

Great Voice Experiences Start with Listening:

Best Practices in Research and
Design for Voice User Interfaces

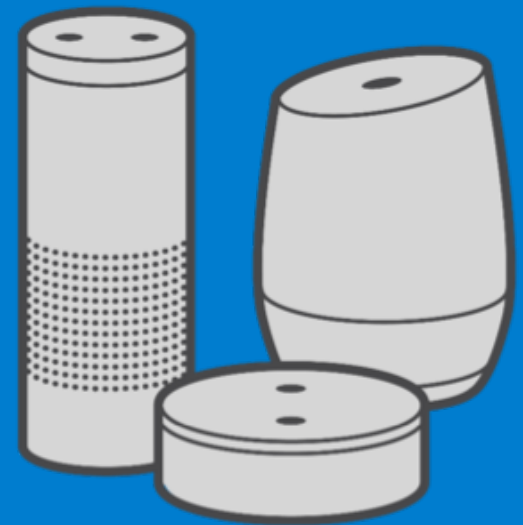


Table of Contents

INTRO

Pages 3-4

- Rise of the smart speakers
- Key findings

The Current Landscape

01

Pages 5-13

- State of smart speaker devices
- Current state of smart speaker applications
- What does this mean for brands?
- Getting found

Best practices for designing voice interactions

02

Page 14-26

1. Fail gracefully
2. Do your homework
3. Be a good host (i.e., setting expectations)
4. Design for context and continuity
5. Use natural commands
6. Make the user feel heard (i.e., confirmations)
7. Have some personality

What's next for VUIs & smart speakers?

03

Page 27-29

- What's next for VUIs & smart speakers

Appendix

04

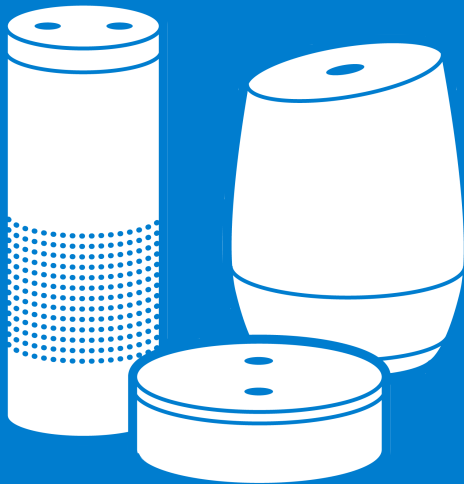
Page 30-35

- About the research
- Smart speaker research team

Rise of the smart speakers

“Human-Computer Interaction takes big leaps every once in a while. The next generation of speech interfaces is one of those leaps.”

–Tim O’Reilly



Whether you call them “smart speakers,” “voice-enabled speakers,” or “voice assistants,” there’s no denying we’re witnessing the birth of a major new technology category. Although voice technology has been around since the 1950s, Amazon introduced the first voice home assistant, Echo, to the US in 2015 with Google following suit in late 2016. While these devices currently fall short of our dreams of our own JARVIS (personal assistant to Marvel’s Ironman), it’s estimated that over 35M Americans already use a smart speaker at least 1x/month—a 129% increase over 2016*. The global market is expected to grow to \$2B by 2020.** In just the past few months, Amazon has introduced a slew of exciting new Alexa-powered technology. Google introduced Google Home Mini and Google Home Max. Apple announced their first smart speaker, HomePod. Sonos and Harman Kardon have been introducing speakers powered by Alexa, Google Assistant, and Microsoft’s Cortana. The blogs swirl with rumors of numerous other companies preparing to enter the fray, including Samsung, and Facebook.

The cause of this flurry of activity? Voice-enabled technology opens up new possibilities for users. Finally we will be able to interact with technology using the most natural interface of all: speech. No longer tethering us to screens, voice interfaces allow technology to become both more present and less intrusive. And yet, internet jokes abound about smart speaker failures. If you’ve spent time living with a voice assistant, you know first-hand they have their limitations. We may daydream sci-fi visions of conversational interfaces, but we live in a world where voice technology is still struggling to deliver satisfying command-and-control transactions.

So what’s going on with smart speakers, what’s working well, what needs improvement, what does the future hold, and what should brands consider when surveying this new landscape? At AnswerLab, questions like this are our reason for being. To answer them for ourselves and our clients, we conducted multi-method research with smart speaker owners to understand not just what they want, but how speaker owners currently behave and why.

*Alexa, Say What?! Voice-Enabled Speaker Usage to Grow Nearly 130% This Year." EMarketer. May 08, 2017. Accessed October 02, 2017. <https://www.emarketer.com/Article/Alexa-Say-What-Voice-Enabled-Speaker-Usage-Grow-Nearly-130-This-Year/1015812>

**Gartner Says Worldwide Spending on VPA-Enabled Wireless Speakers Will Top \$2 Billion by 2020." Gartner. Accessed October 02, 2017. <http://www.gartner.com/newsroom/id/3464317>

KEY FINDINGS

We surveyed 1000 smart speaker owners and spent time in the homes of 10 owners to explore smart speaker use cases, opportunities for improvements, and design best practices.

