

MINDSHARE FUTURES



J. WALTER
THOMPSON
INTELLIGENCE

J. WALTER THOMPSON LONDON



SPEAK
THE FUTURE ANSWERS TO YOU
EASY

A consumer trends and insight report on voice technology and its impact on brands.

Opening letter

There has always been something that gets in the way of our relationship with technology; the keyboard, the mouse, the screen. We're now ready for the most natural and intuitive form of interaction—the voice. It's time for humanity and technology to Speak Easy.

Developments in speech recognition and natural language processing (NLP) mean we can now talk to computers in a way that was considered science fiction just a few years ago.

According to Google, 20% of mobile searches on Android are by voice. Amazon has created a new voice hardware category and sold 11 million Echoes, and Google launched its own equivalent, Home, in the UK in April this year.

The rise of a voice-activated world will allow us to get closer and more intimate with technology. Voice interaction will redefine not just how we live our lives, but the digital advertising landscape and how brands reach consumers. Now is the time for marketers to learn how to Speak Easy...

JEREMY POUNDER

Futures Director
Mindshare

ELIZABETH CHERIAN

UK Director
The Innovation Group
J. Walter Thompson

About this report

Speak Easy is a trends and insight report, carried out in equal partnership between J. Walter Thompson Innovation Group London and Mindshare Futures, which explores voice technology and its implications for brands. Our research comprised several methodologies and covered the period January 2017 to March 2017.

NEUROSCIENCE EXPERIMENT

In partnership with Neuro-Insight, we used Steady-State Topography (SST) brain-imaging technology to measure how the brain responds to voice technology for a series of tasks versus text or typing alternatives. 102 smartphone users and Amazon shoppers between the ages of 18 and 65 took part.

QUALITATIVE

Over 30 UK respondents took part in a two-week self-ethnography project capturing their own behaviors and attitudes in a series of voice technology tasks. We then ran two focus groups with 12 of these participants. A mixture of Early Adopters and Early Majority voice users were recruited.

EXPERT INTERVIEWS

We conducted in-depth interviews with experts across sectors including artificial intelligence (AI), neuroscience, marketing, sound design and radio.

QUANTITATIVE

We carried out a quantitative survey using SONAR™, J. Walter Thompson's proprietary market research tool, surveying over 1,000 UK smartphone owners aged 18+ and 100 Amazon Echo owners.

SECONDARY RESEARCH

We carried out extensive desk research that synthesizes international cross-category case studies.

Contents

Voice Today	1
The Future Voice Consumer	19
Brand Futures	29
Implications	40
Thank You	42



Voice Today

Resonantia by Jeff Louviere and Vanessa Brown explores the relationship between photography and music by visualizing 12 musical notes and converting the photographs into audio.

What is voice technology?

We stand on the cusp of a fundamental shift in how we relate to computers. Over the past 50 years, we have been on a journey that has simplified our modes of interaction, from punch cards to keyboard, mouse and then touch. Each step has made engagement with technology simpler and more natural, leading us to perhaps the simplest and most natural interaction mode of all—voice.

Voice is generating excitement because it now works in a way that it did not just a few years ago. Advances in speech recognition (identifying the words spoken and converting them to text) and natural language processing or NLP (understanding the intent behind the words) have made voice a viable consumer proposition. Speech recognition error rates are now at human parity at 5%. We can expect further improvements as machine learning benefits from both continuing investment from the tech giants and the datasets produced by widespread consumer adoption. As a result, the virtual assistant market is set to be worth more than \$3 billion by 2020.¹

“There is a significant difference between [speech recognition of] 95% and 99% accuracy ... it’s the difference between you hardly using it and using it all the time without thinking about it.”

*Andrew Ng,
former chief scientist, Baidu².*

Voice has been given further impetus by the ambitions of GAFA (Google, Amazon, Facebook, Apple) to control their own consumer platforms. Just as Google bought Android so it would have some control of mobile as a platform, so Amazon developed Echo and Alexa to retain influence over the development of voice and its impact on search and shopping.

While development has thus far been driven top down from Silicon Valley, if we are truly about to enter the voice era, then understanding consumer appetite is critical.



“A smartphone is not a neutral platform in the way that the desktop web browser (mostly) was—Apple and Google have control over what is possible on the mobile internet in ways that Microsoft did not over the desktop internet. This makes internet companies nervous ... This is a significant driver behind the Kindle Fire, Alexa, Facebook Messenger bots and all sorts of other projects.”

*Benedict Evans, partner and mobile analyst,
Andreessen Horowitz³*

¹Globe Newswire, “Intelligent Virtual Assistant Market Worth \$3.07Bn by 2020,” December 17, 2015, <http://bit.ly/2nCJLNR>

²Contagious, “Bedtime Story Sounds,” December 13, 2016, <http://bit.ly/2mM1KgJ>

³Benedict Evans “Voice and the Uncanny Valley of AI,” ben-evans.com, March 9, 2017, <http://bit.ly/2maeYU7>

The Landscape

The current voice landscape is dominated by tech giants:

AMAZON

The UK launch of Echo ❶, and the smaller Dot ❷, at the end of September 2016 kickstarted consumer interest by creating a new device category for voice. Powered by voice assistant Alexa, Echo is estimated to have sold more than 11 million units globally and several hundred thousand in the UK to date.⁴ Alexa hosts third-party Skills, which function much like apps but over a voice-user interface (VUI). They can deliver entertainment and information, execute an action, or enable control of other devices. Alexa is also available to hardware developers as Alexa Voice Service to build into their own products in an attempt to stimulate the market.

