepy!
 - IVRs utilized this successfully
 - Caller ANI + frequent flyer acct = immediate upcoming travel info
 - Caller ANI + acct status = prescription reorder status, or bill payment



Let's Play Ball

If I ask a voice-enabled IA for a baseball score...

User: "What's the Giants score currently?"

VUI: "The San Francisco Giants or the New York Giants?"

Is this a good response? Why or why not?

What knowledge would impact your decision?

“Don’t waste my time”

- Be concise – what does the user really want to hear?
- Non-verbal audio / ‘earcons’
- A long interaction may still be a great success
 - IVR: stock price information
- Implicit confirmation – acknowledge while moving dialog forward

VUI: "Ok, which bill do you want to pay?"

VUI: "From San Francisco to where?"

“I’d be more forgiving if the thing gave me useful responses or understood what I say”

- Confidence level + Tiered handling = More trustworthy responses
- Trust: If users ask for something unambiguous, don’t give them something else
Forgiveness: If you acknowledge their request, users are more forgiving...

User: “Play Let it Be”
VUI: “Okay. Let it Be Love”

User: “Play Let it Be”
VUI: “Here’s what I found: Let it Be Love”

...even if you can’t fulfill it

VUI: “I can’t do that”

VUI: “I don’t know how to XYZ yet”

- Skip lists → One Step Correction

Without

User: “Call Johan”
VUI: “Calling Johanna Smith”
User: “No! Call Johan!”
VUI: “Calling Johanna Smith”

With

User: “Call Johan”
VUI: “Calling Johanna Smith”
User: “No! Call Johan!”
VUI: “Calling Johan Edklund”

“If they’re going to spend engineering resources, snark is not where they should spend them”

- Responding to formulaic discourse

IVR: “... Hang up now.”
User: “Goodbye.”
IVR: “Thanks. Goodbye.”

IVR: “... “
User: “Thank you.”
IVR: “You’re welcome.”

- Personas – Be careful with wording & regional dialect
- Humor is tricky
 - Allow the user to personalize
 - Be sparse & choose wisely

“What she says and does doesn’t match how she says it”

- Pronunciation Pitfalls
 - Incorrect pronunciation of locations or artists
 - Prosody – beware of incorrect focus

- The dangerous ‘A or B’ structure:

“Do you want A or B?” vs. “Which do you want: A or B?” vs. “Do you want A?”

- Confirmation should match the question – and the context

Poor

VUI: “Say the street address using single digits”

User: “One four two five Broadway”

VUI: “Fourteen twentyfive Broadway...”

Better

VUI: “What’s the address?”

User: “Fourteen twentyfive Broadway”

VUI: “Fourteen twentyfive Broadway...”

You Must Remember This...

You can't control what users say –

but you can often constrain it with prompting & flow

Good recognition is not enough –

you need to interpret & handle the results

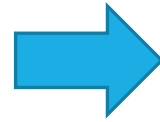
- Look for trends and unexpected responses you need to
- Understand why they said what they said
- Adjust

Learn from IVR techniques & modify for today's conversational apps

Your users will be happier – therefore you will be too!

Tomorrow's Conversation: Know Your Purpose and Context

From: Goal-oriented IVRs/Apps – Self-service, call routing, customer service, simple command & control, directed, mixed-initiative, and open-ended 'How can I help you?' prompting



To: Broader dialogs and conversations: assistants, planners, topics of interest, queries, reviews, referrals, conversationalists, mixed-meandering, and new platforms and technologies



Identify requirements: business, user, demographic, language, social, cultural, functional, technology, resources

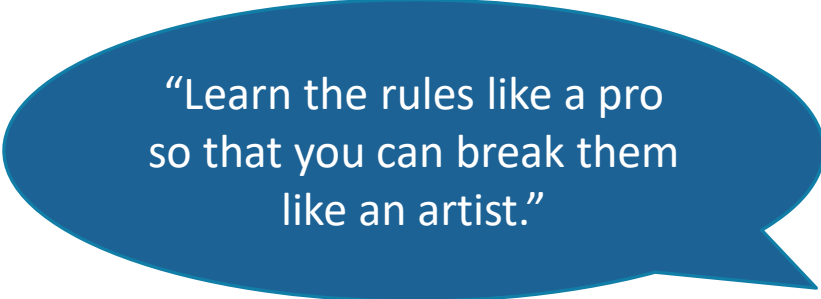
Aim-Less* does not mean Purposeless... Social purpose of conversation

* -Daniel Meneke

"Dave, this conversation can serve no purpose anymore.
Goodbye." -Hal9000, 2001: A Space Odyssey

Apply Principles and Best Practices

- Leverage years of IVR experience in methodology, requirements gathering, design, usability testing and tuning
- Find creative ways to collect data, analyze to understand how conversational interfaces are changing
- Look to highly relevant treatments in cognitive science, linguistics, psychology, playwriting, storytelling, etc.
- Industry Resources: Experts, books, conference talks, white papers, articles, etc.



“Learn the rules like a pro
so that you can break them
like an artist.”

-Pablo Picasso

Yesterday's Enterprise – VUI Design Lessons from Yesterday's IVRs Applied to Today's Conversational IAs

Ann Thymé-Gobbel,
D+M Group (Denon Marantz)
ann.thymegobbel@dmh-global.com

Sondra Ahlén,
SAVIC Voice Interface Consulting
sondra@savicvoice.com