As one participant succinctly put it: *“If it doesn’t make it easier, we’re not going to talk to this thing just because we can. If it’s more complicated than a website or app, then forget it.”*

- In general, most **participants were very satisfied with their devices** and use them frequently.
- **Participants have fairly low expectations of the current technology:** participants were happy with their smart speakers, but:
 - Over **70% have experienced problems or frustrations.**
 - **25% do not think the designers of these experience thought about the needs of people like them.**
- **Third-party voice applications are confusing users:**
 - Many participants were unclear on the distinction between voice applications and native functionality.
 - **20% aren’t using voice applications.**
 - Participants don’t have one place they go for information on voice applications.
- **Brands need to step up their game.** As with all early-stage technologies, users will only blame the technology or themselves before they start to blame the app developers for a limited time.
- **Security is a concern; privacy, less so.** Most smart speaker owners in our study did not have strong privacy concerns about the 'always on' nature of the devices, but many did have security concerns regarding unauthorized access to information they wanted to keep secure.
- **Design best practices do exist.** Best practices in designing for voice interfaces will continue to evolve, but clear lessons are emerging and following them will help brands ensure they get the most value from their investments in voice interactions.
- The most important tool available to digital leaders and practitioners--both to avoid costly mistakes and to create compelling experiences--is user research.
Great voice experiences start with listening.

The Current Landscape

*State of Smart Speaker
Devices*

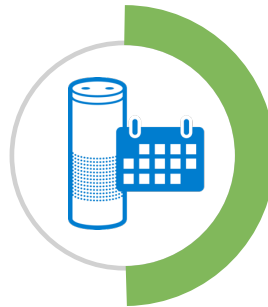
*State of Smart Speaker
Applications*

*What does this mean
for brands?*

State of Smart Speaker Devices

We discovered that over one-third of those surveyed received their smart speaker as a gift. Why is this important? Since this group is less invested in their smart speaker (literally), they are also less invested in learning the technology. They scored lower in measures of use, knowledge, and satisfaction. This makes it even more important that those designing for voice do so with less technically sophisticated/more passive users in mind.

Smart speaker ownership is growing exponentially.



Almost **50%** of those surveyed began using a smart speaker within the past 3 months



63% plan to purchase another smart speaker



77% have suggested buying a smart speaker to people they know



69% use their smart speaker at least once per day



50% say they use it more now than they did during their first month of ownership

State of Smart Speaker Devices

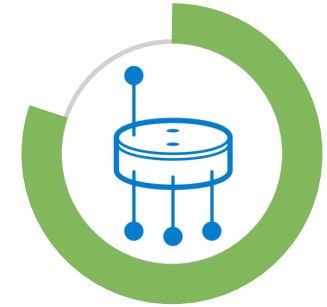
Smart speaker owners love their devices.



86% are satisfied or very satisfied with their smart speaker



82% agree their smart speaker is easy to use



80% agree that their smart speaker has the features they need

Usability scores are high, but 1 in 4 people felt left out or frustrated to some degree.



25% don't think it was created by someone who thought about the needs of people like them



26% have technical issues

State of Smart Speaker Devices

These high satisfaction scores coupled with evidence of frustration or technical challenges support what we heard in interviews: current expectations of the technology are fairly low. Participants we interviewed often shrugged off problems as the sort of inconveniences one puts up with in new technology. In fact, 28% of the smart speaker owners we interviewed who purchased their device themselves did so to experiment with new technology. This was the second most popular reason for getting a smart speaker, after entertainment purposes (30%). Expectations may be low, but we don't expect them to remain that way for long.

Much has been made of privacy and security concerns with smart speakers. Indeed, this may well keep some from purchasing these devices in the first place. But among the smart speaker owners we studied, most did not have strong privacy concerns about the 'always on' nature of the devices. In interviews, some described it as “weird” or “a little creepy,” but they followed this assessment by saying that the convenience afforded by the technology outweighed any concerns they might have. But while most did not have privacy concerns, many did have security concerns. When surveyed about problems or concerns, only 13% included concerns about privacy, but 34% did not agree with the statement “I feel my smart speaker interactions are secure.” In interviews, participants often expressed reluctance to link their smart speaker to financial accounts or other personal data points because they were worried about unauthorized access to information they wanted to keep secure.



State of Smart Speaker Applications

While Amazon calls their third-party integrations “skills” and Google calls them “actions,” we’re going to refer to this category as “applications.”

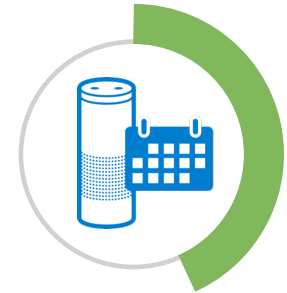
Roughly 20% of those surveyed had never heard of skills or actions and another 30% had heard of them but were unclear on what they were.

Most smart speaker users don’t think in terms of skills, actions, or applications—they think in terms of what they want to do.

Once the concept of third-party integrations was explained, over 80% of those surveyed said they did in fact use voice applications. The remaining 20% use only the out-of-the-box functionality that is a part of each device’s operating system. To better understand the current state of voice applications, we looked at how users discovered them, how they learned about them, and how they used them.