GOOGLE

Google launched Home ❸, its domestic hardware equivalent to Echo, in the United Kingdom in April 2017. Home allows third parties to create Conversation Actions, which are the equivalent of Amazon's Skills. It will be powered by Assistant, the voice assistant that is currently baked into the Pixel phone and the Allo messaging platform. Over time we can expect it to be built into all Android phones. Irrespective of the assistant interface, voice capabilities have been built into the Google search and YouTube apps for some time now.

APPLE

Apple is a voice pioneer, having launched Siri in 2011, but is increasingly seen as a laggard. While it has integrated Siri into wearables (the Apple Watch and AirPods), its challenge is a comparative lack of machine-learning capabilities and user data to make a voice assistant intelligent. Its strategic focus on data privacy and security reflects this.

OTHER ASSISTANTS

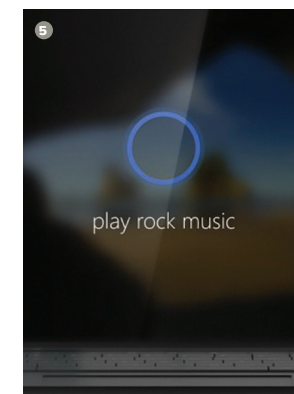
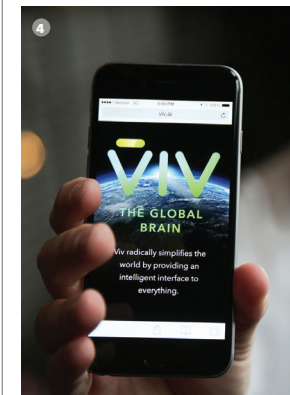
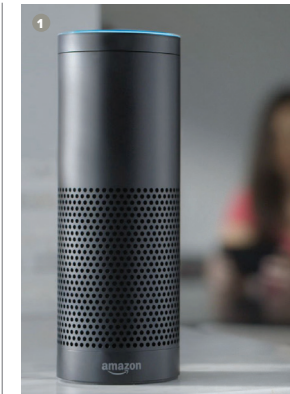
Facebook's assistant M is currently text-based and focused on Messenger, but has the user data and machine-learning skills to make the evolution to voice. Samsung launched its voice interface Bixby for the Galaxy S8 in March 2017, with rumors that it will integrate its acquisition Viv ❹ (the AI built by Siri's original developers), and will look to build voice interfaces into all its consumer products over the coming years. Microsoft has developed Cortana ❺, which works across Windows platforms but currently lacks a strong mobile platform or smart speaker to achieve scale.

BEYOND ASSISTANTS

Going beyond assistance, we are seeing manufacturers such as Mattel's Hello Barbie ❻ or Ford building in voice capability to enhance their products ❼. Voice-activated customer service is also becoming more commonplace. Starbucks is currently trialing a "barista bot" which will allow customers to place orders simply by talking into their phone.

"We're going to see coalescing around a small number of virtual assistants, and it'll just make good business sense to fit within that ecosystem."

*Duncan Anderson,
chief technology officer,
IBM Watson Europe*



"Google search is good, but I feel nervous about leaving it to turn on automatically when I say 'OK Google' because I get nervous that it's listening to me all the time."

JL, focus group participant

⁴Ángel González, "Amazon Has Sold More Than 11 Million Echo Devices, Morgan Stanley Says," Seattle Times, January 19, 2017, <http://bit.ly/2nV4Oab>

Who is using voice and how often?

With limited data available from GAFA (Google has stated that 20% of searches on Android in the United States are by voice,⁵ but has not released a comparable statistic for the United Kingdom to date), we conducted our own primary survey of UK smartphone users to understand voice usage behavior and attitudes.

CONSIDERABLE LEVELS OF VOICE USAGE EXIST ALREADY.

37% of smartphone users use voice technology of some kind at least once a month and 18% use it at least weekly (see Fig 1). There is a sizeable proportion, 23%, who say they have used it once or twice, reflecting some of the teething problems with voice technology, but only 12% say they would never use or consider it. Among early adopters of the Amazon Echo, 60% say “Alexa has become integral to my daily routine.”

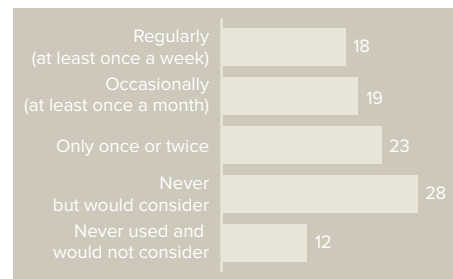


Figure 1
Usage of voice technology
(% smartphone users)

Source: Speak Easy survey Feb 2017; n = 1002 UK smartphone users Q: Would you ever consider using voice technology?

VOICE USERS SHOW STRONG DEMOGRAPHIC SKEWS.

Current voice users are significantly more likely to be young, male and affluent: 50% of 18-34-year-olds, 43% of men and 48% of people with a household income of more than £50,000 use voice monthly or more versus the population average of 37% (see Fig 2).

