SUTHERLAND

VOICE IS BACK

PART 2

The voices-everywhere dilemma

Andrew B. Zimmerman, President, Sutherland Digital, Technology, and Media & Communications
Owen Daly-Jones, Senior Vice President & Global Head, Sutherland Labs

SMART SPEAKERS LIKE AMAZON ECHO AND GOOGLE HOME ARE MAKING BIG WAVES IN THE CONSUMER WORLD.

It's a fascinating tech phenomenon because interactive digital voice systems have been around since the 1950s, and voice recognition is arguably commonplace on smartphones, telephone customer service systems and in-car infotainment. It's not new – nor has it even been particularly noticeable before - but unit sales of new voice assistant devices are skyrocketing faster than smartphones when they first appeared. So, there's clearly something new and noticeable happening to digital voice now. Answering precisely what that something is – and leveraging that insight commercially – means understanding voice customer experiences, but that's more complex than it first sounds.

The challenge is digital voice 'smart home' devices and assistants (a wide range of gadgets and embedded services, sometimes referred to as 'voice first') are unlike previous must-have consumer technologies like smartphones, tablets and PCs. That's because no matter where you are, it's likely you have your preferred mobile device nearby, and using it is a task-focused, one-to-one user experience. In contrast, we share smart speakers and use them in specific locations, usually while we're doing something else like cooking, reading, relaxing on the sofa, showering or doing whatever. We could be doing whatever alone or in a group, and the way we do whatever varies depending on whenever and wherever we do it. That's a huge range of variables that affect the things we need a voice assistant to do, the way we ask it to do them, and the way we need it to respond.

The sheer diversity of digital voice customer use-cases presents an interesting design dilemma:

How do you create delightful voice experiences for customers of all ages, that use the same digital voice devices and assistants while they're doing whatever, wherever and whenever?
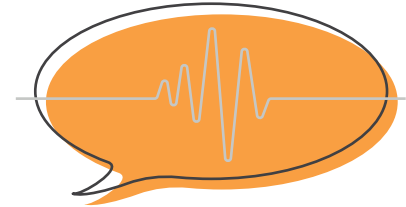
That's the 'voices-everywhere dilemma', and mastering its complexity means going into the field to study your customers while they engage with their new digital voice companions.

UNDERSTANDING THE CONTEXT OF VOICE INTERACTIONS

Digital voice devices such as Amazon Echo and Google Home are most commonly used in hands-free multitasking scenarios. Multiple studies report the kitchen is the most popular location for smart speaker use in the household, followed closely by the living room. This is the first clue that digital voice experiences comprise many diverse, context-specific use cases that need careful consideration, because the kitchen and living room encompass a wide range of different possible activities.

We use these spaces to passively consume media (like relaxing with the TV) and consume it in the background when we're busy (like a radio in a kitchen). However, we also use our new voice-activated companions with more active kinds of listening, like asking them to tell jokes or play games, or assist with the cooking process (using them more like a game console or smartphone app). These activities could involve a single individual or a group, across the whole household age range, which means designing digital voice experiences isn't simply about understanding user-device interactions, it's about understanding a day in the life of the household.

> "People usually start out with one smart speaker, so the kitchen or living room – both high traffic, public spaces - tend to get priority. Then the device, especially the smaller ones like Echo Dot or Google Home Mini, gets trialed in different places. Kids might take them around the house to play with, adults might try them out in bedrooms too for alarms, news and music. It is a very organic relationship."
>
> **DR. IMOGEN CLARK**
> DESIGN RESEARCHER
> SUTHERLAND LABS

User voice experiences also vary depending on the type of hardware, and the voice assistant we're using on it. On smart speakers such as Amazon Echo and Google Home, we're most likely to set a timer, play music, check the news or set an alarm. On smartphones, Apple and Google's voice assistants are used mostly to search the internet and get directions. You can use Alexa on mobiles too, but 'she' often gets asked to do same things she does at home via a smart speaker. Again, it's highly contextual, however one common feature in all these scenarios is voice commands are mostly simple requests. Buying products or managing calendar are possible with most digital voice systems but are much less frequently accessed because voice interfaces still have a long way to go, usability-wise.

> "Most users want to access more advanced voice tasks, but success is lower because the commands use more complex syntax that's more likely to misunderstood by the assistant. Also, noisy acoustic environments and accents can significantly reduce the efficacy of voice recognition, meaning it's quicker to do some things the old-fashioned way... with a screen."
>
> **SIMON HERD**
> DIRECTOR OF DESIGN RESEARCH, SUTHERLAND LABS

THE DIGITAL VOICE DESIGNER'S TOOLKIT

ETHOGRAPHIC FIELD WORK

Ethnographic field work means observing customers using digital voice at home, in their cars, at work – anywhere there's a need to design or improve a customer experience. It is a way to record the real-world context of a voice experience, considering physical factors like room acoustics; logical factors like using trigger words and commands; and emotional factors that affect user experiences, because talking to a voice assistant encompasses many different user modes and moods (e.g. elbow deep in making a soufflé, relaxing with a good book, doing your 3rd Grade math homework, rushing to go to work, and so on). Ethnography provides designers with a 360-degree insight into the relationship between the voice user interface (VUI) and whatever else the user is doing at the same time, capturing the voice assistant user needs of the whole household from younger children to teens, parents and grandparents.