72% said they were likely to add a new application in the next 3 months



42% of application users said they seek out new applications more now compared to during their first month of ownership, indicating a desire to expand their voice assistant’s functionality

Current State of Smart Speaker Applications

“It’s the same as picking up my laptop and going to their website. If the website’s crummy, I’m going to blame the company, I’m not going to blame my laptop.”

—David, participant

One statistic that really grabbed our attention is 63% of those who added voice applications said they had encountered some form of problem or frustration with the applications they use. The most common frustrations with applications:



22% said it’s difficult remembering the exact wording for commands



19% said the skill often does not understand them



19% said the initial setup of skills are difficult

We’ll get into the specific challenges users face in ‘Best Practices for Designing Voice User Interfaces,’ but for now, suffice it to say that brands who fail to deliver a satisfying experience at each touchpoint will suffer damage to their reputation across all touchpoints.

Again, smart speaker owners are happy with the experience, but encounter challenges frequently. Currently, the top two reasons people assume they experience problems with the technology are their own mistakes (37%) and limitations of the current technology (27%). But this won’t last. As technologies mature and design conventions become established, users are less likely to blame the technology or themselves and more likely to blame the app developers.

What does this mean for brands?

How brands should approach developing voice applications

“If it doesn’t make it easier, we’re not going to talk to this thing just because we can. If it’s more complicated than a website or app, then forget it.”

-Jenna, participant

Our research makes clear that customers and prospects expect big brands to have a presence in smart speakers. Customers want to interact with brands in ways that are relevant to them and that leverage the unique opportunities of the platform. However, brands should first ask themselves whether voice is the right interaction for the task.

New communication technology rarely replaces previous technologies. At least not initially. Rather, it adds a layer on top of existing technologies, filling in the gaps and extending their reach. Voice interfaces can’t do some things as well as graphic interfaces and things they do better.

Where voice succeeds:

- **Voice is great** for simplifying complicated requests. Tasks that require taking out one’s phone, opening an app, and going through multiple steps of simple input in a graphic interface can sometimes be completed with a single sentence in voice. Consider, for example, setting a timer or playing a specific album.
- **Voice is great** for hands free contexts. Voice opens up new opportunities for interacting with users where they are—cooking, driving, doing home maintenance, caring for a baby, gardening, etc.
- **Voice is great** for bringing people together. We heard numerous stories of families gathering around their smart speaker to play games, hear jokes, and play music. While many tech devices seem to isolate us further from one another, smart speaker applications have the ability to engage with a group.

Where voice fails:

- **Voice is not good** for tasks requiring complex outputs. Consider what it would be like to use voice for comparison shopping on an ecommerce website, scanning industry news to determine what of it is relevant to you, or reviewing charts and graphs intended to support decision-making.
- **Voice is not good** for tasks requiring complex inputs. Long forms and/or inputs with multiple variables can quickly become overwhelming. Imagine completing your tax forms using voice alone. Not fun.
- **Voice is not good** for situations where auditory privacy is necessary. Participants in our study said they were not comfortable discussing most financial matters, health concerns, or personal details in shared spaces. And many contexts (e.g., the office) are not places where we can speak freely.

What does this mean for brands?

How brands should approach developing voice applications



AnswerLab recommends:

- Voice applications are another touchpoint—ensure your brand’s experience is consistent across devices and platforms.
- Do not simply “port” your web or app experience to smart speakers. There is no need to try and provide all the same functionality in a voice application that you do through other interfaces.
- Consider your web, app, or product’s functionality through the lens of what voice does well and what it does not do well.
- Conduct research to learn what users want to do with voice and build for that.
- Remember: if the smart speaker experience doesn’t reduce friction as compared to existing methods, users are unlikely to use it.

Getting found

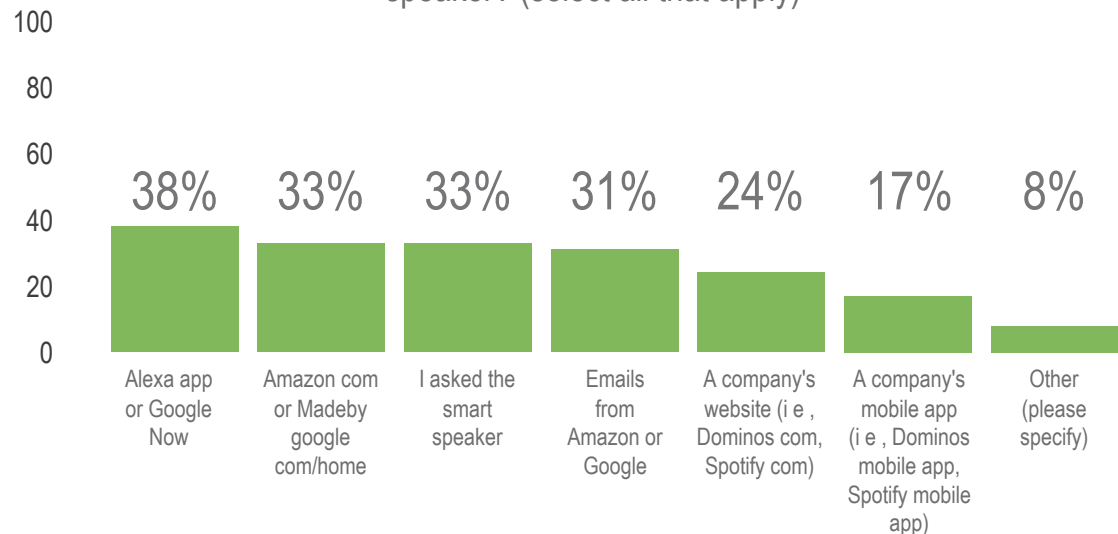
How do application users learn about those applications?