NO CLEAR DOMINANT VOICE PLATFORM YET.

Google voice search and Siri have the highest usage penetration, with 11% of smartphone users using them at least weekly. Alexa, launched in September 2016, is already not far behind, with 4% of smartphone users using it weekly.

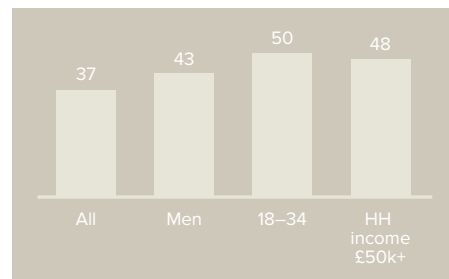


Figure 2
Usage of voice technology by demographic
(% smartphone users)

Source: Speak Easy survey Feb 2017; n = 1002 UK smartphone users Q: Would you ever consider using voice technology?

⁵Jennifer Slegg, “20% of Google’s Mobile Search Queries Are Now Voice Queries,” The SEM Post, May 19, 2016, <http://bit.ly/2n0oAli>

Why do people use voice?

A common motivation for using voice is its novelty, with 45% of regular users saying they do so because “it’s fun.” This could be thanks to the proliferation of Alexa Skills offering bad-taste jokes, pick-up lines, animal facts and the like. But as voice develops we expect the novelty factor to recede.

“Voice is just how we interact with society and the world.”

Joseph Evans, senior research analyst, Enders Analysis

Currently the core reason for using voice is efficiency. The top three drivers cited in our study were “convenient,” “simple to use,” and “faster than typing” (see Fig 3). In fact, humans can speak 150 words per minute, so it’s more than three times faster to shop using a microphone than using apps.⁶ It’s no wonder that 76% of all regular voice users⁷ say that “using voice

technology feels really natural now and I don’t even think about it.”

When it works, voice technology allows people to accomplish tasks quickly and easily, thus streamlining their lives. Duncan Anderson, chief technology officer of IBM Watson Europe, explains, “It’s about being super-helpful, super-efficient, not getting in the way, building something that allows me to get the job done with minimal fuss.” Indeed, 76% of regular Alexa users agree that “Alexa makes my life easier.”

87% of regular voice users agree that “when voice technology works properly, it really simplifies my life.”

But voice promises more than just efficiency. As voice assistants become more intelligent and sophisticated, they will play an advisory role, acting as a “digital butler” for everyday life management.

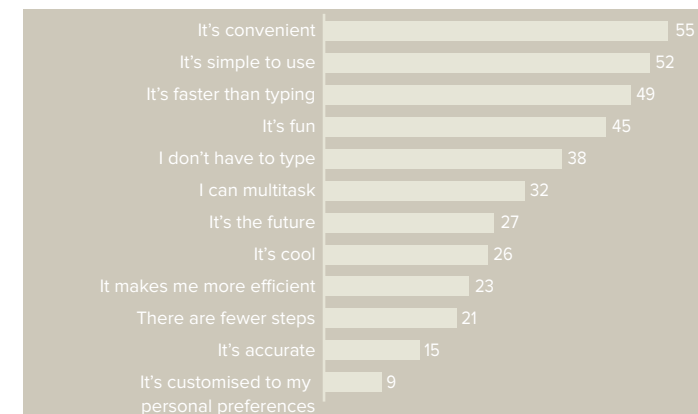


Figure 3
Reasons for using voice
(% voice users)

Source: Speak Easy survey Feb 2017; n = 292 UK regular voice users (at least once per week) Q: Which of the following are reasons that you use voice commands?

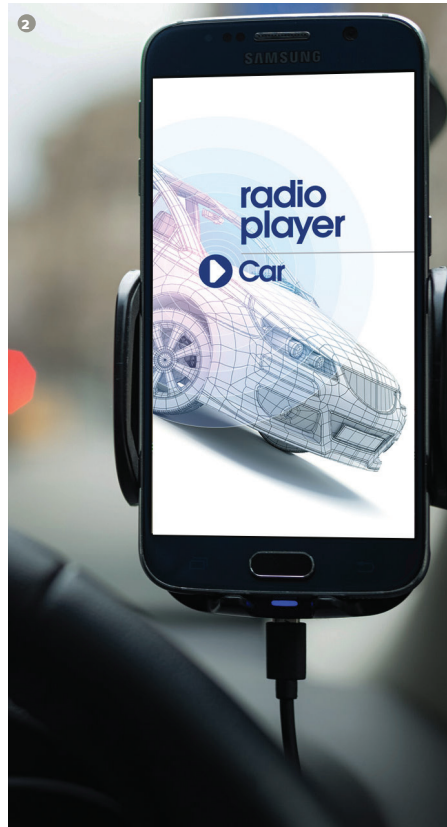
⁶Mary Meeker, “Internet Trends 2016 – Code Conference”, KPCB, June 1, 2016, <http://bit.ly/1otPLCi>
⁷Regular voice users are defined as people who use voice technology once a week or more often

Where do people use voice?

Voice usage is currently skewed towards private spaces, particularly the home where a large proportion of routine tasks, highly suited to voice, take place—hence the development of both Home and Echo. But this also reflects the reluctance users express about using voice in public spaces. This may erode over time as people become desensitized to the experience of using voice in public. Voice-responsive headphones such as Apple's AirPods 1 may also make public voice interactions semi-private by concealing the response from others.

