The Conversational UI and Why it Matters

Table of contents

<table>
<thead>
<tr>
<th>Table of contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation basics</td>
<td>2</td>
</tr>
<tr>
<td>Conversation—a study in cooperation</td>
<td>2</td>
</tr>
<tr>
<td>Creating better conversations</td>
<td>2</td>
</tr>
<tr>
<td>Validating user input and managing expectations</td>
<td>3</td>
</tr>
<tr>
<td>When a conversation breaks, repair it</td>
<td>3</td>
</tr>
</tbody>
</table>
We’re entering a promising new era of computing, where advances in machine learning and artificial intelligence are creating a resurgence of interest in voice interfaces and natural language processing, unlocking the potential for conversation as the new mode of interaction with technology.

For the most part, the problem of recognizing spoken input has been largely solved, opening up a new challenge: How to build a user experience that’s modeled after natural human conversation.

This multi-part series outlines the basic mechanics of conversation, introduces core principles to design by, and presents you with a practical voice user interface (VUI) toolkit to start creating conversational experiences that engage, delight, and truly help your users.

**Conversation basics**

We can discover the key ingredients for a good VUI conversation simply by deconstructing the rules and conventions that we are mostly unaware of in natural human conversation.

Building blocks for successful VUI conversations include:

- **Turn-taking:** We take turns in a conversation based on subtle signals that we pass back and forth. Without effective turn-taking, we might talk over each other, or our conversations can get out of sync and hard to follow.

- **Threading:** In natural speech, all elements of a conversation are usually woven together in a coherent thread that includes context and the way the conversation evolves over time. Threading helps us keep track of that conversational flow.

- **Inherent efficiency of speech:** People often use verbal shortcuts because they intuitively understand what’s being said—essentially, we can “read between the lines” in a conversation and some things can be left unspoken. But the implication for VUIs is having to compensate for the seemingly illogical, non-mathematical nature of human speech.

- **Variations in user behavior and comfort level:** People use a variety of words and styles to say the same thing, depending on their own situational context and expectations of a VUI based on previous experience. VUIs should support these variations so all users can have a frictionless experience.

Read more on conversation basics in Understanding How Conversations Work: The Key to a Better Voice](https://developers.google.com/actions/design/principles)

**Conversation—a study in cooperation**

Turn-taking, context, and threading are all part of cooperative conversation, an idea popularized by linguistics philosopher [Paul Grice](https://plato.stanford.edu/entries/grice/). Grice called this the Cooperative Principle. He also developed Grice’s Maxims to define the essential conversational rules he observed—namely that people should be as truthful, informative, relevant, and clear as possible when talking with each other.

VUI should try to follow these inherent rules of cooperation as well—and be ready to support wary users who’ve had bad experiences with other voice interfaces in the past.

Read more on the Cooperative Principle in Be Cooperative… Like Your Users". 

1 https://plato.stanford.edu/entries/grice/ 
2 [http://g.co/dev/actions-be-cooperative](http://g.co/dev/actions-be-cooperative)

**Creating better conversations**

A good VUI doesn’t follow a stale script, and shouldn’t be based on the old touch-tone phone systems that force users down narrow paths. It also shouldn’t try to “teach” people what to say to protect them from veering off the so-called “happy path.”

Instead, it should focus on the intuitive power of language and meaning, using everyday language to communicate with users. VUIs should also avoid stating the obvious or talking down to users—people
don’t appreciate a device that sounds like it thinks it’s smarter than they are.

Read more on building an intuitive VUI in Unlocking the Power of Spoken Language.
(http://g.co/dev/actions-spoken-language)

Validating user input and managing expectations

Once someone makes a request, a VUI can use acknowledgers—words or phrases like “OK,” “Sure,” “Alright,” “Thanks,” and “Got it”—that show the VUI is engaged and listening. Randomizing acknowledgers can help make the experience feel more fluid and natural.

After acknowledging, the system can then seek either explicit or implicit confirmation that it understood.

With explicit confirmation (which is usually used when something major is at stake, such as buying a plane ticket), the VUI asks the user for verbal agreement before proceeding.

With implicit confirmation (used for lower-risk situations, such as streaming a song), the VUI incorporates a key element of the user’s request into its response to validate and instill confidence in the user, but doesn’t require verbal approval per se.

Read more in Instilling User Confidence Through Confirmations and Acknowledgements.

When a conversation breaks, repair it

Things can go wrong with even the most well-designed VUI. When conversations break down, either party can initiate a repair. Though humans usually spot and repair their own mistakes, a VUI must also be able to repair the conversation based on the flow and nature of the interaction.

Read more on repair in Understanding How Conversations Work: The Key to a Better Voice UI.
(http://g.co/dev/actions-understanding-conversations)

VUI Do’s and Don’ts

Do this …

• Follow basic conversational rules and everyday speech patterns
• Apply the principles in Grice’s Maxims
• Suggest through intuitive examples of what people can say (without “teaching” them)
• Use acknowledgers to show that the system is listening
• Randomize acknowledgers to make the VUI sound more natural
• Use explicit confirmation for clarity around important requests, and implicit confirmation if a request is less risky to fulfil
• Use “errors” as opportunities for more meaningful (and natural) interactions

Don’t do this…

• Model responses on touchtone apps and phone-tree scripts
• Try to teach people what to say
• State the obvious
• Talk down to users or offer responses that sound mechanical
• Use “Go back” instructions
• Accommodate just one speaking style
• Use language that talks down to users

© 2016 Google Inc. All rights reserved. Google and the Google logo are trademarks of Google Inc. All other company and product names may be trademarks of the respective companies with which they are associated.