Getting your voice application found is currently a challenge. Many of our survey respondents were unclear on the distinction between native applications and third-party integrations. In our in-home interviews, even those who did understand the distinction sometimes struggled to successfully launch the third-party voice application. Most, when trying to order through the Domino's Pizza voice app, were instead given a list of nearby restaurants.

Survey respondents who had enabled third-party applications found them through a variety of means [chart below]. Participants found no single source for learning about and finding applications. Further, the traditional model of an “app store” doesn't work well over voice. As mentioned previously, it is not easy for users to scan large or complex sets of information in voice. And as we discovered in our research, many users don't frequent the smart speakers' mobile apps.

In short, you can't rely on device manufacturers to get your voice application noticed. If you've committed the resources to developing an application for this platform, make sure you've committed resources to marketing it effectively.

Question: Where did you go to determine if a skill/action existed for your smart speaker? (select all that apply)



Total sample; Unweighted; base n = 813

Best practices for designing voice interactions

1. *Fail gracefully*

2. *Do your homework*

3. *Be a good host
(i.e., setting
expectations)*

4. *Design for context
and continuity*

5. *Use natural
commands*

6. *Make the user feel
heard (i.e.,
confirmations)*

7. *Have some
personality*

Best practices for designing voice interactions

In addition to our research on users' wants, needs, and behaviors around voice technology in general, we also surveyed them about specific challenges and conducted in-home usability tests. In this section, we provide an overview of the things every brand needs to consider when approaching a VUI/smart speaker project.*



*For a detailed guide, we recommend Cathy Pearl's [Designing Voice User Interfaces](#) (O'Reilly Publishing, 2016).

1. Fail gracefully



“To err is human...” and we can expect no more from technology designed and built by us. The key is to fail gracefully. Users will forgive many technical limitations and errors if the system responds in a way that helps them to understand what happened and what to do next. We start with failing gracefully for two reasons: First, the way any digital technology handles fail states is critical to users’ perception of the experience—and this is especially true with the intimacy of voice. Second, virtually all the remaining guidance in this section is built around avoiding these very errors.

Branded Interaction on Device: Would you like to track your order or place a new order?

Alysha: Yes.

Device: I’m sorry, can you please repeat what you said?

Alysha: Yes.

Device: I’m sorry, I didn’t get that. Can you please repeat what you said?

Alysha: [shrugs] No?

Device: I’m sorry, I didn’t get that. Can you please repeat what you said?

Alysha: [quietly to interviewer] I’m irritated now. I don’t even remember the question.

[loudly to device] Yes.

Device: I’m sorry, I didn’t get that. Can you please repeat what you said?

Alysha: OK, Google, stop.

In the example above, Alysha didn’t realize she was being asked whether she wanted to track an order or place a new order, she thought she was being asked if she was interested in either. This specific error could have been addressed with how the query was written (e.g., “Which would you like to do, track an order or place a new order?”). Alysha’s frustration could have been ameliorated by repeating the question after the second error instead of expecting her to remember a query she was clearly struggling to answer. Further, the experience would have been less frustrating, and more human, by varying the error response (e.g., “Forgive me, I don’t understand,” or “Can you say that again?”). Ultimately however, no matter how carefully scripts are drafted and tested, things can—and will—go wrong.

1. Fail gracefully

Never suggest the user is at fault. Apologize and try to help.

When errors occur:

Look for opportunities to not speak.

A soft chime and/or change of indicator light can tell users they weren't understood without the system telling them so. This is gentler, takes less time, and avoids too many "I'm sorry"s.

Be helpful.

- Be specific about what went wrong and what the user can do to resolve it.
- Ask clarifying questions. "I think you said 'X,' is that right?," "Did you mean...?"
- Offer suggestions. "I can't find X, would you like information on Y?" "I'm not able to do that, but there may be an application that can. Would you like to search for voice applications that can help with that?"

Be humble. Never suggest the user is at fault. Apologize and try to help.

Be human. Consider using light-hearted messages where appropriate (as noted in best practice #7, humor can lighten the mood when errors occur, but proceed with caution. Humor should only be used when you know the user is likely to be in a low stress situation and engaged in a low risk task.)

Finally, if there is an error loop occurring (as with Alysha's interaction above), don't keep repeating it. After several failed attempts, offer a friendly apology and release the user from a frustrating back and forth. For example, "I'm sorry for my limited intelligence. Smart people at [brand] are working to make me better every day though! In the meantime, you may want to go to their website or mobile app."

2. Do your homework

“This doesn’t make any sense. They should’ve talked to me first. I could’ve told them this wasn’t going to work.”