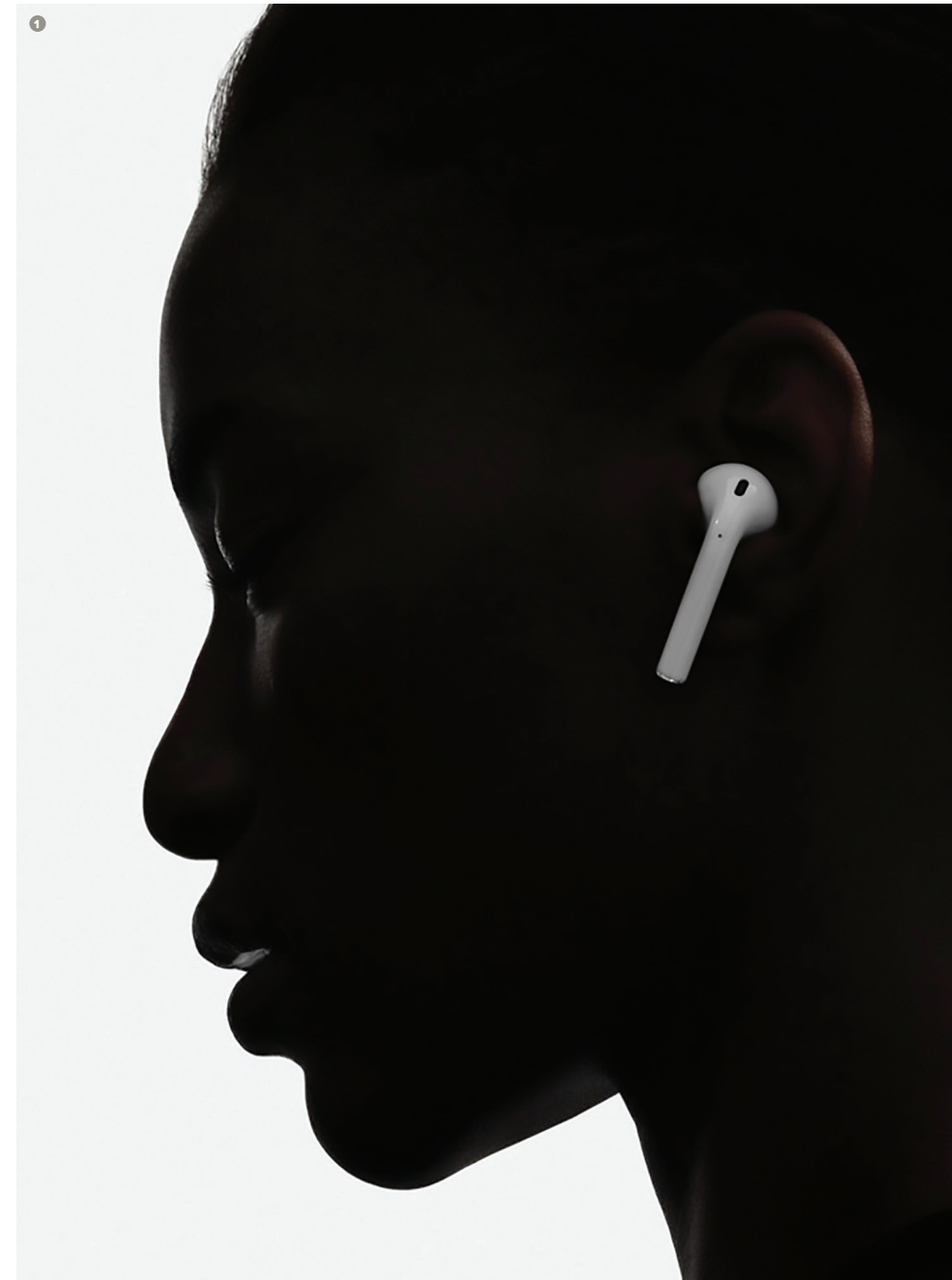
The car is also likely to prove a natural home for voice technology as it is a private space where the benefits of hands-free and multi-tasking are strong. A number of manufacturers, including Volkswagen and Hyundai, are already integrating Alexa into their products.

We expect the home and car to be the dominant locations for voice, with retail and other public spaces to follow in time.



“I think voice control will be crucial in cars of the future.”

Michael Hill, founder and managing director, Radioplayer 2



What are people using it for?

TASKWARRIOR

Early voice usage centers on specific tasks or commands. Aside from the novelty factor of asking fun questions, the majority of use is oriented towards queries with a single, known outcome—play music, find out the weather, give me directions (see Fig 4). Connected home tasks such as switching on lights have very low usage levels with only 15% of regular users saying they have ever used voice for this.

Over time we expect voice to be used for more exploratory interactions, where users are browsing or looking for inspiration, such as researching holiday destinations. The key to moving beyond task execution towards providing suggestions or inspiration will be the assistant's contextual intelligence and ability to refine queries through back-and-forth questioning.