USER DIARY STUDIES

Diaries are particularly useful because voice assistants are often performing background tasks (like playing music) so users tend not to pay much attention to the interaction because they're usually focused on doing something else at the same time. Diary studies help reveal how the 'on in the background' device fits into the ebb and flow of household life.

VOICE ASSISTANT APPS

Most voice assistants have companion apps that provide a history of their conversations and responses, which gives a vital insight into how successful – and unsuccessful – user interactions have been. These records compliment ethnographic observations and diary studies to help design better conversational responses, and refine the linguistic interactions required for executing commands.

DIGITAL VOICE EXPERIENCES
CREATE UNEXPECTED FRICTION
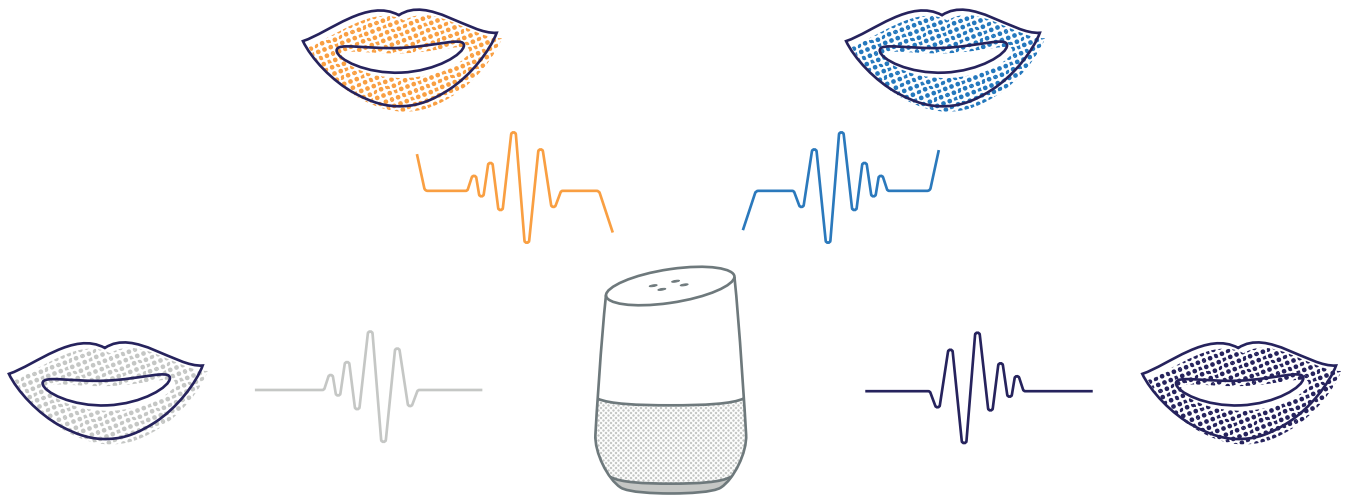
At first, household digital voice devices appear to be familiar one-to-many products like TVs and radios (i.e. a device the whole family can share.) But when it comes to setting them up and controlling their features, these are generally one-to-one devices linked to an individual user, more like smartphones and PCs. This can make digital voice a confusing beast.

Our research shows that although the whole family might use the same device, usually there's one person in the household that configures it with their personal ID. When you consider these devices are following a slightly unusual adoption curve, with 25-34-year-olds being the heaviest adopters but often buying them for older relatives and children, this creates confusing problems arising from managing multiple user accounts on one shared device, or multiple devices on one user ID.