—Amanda, participant

The most effective approach to avoiding errors, both in execution and strategy, is to do your homework. One-quarter of our survey respondents didn’t agree that smart speaker interactions were created by someone who thought about their needs. This was even more pronounced among respondents who were older, less tech-savvy, or women. And the best way to know about someone’s needs is to observe and listen.

Far too often, user research is confused with usability testing and only comes at the end of the development cycle. But how do you know you’re solving the right problem in the first place? How do you know that your idea for addressing that problem is a good one, that you’ve anticipated all the use cases, contexts, and things that could go wrong? How do you know if you’re on the right path—or the wrong path? By the time you’re usability testing, it’s too late (and too expensive) to ask these questions.

For additional advice on conducting voice interaction research:

Download our [Best Practices in Voice Interface Research](#) which includes:

- *An Iterative approach to designing for voice*, where we identify and explain what research methodologies we recommend at each stage of the design process.
- *Recommendations for successful voice interface research*, helpful tips for researchers who are new to this platform.

3. Be a good host (i.e., setting expectations)



Device: Welcome to Twitter Reader. I can tell you what is happening, can read your latest mentions, the last tweets you liked, or were posted by you. I can also search for trends by city. When I am reading out a list of tweets or trends, you can ask me to tell you more about the tweet or trend by number. Now, how can I help you?

Heather: Wait, what can you do?

Device: Here are the latest tweets from your timeline...

It helps to think of voice applications a little like a host would: welcome people, make them feel comfortable, and get them where they're trying to go as elegantly as possible. In designing tech experiences, making users feel comfortable should include many of the same practices that go into web and software design: consistency, clarity, and comprehensibility; users should know where they are and what they can do; they should have as much information as they need and only as much information as they need, and they should have it when they need it.

3. Be a good host (i.e., setting expectations)

“If I have to go online to learn how to use [a voice application], I’m already online—I’ll just use that to do what I wanted.”

-Bilee, participant

That said, while it’s important to be unambiguous about what users can do, don’t be too prescriptive in explaining how to accomplish these tasks. Instead of listing everything they can say, give some suggestions and design for all likely variations.

Also, avoid “onboarding.” Users just want to do what they came to do. This is even more important in voice applications, because the ability to “x out” of an onboarding experience may not exist or may not be obvious. Instead, welcome users quickly, get them into their task efficiently, and ensure contextual help is always available should they have questions.

To ensure your users feel comfortable:

- Set expectations at every decision point. Namely, what users can do and suggestions for things they can say.
- Look for opportunities to provide context-aware self-help within the voice application. Where possible, avoid sending users off to look at the device app or your own website for help.
- In multi-step processes, let the user know where they are in the process.
- Follow our advice for writing effective commands, including conducting research to identify all the various ways users might choose to issue commands.

4. Design for context and continuity

Chris: Alexa, what's the yellow light mean?

Echo: The yellow light means you have a new message or a notification. You can say 'play my messages' or 'read my notifications.'

Chris: Alexa, play my messages.

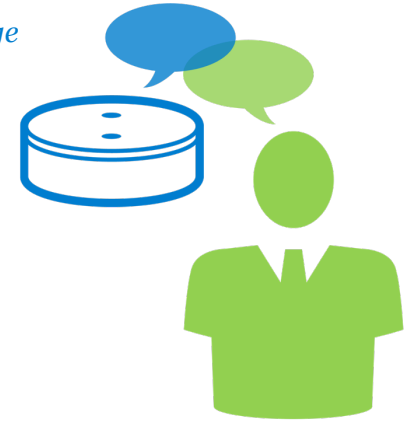
Echo: No messages from today. You have one notification. You can say 'read my notifications.'

Chris: Read my notifications.

Echo: [silent]

Chris: Alexa, read my notifications.

Echo: [reads notifications]



In this case, Alexa does a great job of being a good host and using natural commands. The user didn't even have to think about what to do when he saw the yellow light, he just asked. The problem is Alexa's failure to maintain continuity. Instead of, "You have one notification. You can say 'read my notifications'," why not simply ask, "would you like to hear it?" and allow the user to say yes or no?

Examples like this abound in our research. Voice technology is still in its early stages. At the time of our research, one couldn't ask Alexa the address of a location and then ask "how far is that from home?" Technical limitations exist, but nevertheless, as much as the technology will allow, good voice interaction design pays attention to context, taking into account where the user is likely to be, both physically and in the process of task completion. Further, good design is consistent. Users should not have to wonder if different commands and responses will produce different results.

To design for context and consistency:

- Look at every transition in your task flows to see what can be combined, simplified, or removed. Don't assume the same flow for web or software should be used for voice.
- If there is a next logical step to the process, anticipate that step, ask the user if they would like to take that step, and listen.
- Observe the physical contexts in which your voice application is likely to be used and design for those.
- Review all your commands, responses, and confirmations to ensure consistency among words and actions.

5. Use natural commands

“That’s the other problem I have. Sometimes even the ones that I like, I can’t remember the commands to get them to work, and it’s [sigh], oh god, now I have to go look it up.”