EARLY DAYS FOR SHOPPING

Using voice to buy products is currently in its infancy: only 18% of regular voice users have bought a product purely using voice “without looking on a website first.” Slightly more (24%) have bought a product by voice “having browsed for it elsewhere.” In the case of Alexa, the lack of a screen means that shopping tends to be focused on repeat items, such as groceries, or low-value items where the user is happy for Amazon to choose the product. Amazon's limited penetration in grocery is a factor in the low levels of shopping at this stage.

Consumers will also take time to get used to voice as a shopping interface and trust it to work more efficiently than the touch or typing alternative. Nonetheless, voice is likely to create countless new opportunities as situations where you'd normally be otherwise occupied—cooking, knitting, children's bathtime—become potential shopping occasions.

“I don't trust it that much and I always log in on my phone to check if the right item has been ordered.”

AB, focus group and online community participant

INSPIRING NEW BEHAVIORS

As voice enables tasks to be performed concurrently, we are starting to see home users co-opt new behaviors into their routines. For some, this is as simple as requesting the news headlines (“flash briefing,” as Alexa calls it) while getting ready in the morning or dictating shopping lists while surveying the contents of the fridge.

“Normally while I try to muster the energy to get out of bed, I get Alexa to read me my flash briefing. I don't have a TV and don't really use newspapers, so being able to lie there and slowly catch up on the news in the morning is perfect for me.”

Online respondent

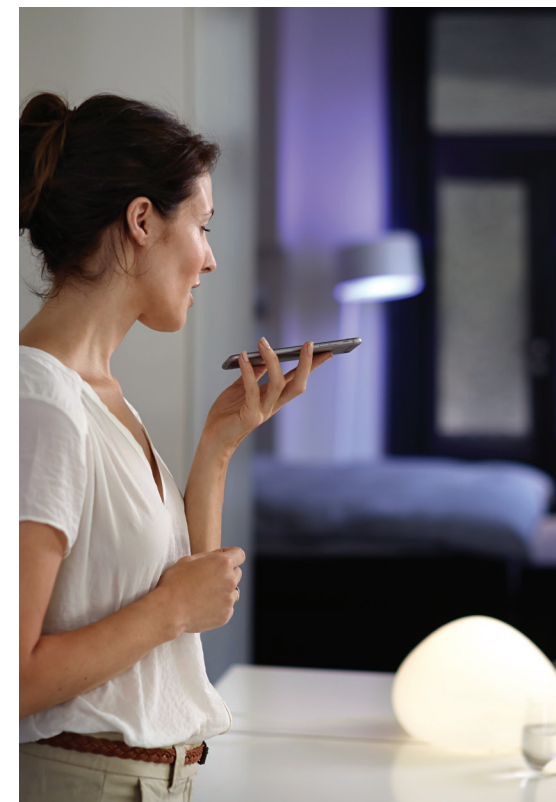


Figure 4
Uses of voice technology (% regular voice users)

Source: *Speak Easy survey Feb 2017; n = 292 UK regular voice users (at least once per week) Q: Which of the following tasks have you used voice services for?*



How is voice used?

Voice usage can be split into two main types. These are tasks carried out entirely by voice, with both the question and response delivered this way (for instance, “Alexa, what’s the weather like today?”), and tasks that are initiated by voice and completed on a screen, typically a mobile (such as a voice search for directions). As voice assistants become more intelligent we expect there to be more opportunity for “100%” voice interactions. In future, booking a restaurant could be achieved through a series of back-and-forth questions to refine the choice. An early example of choice refinement comes via the Johnnie Walker Skill for Alexa, which poses questions about flavor preferences and price to deliver the perfect whisky recommendation for users.

The interplay between voice and screen is likely to remain critical, however, particularly for tasks that depend on a display of visual information such as fashion or travel research. We are starting to see devices that blend voice and a screen, such as Hello Egg ❶, a standalone voice-enabled cooking assistant, and Baidu’s Xiaoyu Zaijia ❷ (meaning Little Fish), which responds using a combination of pictures, text and speech. There will always be activities where voice plays little or no part, at least for the foreseeable future, such as buying the wedding gown of your dreams.



What stops smartphone users from using voice?


Lack of awareness is not a significant barrier, with only 7% of those not using voice citing the reason as “I wasn’t aware of it.” Instead, smartphone users express a reluctance to use voice in public due to embarrassment, with 22% saying “I’d feel stupid using it.” As with Bluetooth headsets and mobile phones, we expect self-consciousness to recede over time. Nonetheless, some tasks (banking being a prime example) will still demand discretion, making them less likely to be carried out through voice—22% say “I wouldn’t want other people to overhear” as a reason for not using voice.

“Perhaps as we use it more in the home and the car, people will start to get more used to it, and the feeling-daft factor fades away a bit.”

Duncan Anderson, chief technology officer, IBM Watson Europe

A more significant barrier seems to be a belief that voice offers no advantages over the touch or typing equivalent, with 48% of non-voice users saying they “don’t see the point.” Some of the teething problems

with speech-recognition accuracy and the precise language required by some assistants to complete tasks are contributing to this perception. Half of regular Alexa users say “I have had to change my behavior to incorporate Alexa in my life.”

Finally, a substantial minority is concerned by the privacy implications of having “listening” devices in the home: 44% of regular voice users say “I am worried about companies listening to the conversations I have with my voice assistant.” This is already driving the development of gadgets to foster privacy in open spaces, such as United States-based startup HushMe , which is seeking crowdfunding for what it dubs “the world’s first voice mask.”