You often hear from digital voice consumers about problems to do with profiles and ID. One person we met had a couple of Echo devices himself, and he'd also bought one for his father and set it up in the father's home. But because all three devices were tied to his Amazon ID, when he tried to call his father's Echo, his two home devices would ring too. Not great."
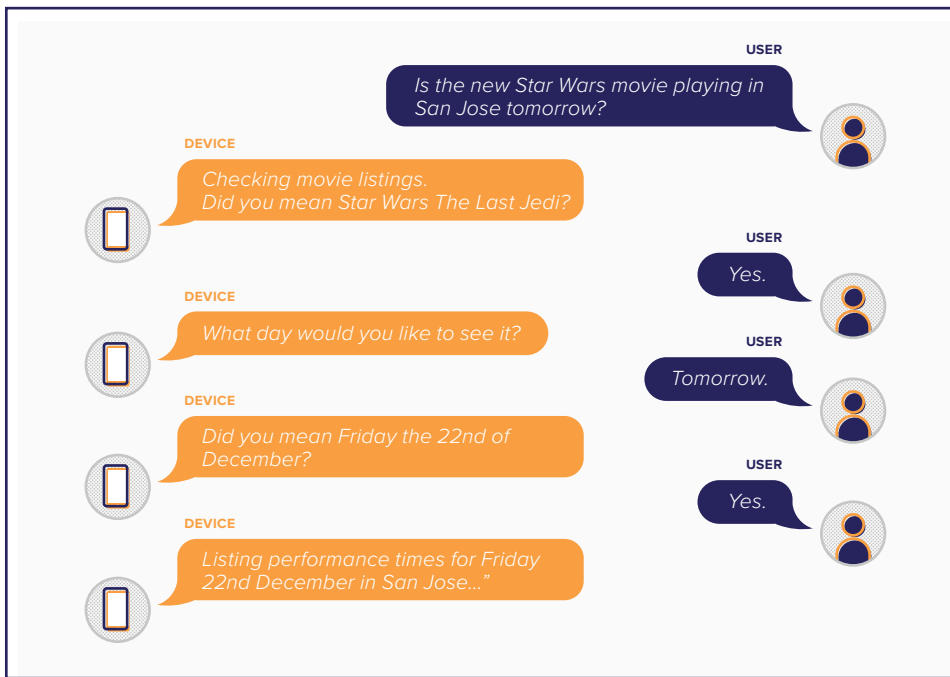
DR. IMOGEN CLARK
DESIGN RESEARCHER, SUTHERLAND LABS

Other typical multi-user friction points arise where one person shares a device with other household members and secures their personal payment details for shopping services (preventing others from accessing the shopping functionality on offer) or limits other users' options to personalize their own content preferences by associating the device with their own apps and preferences. In households with multiple devices, this creates a situation where identical devices have variable utility depending on who's asking.

To combat customer pain points, Amazon Echo and Google Home address these challenges with a voice setup experience in which each household user 'trains' the device to recognize their voice and associate it with their own respective profile. This process requires relatively complex configuration steps and isn't optimized for children, so it's considerably more complex than managing a family game console or household Netflix account, for example, while digital voice assistants aim to address every user and their respective customization preferences, the technology is not one-size-fits-all — *yet*.

STUBBORN CUSTOMER PAIN POINTS

Many digital voice pain points come from the challenge of processing natural language. Capturing data from voice interactions isn't always easy because humans conversationally give and receive data in a multiplicity of ways. For example, booking movie tickets: You might start with the name of the movie you want to see, the number of tickets you'd like to buy, the theater location nearest you, or a specific time and date term like 'tomorrow' or 'late afternoon'. You might also present the request as a question *"Can I book two tickets for…"*, a demand *"Give me two tickets for…"* or a neutral statement *"two tickets for…"*

Successful voice user interface VUI conversations navigate this linguistic complexity by guiding the user gently through structured scripts, like this:

USER
Is the new Star Wars movie playing in San Jose tomorrow?

DEVICE
Checking movie listings.
Did you mean Star Wars The Last Jedi?

USER
Yes.

DEVICE
What day would you like to see it?

USER
Tomorrow.

DEVICE
Did you mean Friday the 22nd of December?

USER
Yes.

DEVICE
Listing performance times for Friday 22nd December in San Jose…"

However, the linguistic data capture challenge is often complicated by the basic problem that successfully recognizing voice inputs doesn't necessarily mean a voice assistant has understood your intent, leading to unpredictable results, e.g. *"Sorry, I can't find 'is the new Star Wars movie playing in San Jose tomorrow' in your music library"*. This leads to user frustration, reducing future interactions for the same task.

Another common pain point is changing a voice command once you've started speaking. It's difficult to 'go back one step', or edit parameters like the duration of a timer, an alarm time or a calendar entry *after* you've set it. Many people get frustrated, set duplicates by mistake, or simply delete and start over.

> *Conversation design needs to emphasize error handling and prevention. Prompting users to give information in a structured, recognizable format is an essential VUI design technique. You could say that where psychology was the dominant discipline underpinning the evolution of graphical user interfaces, linguistics is equally important to VUI and chatbot design.*

SIMON HERD
DIRECTOR OF DESIGN RESEARCH
SUTHERLAND LABS

"The need to remember trigger words and intricate sequences of commands currently compromises more complex interactions. Practically, this means that the full range of everyday tasks are often easier to execute on smartphones and laptops."