—David, participant

Many participants in our research said they had a hard time remembering the right commands. In traditional web design, good design focuses on recognition over recall, but with voice interactions, this gets flipped. Since there’s no visual stimuli to drive recognition, the user is forced to recall important aspects of the interaction. The less we ask of users in this regard, the better the experience. Users shouldn’t have to learn a new language to interact with your voice application. In this way voice interfaces can add to the user’s cognitive load instead of making her life easier.

To make things easier for users:

- Commands should be natural and easy to remember.
- Commands should be consistent across applications.
- Allow multiple commands for the same action.
- Commands should be sufficiently unique to help with recall and to avoid errors of misinterpretation.

Spend time listening to people talk to learn how they think and speak about the tasks you’re trying to enable. Depending on where you are in the development process and the investment you’re able to make, you might do this through a “wants & needs” focus group or an open card sort, or by simply observing people in their natural environment like their homes or offices. As much as possible, voice interface commands should match your users’ thought processes and vocabulary. We recommend conducting usability testing with a range of users prior to launch in order to make sure you’ve captured the full range of expressions a person might naturally use when interacting with your voice application.

6. Make the user feel heard (i.e., confirmations)



Device: Do you want to log into your profile for a faster checkout experience?

Brian: No.

Device: Sounds good!

[Brian laughs. To interviewer:] It's just very inhuman. It's like, OK, I'm glad we agree. Just do what I asked, you robot. It's trying too hard.

We all want to be understood, and when we ask for something, we want a response. When talking with one another, we often confirm we heard what someone said not with verbal confirmations but with a change in eye contact or a nod of the head. Smart speakers can't do this.

Confirmations can take many forms and practitioners should consider which confirmation is most suited to the interaction in question. Our overall guidance here is that confirmations should be kept to as few as necessary to reassure the user and as brief as possible in order to keep the conversation moving. (Keep in mind that error messages are also confirmations, even if what they're confirming is something that didn't or can't occur.)

6. Make the user feel heard (i.e., confirmations)

Confirmation types and when to use them:

Explicit confirmation: For example, “I heard you say [X]. Is that correct?” For interactions where a mistake would be significant, it is critical to make sure the user’s command was understood. Placing an order that would result in the user being billed or calling someone from their contacts list are important to get right. This can also be used when the system’s natural-language processing isn’t 100% confident it understood a command.

Implicit confirmation: In implicit confirmations, the question is implied in the response. For example, “What’s my commute look like?” “Traffic is heavy and your commute is estimated to take approximately 42 minutes.”

Sometimes, instead of a confirmation, an acknowledgement will do:

Nonspecific confirmation: For example, “Okay.” A general confirmation is most effective when the command is simple and straightforward and a mistake would not be critical.

Nonverbal confirmation: For example, a device’s lights turn red and it plays a dissonant chime. Although smart speakers have a very limited repertoire of nonverbal communication, they do have lights and chimes. Think of how much R2-D2 and BB-8 conveyed with the same vocabulary! Use a nonverbal confirmation when the fail state is not critical and the acknowledgement may be conveyed unobtrusively.

And sometimes, a confirmation isn’t needed at all:

No confirmation: Occasionally a confirmation isn’t necessary because confirmation is conveyed through other means. For example, when asking a smart speaker to turn on a light or to pause music that’s currently playing.

7. Have some personality

“I love Dom [the voice persona for Domino’s Pizza]. I already have a relationship with Dom [through the app and website] and Dom has not done me wrong. I already have a high opinion of it, so it messing up, I know it’s going to get it right eventually. I’ve never used the Progressive thing so it messing up on the first time makes me think it’s a miss. If it was obviously Flo that would’ve been cute, that would’ve upped my opinion of it from the start. If it had used a voice I recognized.”

—Amanda, participant



Web content and ad content are often referred to as having a “voice” based on the tone of the writing—voice interfaces literally have a voice, and many of the same rules apply. But unlike “voice” used to describe writing, users look to voice interactions to be more entertaining than screen-based interfaces, perhaps to fill in for the lack of visual stimuli, perhaps because an audible voice is so personal that many anthropomorphize the voice-powered assistant. For this reason, so many of our participants talked about the voice interface’s “personality.”

You don’t need to have a recognizable voice persona, as Domino’s does. And if you don’t have one, don’t force it. As technology writer Cennydd Bowles* points out, “Marketers can’t resist an opportunity to force a damn relationship on you. Truth is, I don’t want to talk to most of my products. They’re dumb utilities. Close and forget. I want a spade, not the experience of digging.” But our research showed clearly that users do respond to the personality of voice interfaces, and your voice application will have a personality whether by design or default.

*Bowles, Cennydd. “What happens next with Conversational UIs – Cennydd Bowles – Medium.” Medium. February 19, 2016. Accessed October 07, 2017. <https://medium.com/@cennydd/what-happens-next-with-conversational-u-is-b9e4699541d5>

7. Have some personality

Almost all our participants identified humor as the hallmark of personality. But humor can be tricky for brands. We recommend exploring the use of humor (if it's consistent with your brand), but proceed with caution. As mentioned previously, humor works best when you can safely assume the user is in a low stress situation and engaged in a low risk task. They will be far more receptive to a dash of humor if they're asking about surf conditions or current movies than they will if they're checking a flight's status or tracking an important package. With humor as with error handling, practitioners should conduct ideation exercises around 'what could go wrong' and test designs before sending them into the market.