Yet, expressing concern is not the same as taking evasive action. In our view, the proportion of people that claim to be concerned exceeds those prepared to modify their behavior as a result. Nonetheless, privacy is a key issue that won’t go away. Developers of voice interfaces will need to make it a priority.

“I personally find it quite convenient to always have [Alexa] there listening—although that does throw up some privacy issues for me, it’s something I am willing to overlook.”

RB, focus group and online community participant



What do people want from voice?

How can voice technology evolve to best serve the needs of consumers?

INTEGRATED AND SEAMLESS

Users of voice are looking for greater integration across platforms, devices and contexts, picking up conversations where they left off and building on prior understanding. This is certainly Amazon's goal in licensing Alexa to third-party hardware manufacturers.

“Web browsing isn't going to go away, apps aren't going to go away, ... video on demand isn't going to go away, ... but voice interactions will be one channel. A lot of this stuff will move to being called up by voice.”

Joseph Evans, senior research analyst, Enders Analysis

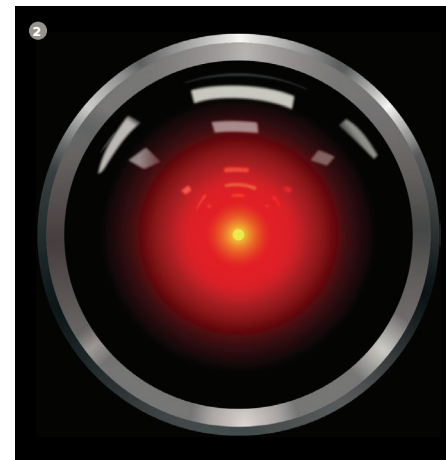
SMARTER SERVICE

Some users are underwhelmed by the current sophistication of voice assistants and feel as though the vision of voice has been oversold. Cultural reference points from HAL 9000 ❶ through Iron Man's JARVIS to Channel Four's *Humans* ❷ inevitably frame expectations. People want voice assistants to show greater understanding, be able to initiate conversations, and preemptively solve problems. Of regular Alexa users, 46% would like back-and-forth dialogue to get to better answers. This is about knowing preferences, remembering past behaviors, and delivering better service as a result.

“I would love the AI to be able to keep track of the topic or subject of the conversation, without always having to reintroduce the subject or topic when asking a new question.”

Online respondent

60% of smartphone users agree that “if voice assistants could understand me properly and speak back to me as well as a human can, I'd use them all the time.”





Hypothetisches Gebilde by Alicja Kwade, 2016,
featured in 303 Gallery at Frieze London 2016.

The Future Voice Consumer

As we look towards the future, a number of trends
will shape how consumers use and feel about voice.

Easing the cognitive load

Voice will be embraced as a less mentally taxing form of interaction

Our research identified efficiency as a primary motivation for using voice. To explore this further, we worked with Neuro-Insight, a leading neuroscience research agency, to investigate the brain's response to voice interactions, compared to touch or typing (see "About this report" for more details).

We found that voice interactions showed consistently lower levels of brain activity than their touch equivalent. Specifically, both left and right brain memory encoding were lower, indicating that the brain was not working as hard to process information (see Fig 5). This was particularly pronounced when respondents were receiving information, suggesting that a voice response is less taxing than its screen-based equivalent (see Fig 6).

This supports the idea that voice, as the oldest form of human communication, is inherently more intuitive and comfortable. Notably, 41% of all regular voice users tell us that they use voice tech when they are feeling lazy. As composer, sound designer, artist and audio specialist Nick Ryan points out, "For most of our cultural evolution as a species, humans have transmitted knowledge and ideas from one generation to another through oral tradition—the voice is therefore perhaps the most innate and intuitive way for us to communicate."

Some companies are already tapping into the human bias towards voice. Travel app HelloGbye allows users to dictate their dates of travel, destination and the number of people they are travelling with to generate a suitable list of flight and hotel options.

We believe that the relative ease of voice interaction will make it increasingly attractive to consumers, particularly those feeling overwhelmed by technology.

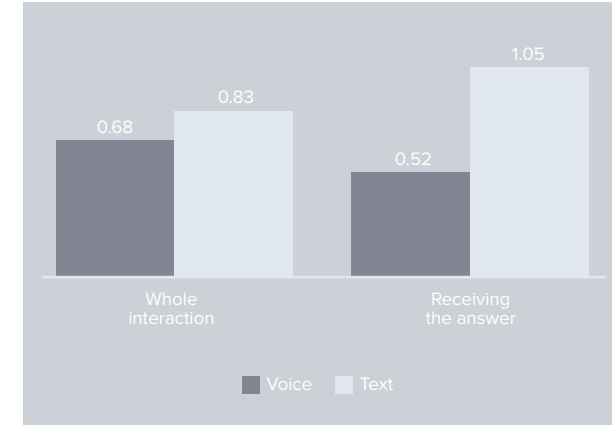


Figure 5
Long-term memory encoding during interaction

Source: Neuro-Insight study Feb 2017; n = 102 UK smartphone users. Brain activity measured using SST headsets; unit of measurement is radians, which equates to strength of brain response

