Understanding How Conversations Work

—

The Key to a Better Voice UI

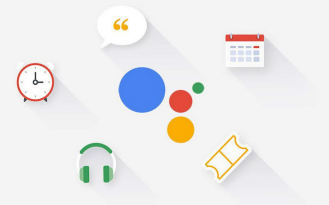## Actions on Google

"
Twenty-five years from now, no one will be clicking on drop-down menus, but everyone will still be pointing at maps and correcting each other's sentences. It's fundamental. Good information software reflects how humans, not computers, deal with information.

Bret Victor, Magic Ink

Let's face it. Right now, most voice user interfaces (VUIs) fall short of the future we were promised by science fiction—surrounded by artificial intelligence and effortless conversations with clever robots and smart appliances.

So how do we get there?

For starters, we have to teach our machines to talk to humans, not the other way around.

Consider this: Conversation has advanced our civilization to where it is today. All human inventions are born from the ideas we communicate through spoken words—an ability we evolved over a very long time. Over 100,000 years in fact. Compare that to the roughly 5,000 year-old infancy of writing, let alone computing.

"So people are obviously not going to change how they talk anytime soon. And their unconscious expectations about how conversations are supposed to work won't go away either."

Whether we're aware of it or not, we all follow specific rules and conventions when we talk. If we can decon-struct what makes for a good human conversation, we might be able to figure out how to build better VUIs.

The 6 steps of conversation

The basic mechanics of a conversation can be broken down into six simple steps:

1. **Open a channel to set up common ground — Speaker A sends a message to speaker B**

2. **Commit to engage — B commits to the conversation with A**

3. **Construct meaning — A and B connect through a set of structured ideas and (often unspoken) contexts**

4. **Evolve – A or B (or both) learned or gain something based on their interaction**

5. **Converge on agreement — If everything works, A and B have reached an agreement; if not, both may move to repair the situation**

6. **Act or interact — Functional action may follow as a result of the conversation, or some unconscious goal may be reached (being less lonely counts).**

Turn-taking

A somewhat obvious, yet important, instrument of conversation is turn-taking, which involves subtle signals we take for granted. Syntax helps listeners predict that a chance to respond is coming, while prosody—a combination of pace, volume, pitch, and silence—signals that a transition point is coming. People use these cues to hand the conversational baton back and forth to each other. Without effective turn-taking, we either talk over each other or the conversation gets out of sync.

The Cooperative Principle

Linguistics philosopher Paul Grice, whose work has also been applied to artificial intelligence, said that to be understood, people need to speak cooperatively. He came up with some basic rules of cooperative conversation called Grice's Maxims, noting that people have to be as truthful, informative, relevant and clear as the situation calls for. See also: *Be Cooperative … Like Your Users*.

Implicature and context

The meaning of a conversation depends on its context. But in normal conversation, often what we leave unsaid carries meaning, too.

Say you ask a friend, "Are you going to the party on Saturday?" and she replies, "I work the evening shift." Your friend is implying that she can't be in two places at once, so you then infer that she won't be coming to the party.

Or in another context, when asked how many people to make a reservation for, and you say "Oh, just me and my husband," you're expecting the other person to infer you're booking for a party of two.

If we didn't have these presumptions and principles operating in the background, our conversations would need to be super-literal to be understood.

Threading

All elements of a conversation should be bound together in a coherent thread. As anyone who appreciates a good knock-knock joke knows, threading the context of each turn in a conversation keeps the relevance going and enhances the overall exchange.

Threading example #1

To do this successfully, the designer should maintain an awareness of each pair of turns in the conversation (called an adjacency pair) as it unfolds:

What time is it?

It's exactly noon.

Threading example #2

It doesn't have to be a question-answer pair, either. Listenership signals form adjacency pairs:

You won't believe what happened!

What??

Threading example #3

As do endorsements or negations of what was just said:

This is the worst line ever!!

I know!

Threading example #4

If a VUI doesn't produce a nice threaded exchange full of bound adjacency pairs, the conversation quickly becomes broken, or simply uninteresting. As such, threading is an essential tool for making an experience more engaging for the user, as in this game example:

Welcome to the Numberdini Quiz! I'm thinking of a number from 0 to 100. What's your first guess?

50

It's lower than 50. Next guess?

How about 9.

Piping hot! Go lower.

Okay ... is it 8?

Keep going.

7?

Yes1 It's 7. How about another round?

Repair

Broken conversations can also result from a lack of common ground. Meaning may also start to unravel through inappropriate contributions that violate Grice's rules of cooperative conversation.

"For example, if a person is asked "Do you know who's going to the party?" and they answer simply "yes," it's uncooperative and unnatural, making it awkward to repair."

Even within a functioning conversation, form or content may be inaccurate, inappropriate, or nonsensical, requiring repair to get things back on track. Either party can initiate a repair, in or out of turn, but there's a general order of preference, and speakers usually spot and repair their own errors. A VUI needs to be able to repair the conversation based on the flow and nature of the interaction. Read more on repair strategies in Being a Better Conversation Partner.

Some final words

Conversation is a principled, mutual process of collaboration and negotiation. All parties involved create and agree upon meanings and operate against a background of rich, nuanced context. Understanding this can give you a theoretical model for designing your own conversational VUI.

Best practices

Remember these basic building blocks
when creating a VUI experience:

→ **Six steps of conversation**   → **Implicature and context**

→ **Roles in turn-taking**   → **Threading**

→ **The Cooperative Principle**   → **Repair**