When designing your voice interactions:

- If you haven't defined your brand personality, do so.
- Listen to how your customers communicate with you and match their tone. For example, are they casual or formal?
- Ensure your voice application's tone, word choice, etc., are in line with your "voice" across other touchpoints.
- Look for ways to differentiate from your competitors' voices.

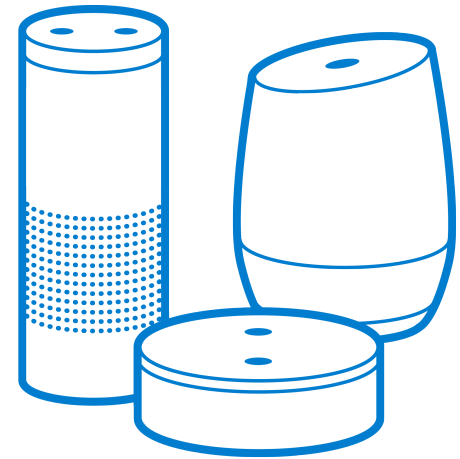
One interviewee, when comparing different voice interfaces she used, said "willingness to help," in addition to humor, defined the personality of her preferred assistant. How do you convey "willingness to help" with a computer program? By using clarifying questions and suggestions instead of error messages. And that brings us full circle to where we started this section on best practices: handling errors with grace.

What's next for Smart Speakers?

We recommend device makers work with strategic insights partners to develop best practices that can be used by those building in this uncharted space. To learn more about research to develop best practices, read how Google partnered with AnswerLab and developed design principles to guide their mobile advertising clients. [Case study](#)

We've entered an exciting time for voice technology. Dramatic advances in natural language processing and machine learning have ushered in the start of true consumer-facing voice interactions. Significant investments by some of tech's most monied companies have propelled explosive growth in the category, both in diversity of offerings and units sold. With over 70% of the market, Amazon is clearly the one to beat. Amazon's product strategy for Echo seems to be to release products almost like betas, trying out new form factors, new interaction models, and new functionalities. But will this serve as a valuable public laboratory or will it undermine confidence in their brand? Perhaps both. Further, when competing against cash- and talent-rich companies like Google and Apple (and maybe Facebook?), first-mover advantage is not a barrier to entry.

Our research provides a snapshot of this changing landscape. We believe many of the sentiments we've uncovered—and certainly the best practices we recommend—will continue to be relevant for some time. Smart digital leaders, product managers, designers, and marketers will certainly want to keep abreast of this new platform as it develops.



What's next for Smart Speakers?

Based on our research, here are some of the developments we're continuing to watch:

Changes to the business ecosystem

Monetization: When Amazon, Google, and/or others provide a path to profitability for developers, we expect the quality of voice applications to improve materially. Will your voice application be ready to compete?

New Entrants: Will Apple's long-anticipated HomePod be able to carve out its own space? Is Facebook really preparing their own entry? Samsung? Microsoft? What will the effect be of these and other entrants into the space?

New functionality

Biometric authentication: Amazon just announced that they can do what Google has been doing, recognize which person is speaking and adjust the experience to that individual. But none of the smart speakers yet have true biometric authentication, and given the security concerns expressed by those we surveyed and interviewed, this will be critical for "v-commerce" and other interactions with sensitive information. Will you be positioned to capitalize on secure transactions in voice?

Visual feedback: Amazon recently launched the Echo Spot and before that the Echo Show, both smart speakers with a screen as part of the interface. Google is rumored to have a similar product in development. Expect to see more experimentation in this space with multi-modal interactions. Will you be prepared for all the usability challenges that come with that?

Always on and everywhere

Always on: Participants in our research voiced few concerns about the technology listening to their conversations. (Where they had concerns was with sensitive information being shared with voice applications.) It was clear from our interviews that convenience was more important to them than privacy concerns. We expect smart speakers will eventually reach a point where they are listening for more than just their "wake word" and they'll remember more about their users. This will open up new possibilities for contextually-aware, continuous interactions. What will this mean for voice applications? Will you be ready?

New contexts: Smart speakers' voice assistants are already making inroads in the automotive space and that will continue. They are already used to manage smart appliances, and we expect they will soon become the hub of the smart home, with microphones throughout the house. The voice assistants in our smart speakers are already on our phones, and while those devices are not yet delivering seamless experiences across devices, they soon will. They are not yet in most computers or mixed reality (VR/AR) devices, but we anticipate they will be. As voice technology expands into new contexts and permeates our lives, those companies that support users throughout their day will be poised to become equally as ubiquitous. Will your company be one of them?

Your customers are talking – make sure you're listening

The landscape will continue to shift as smart speakers gain functionality and conversational interfaces improve. Along with these improvements, your customers' needs and expectations will also increase.