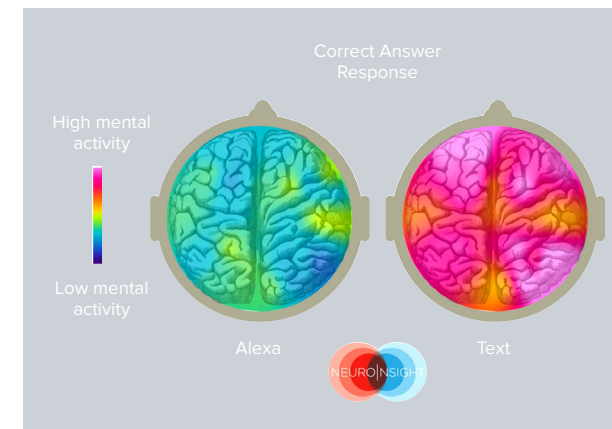


Figure 6
Brain activity maps when receiving a response

Source: Neuro-Insight study Feb 2017; n = 102 UK smartphone users. Brain activity measured using SST headsets; unit of measurement is radians, which equates to strength of brain response; respondents were asked to carry out a series of tasks either speaking to Alexa or typing into an iPad

The Digital Butler

Voice assistants will take on a more proactive role in managing our lives

If voice will be embraced because it is less cognitively draining, it is possible to envisage how this bias towards the “path of least resistance” will extend to our wider decision making. If it becomes easier to speak to a voice assistant than type a search query, it may also prove easier to follow the assistant’s proactive suggestions than think through the alternatives. The voice assistant may evolve to take on a much more prominent decision-making role, managing the user’s life proactively. In effect it would become the “digital butler” or “pocket concierge” helping us at every turn.

Already, people depend on their virtual assistants and other AI helpers to ease the burdens of everyday life. For example, we follow our satnavs when they suggest alternative routes due to traffic. Lily, a conversational personal stylist app, uses AI to understand how a user views their body and recommends clothes based on these perceptions.

Almost a third of respondents are excited about a future where their voice assistants will anticipate what they need and take actions or make suggestions.

58% of regular Alexa users agree that “voice technology will make humans smarter.”

For brands, the opportunity will come in being selected for recommendation ahead of their competitors by the digital butler, a topic explored further in “Brand Futures”.



BRiN uses AI to understand business problems and provide smart solutions through a smartphone app.

Liberation from the screen

Voice will free us up to interact better with the world around us

Voice will enable people to interact more directly with the world around them as their eyes move away from their screens. This presents brands with new opportunities to engage people.

Over the coming years we can expect voice technology to become increasingly ubiquitous in the home. CNET estimates that there were more than 30 significant Alexa hardware integrations announced at CES 2017 alone. Outside the home, voice-enabled headphones akin to the AirPods 1 will drive usage in public spaces. Vinci's Alexa-enabled headphones 2, for instance, allow users to control their music or ask for directions on the go using voice.

As we are freed from the screen we will carry out tasks hands-free, something Duncan Anderson of IBM Watson Europe

sees as a boon: "Voice interfaces have an advantage over text-based interfaces because they don't need the screen."

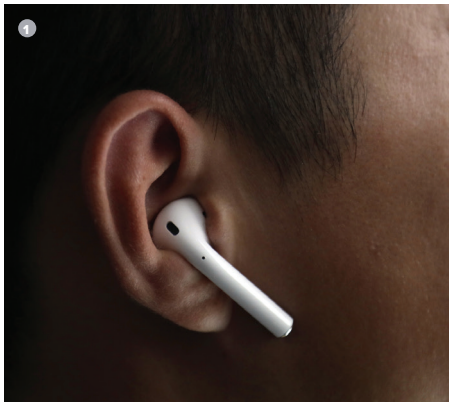
Voice could even power better human interactions. Already 44% of smartphone users think "voice technology will help people interact more with each other as they won't be always looking down at a screen."

That is not to say that liberation from the screen will be absolute. While voice may free up some of our screen time through streamlining tasks, many may choose to redirect their attention back into visual communications and social media. Looking further ahead, though, if mixed reality (MR) becomes mainstream, even referencing a visual won't require users to look down at their smartphones.

But with more capacity to "look up," people may become more open to interaction with brand communications in the world around them, whether in the form of voice-activated outdoor sites or even embedded brand content in internet-connected objects.

For brands, the challenge will be twofold. Firstly, they will need to ensure that services or content can be easily accessed through voice in a simple and intuitive manner. Secondly, they will need to think how they can capture the attention of consumers through content accessible in the Internet of Things.

54% of smartphone users agree that "it would be much easier if technology could speak back to me."



Craving intimacy

Voice users will look for deeper understanding from their voice assistants and develop stronger emotional connections

As advances in machine learning improve voice assistants' capabilities and users feel better understood, we expect people to grow closer to their assistants.

We can already see a consumer appetite for voice assistants to be able to understand them more fully—60% of smartphone users agree that “if voice assistants could understand me properly and speak back to me as well as a human can, I'd use them all the time.”

People want the assistant to know their every preference and deliver an experience that caters to their every whim. As one of our respondents put it, “I'd like voice technology to understand me on the level that humans understand each other.” Around a third (32%) of smartphone users are excited about a future where “my voice assistant will anticipate what I need and take actions or make suggestions.”