DR. IMOGEN CLARK
DESIGN RESEARCHER, SUTHERLAND LABS
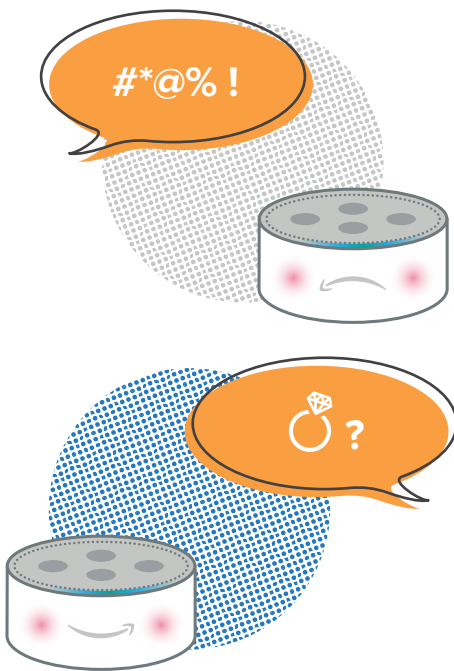
Human voices inject another unpredictable variable into voice interactions, namely emotion. People respond more emotionally to voice interfaces: They are both more easily frustrated and delighted. This explains why Alexa received over 250,000 marriage proposals in less than a year after launching in the US. This places more pressure on VUI designers to get interactions right, because people don't shout and swear at a badly designed screen layout, but they do at voice assistants.

Additionally, swapping between different voice assistants can be 'emotional', as different voice assistants can give you different answers to the same question. Ask Google Home *"What temperature should chicken be cooked to?"* and you get a temperature in Fahrenheit. Ask Siri and she offers you search results to browse on a screen. Ask Alexa and she doesn't understand the question. Well, I guess she's too distracted by turning down proposals from amorous early adopters.

THE FUTURE OF VOICE?

The most interesting thing to consider about the digital voice evolution is the wide range of unexplored use cases still waiting to be discovered, both at home and in the office. Alexa for Business, for example, will let companies build out their own skills and integrations for both practical and business use cases. Soon enough, Alexa will control office lights and room temperature. It will also manage company calendars, reserve conference rooms and even arrange travel itineraries. As voice interfaces become increasingly important, the popularity of Alexa in the workplace is a matter of when, not if.
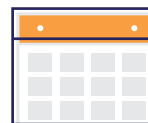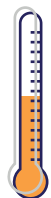
"The potential applications at work are huge, and there's still a lot of unexplored home use cases too. Our studies into streaming content for one client revealed their customers often watch TV on mobile tablets while they're in the bath. But touchscreens and wet hands? Not ideal. Accidentally splashing your new iPad? No way. We should expect to see voice for bathroom, soon."

SIMON HERD
DIRECTOR OF DESIGN RESEARCH, SUTHERLAND LABS

With today's rapidly changing landscape, it's hard to predict what the limitations for digital voice technology will look like. One thing that won't change, however, is the importance of continued ethnographic research into the way we interact and engage with digital assistants. Why? Because the better we understand the pain points of our digital assistant relationships, the closer we are to addressing a flaw in human biology — having too much to do at once with only one pair of hands to help.

In Part 3 of the Voice is Back series - 'Customer Support Is In the Air' — we consider why and how brands should address incorporating the digital voice channel for its customer support needs.

Stay ahead of your competition and evolve with your customers as digital voice assistants evolve into customer support agents sitting right next to your customer.

It's time to talk.

Email us at sales@sutherlandglobal.com, or call 1-800-388-4557 ext. 6123.

ANDREW B. ZIMMERMAN

President, Sutherland Digital, Technology, and Media & Communications

Andy is President, Sutherland Digital responsible for digital practices and technologies including design services, digital management consulting, analytics and digital platform offerings. Prior to joining Sutherland, he was CEO of Frog Design, Global leader of Accenture Mobility and Accenture New Services, and Managing Director of Idealab, a leading incubator.

OWEN DALY-JONES

Senior Vice President & Global Head Sutherland Labs

Owen is a trained psychologist and anthropologist, interaction designer, product designer, and Fellow of the UK's Royal Society of Arts. He founded and leads Sutherland Labs, the creative hub from which Sutherland's Design Thinking approach is driven, and inspires and leads our global design and creative teams. His endless fascination with human behavior and 25 years in the industry make him ideally placed to advise our clients on everything from service propositions to digital transformation.

As a process transformation company, Sutherland rethinks and rebuilds processes for the digital age by combining the speed and insight of design thinking with the scale and accuracy of data analytics. We have been helping customers across industries from financial services to healthcare, achieve greater agility through transformed and automated customer experiences for over 30 years. Headquartered in Rochester, N.Y., Sutherland employs thousands of professionals spanning 20 countries around the world.