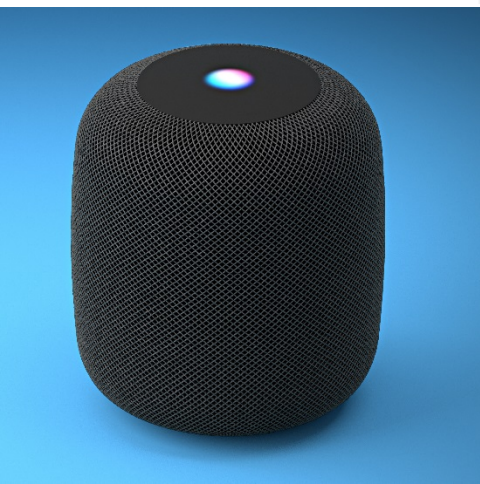
Do you have questions about designing for voice interactions or want help navigating your voice interaction strategy?

HOW TO WORK WITH US

AnswerLab can support your voice experience efforts in the following ways:

- Bring our Smart Speaker team on site for a Q&A as you begin exploring your strategy for designing voice interactions.
- Gain ongoing smart speaker insights—stay tuned for our industry-specific findings and recommendations.
- Engage with AnswerLab workshops to help you define your digital strategy and plan for the user insights you'll need.

Interested in working with us? Contact us at answerlab.com/contact-us



About AnswerLab

AnswerLab delivers insights and advice to create exceptional digital experiences. The world's most innovative brands rely on our research to improve user engagement, reduce development costs, and increase conversion rates. We partner at each stage of the product development cycle, helping digital leaders envision new experiences, optimize existing ones, and measure their impact.

About the research

We reviewed the current state of the smart speaker user experience through a multi-method research study including both qualitative and quantitative methods.

Qualitative Research

We carried out in-depth in-person interviews with 10 smart speaker owners in Sacramento. Why Sacramento? It's one of the 15 most demographically representative cities in the United States. Participants for this segment of our research included people representing:

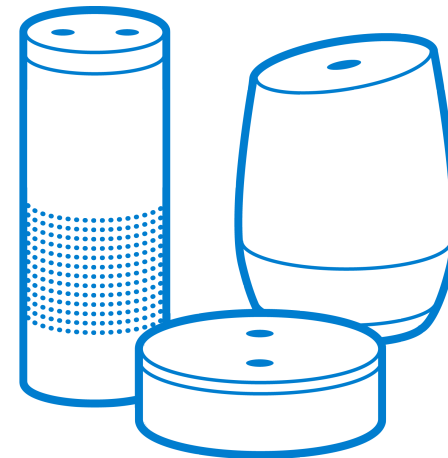
- A range of device ownership including the Google Home device and Amazon's Echo and Echo Dot devices
- Ages 24 to 60
- Mix of household incomes (\$20k - \$200k)
- Mix of education levels (high school graduate or higher)

Participants submitted seven days of diary entries, providing a snapshot of a typical week of activity. Following the diary submissions, in-home interviews allowed us to observe and investigate smart speaker usage in context. The interviews were 90-minute sessions that included both general exploratory questions about participants' current and desired use of smart speakers as well as usability tests of several representative voice applications.

Quantitative Research

We ran an online survey of 1,000 smart speaker owners throughout the United States with a panel provided by the market research company Lucid. Participants for this segment of our research included people representing:

- A range of device ownership including the Google Home device and Amazon's Echo, Dot, Show, and Tap devices
- Ages 21 to 75
- Smart speaker ownership ranging from one month to two years



Smart Speaker Research Team

Chris Geison

UX Researcher,
smart speaker
research and
insights lead

Chris Geison is a UX Researcher at AnswerLab where he leads research to help fortune 500 clients identify and prioritize insights that improve their business results. Drawing from his experience in digital strategy at Charles Schwab and a fascination with motivation and change honed working in behavioral health, Chris has led studies ranging from the role of emerging technologies in the future of banking to the behavioral effectiveness of workspaces, with a particular focus on conversational interfaces.

Ryan Haupt

Principal UX
Researcher,
quantitative
survey lead

For more than a dozen years, Ryan has been advising the world's top digital properties on how to improve their user experience. He leads research across a variety of industries, including telecommunications, financial services, automotive, retail, and design agencies. Ryan specializes in customized, quantitative and behavioral research methodologies across mobile, tablet, ad desktop. Before joining AnswerLab, Ryan was a Senior UX Consultant at Keynote Systems, Inc., a leading provider of UX research software. Early in his career, he worked at Accenture as a developer, team lead, and technology manager for complex backend system integrations.

Smart Speaker Research Team

Amy Buckner Chowdhry AnswerLab CEO

Amy Buckner Chowdhry founded AnswerLab over a decade ago to help the world's leading brands build better digital products. Under her watch, AnswerLab has grown to become a trusted UX insights partner to companies like Google, Facebook, Amazon and more.

Lin Nie UX Researcher

Lin Nie is a UX Researcher with a PhD in Experimental Psychology. Before AnswerLab, she consulted for Amazon's most profitable product, and led foundational user research for startups. Her research in cognitive science and artificial intelligence has appeared in Wired and Slate.

Beth Devine Project Manager

Beth manages research project logistics for Fortune 500 clients representing a variety of industries, including e-commerce, retail, financial and pharmaceutical. With a background as an organizational psychologist, she has a decade of experience in management research and theory. Beth is also skilled in qualitative and quantitative research methods and data analysis.