For this relationship to flourish, trust is a prerequisite. Providers need to earn trust through a track record of successful service before dependency can grow and a deeper relationship can emerge. They also need to alleviate users' privacy concerns and demonstrate responsibility with personal data. For some, this level of intimacy equates to an infringement of privacy that is unacceptable. However,

we don't believe that this will be enough to dissuade the majority of consumers from engaging. As one focus group participant explained, “You can build trust by, hopefully, making sure no one's ripped off while giving them access to do amazing things.”

Once trust grows, emotional bonds can grow, too. At present, emotional attachment to our voice assistants is limited. Our neuroscience experiments found that the emotional response to voice assistants is considerably lower than for both a face-to-face human interaction and a touch or text interface (see Fig 7). This is likely to be a function of their current rudimentary personalities and a lack of familiarity.

Yet the emotional response to Alexa grew during the course of our experiment as people became more comfortable using it (see Fig 8), pointing to the latent potential for a closer relationship.

This is borne out further by some of the early adopters of voice, for whom a deeper emotional attachment is starting to develop. Over a third (37%) of regular voice technology users say that they love their voice assistant so much that they wish it were a real person. Even more astonishing is that more than a quarter of regular voice

technology users say they have had a sexual fantasy about their voice assistant.

“It is interesting, when something acts naturally and human back to you, how much we imbue it with sentience, with human personality.”

Martin Reddy, cofounder and chief technology officer, PullString

While extreme, this points to an anthropomorphization of the voice assistant that chimes with the 70% of regular voice technology users who say “I want to feel like I'm talking to a real human when I talk to my voice assistant.” Gatebox, a Japanese device akin to the Echo, gives an indication of how this could develop. It contains Hikari, a three-inch holographic girl who can recognize the user's face and voice and is designed to be a home companion.

For brands, the opportunity will be navigating this ever-closer relationship as voice assistants become increasingly powerful gatekeepers to the consumer.

26% of regular voice tech users say they have had a sexual fantasy about their voice assistant.



Figure 7
Emotional response during interaction

Source: Neuro-Insight study Feb 2017; n = 102 UK smartphone users. Brain activity measured using SST headsets; unit of measurement is radians, which equates to positivity of brain response

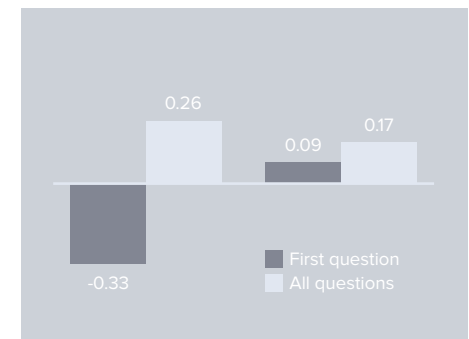


Figure 8
Emotional response by question

Source: Neuro-Insight study Feb 2017; n = 102 UK smartphone users. Brain activity measured using SST headsets; unit of measurement is radians, which equates to positivity of brain response



Brand Futures

As voice becomes an increasingly important mode of interaction, what will it mean for brands? We've identified five key themes that brands will need to consider if they are to thrive in a voice-activated world.

Translation of Sound by Vaclav Mlynar, Kawther Alsaffar, Kristian Knobloch and Ziyang Zhang, 2015, is a Royal College of Art project which transforms the aural into the visible.

Grappling with the voice gatekeepers

Brands will need to reappraise relationships with Google and Amazon as voice strengthens their role as consumer gatekeepers

It seems unlikely that the interruptive model of paid advertising will translate easily to voice. Since a primary expectation for using voice is a more streamlined experience, the insertion of paid-for ads directly into a conversation with an assistant will jar, as Google recently learned. Google Home users kicked up a fuss when Google Assistant reported that “Disney’s live-action *Beauty and the Beast* opens today” after it delivered the time, weather and travel update. Google quickly removed the ad.

“I think the reality is that we’re just not interested in advertising, it’s a very different world. I don’t want to talk to my device, then it giving me an advert for five minutes—that’s just never going to work.”

Duncan Anderson, chief technology officer, IBM Watson Europe


Currently, in the Echo or Home ecosystems, brands can create services (Skills on Echo or Conversation Actions on Home) which are activated by the user. Amazon’s Alexa app displays new Skills to download in much the same fashion as a traditional app store, as well as sending email updates promoting them. While Skills and Conversation Actions offer companies the opportunity to build

a branded experience over a VUI, their future effectiveness is much in doubt for two key reasons.

Firstly, in the short term the technology is not able to deliver a unique voice that’s also interactive. Other than opting to include pre-recorded broadcasts, brands have little scope to differentiate themselves from either their competitors or the virtual assistant itself. Martin Reddy, cofounder and chief technology officer of PullString, explains that “to do that there’s a huge amount of work ... There’s a massive effort to build just the synthetic voices of today, and those tend not to have a lot of emotional range.”

Secondly, like Alexa herself, each Skill requires users to invoke particular commands to make it function properly, which they find cumbersome. As one focus group participant explained, “I found that a lot with the Dot especially, you needed to ask for things in a very formulaic way. And there were really obvious things you couldn’t get it to do unless you asked in very specific way.” There is already a challenge in discovering Skills within the Alexa platform and a tendency for users to abandon Skills in the same way many mobile apps are forgotten over time.

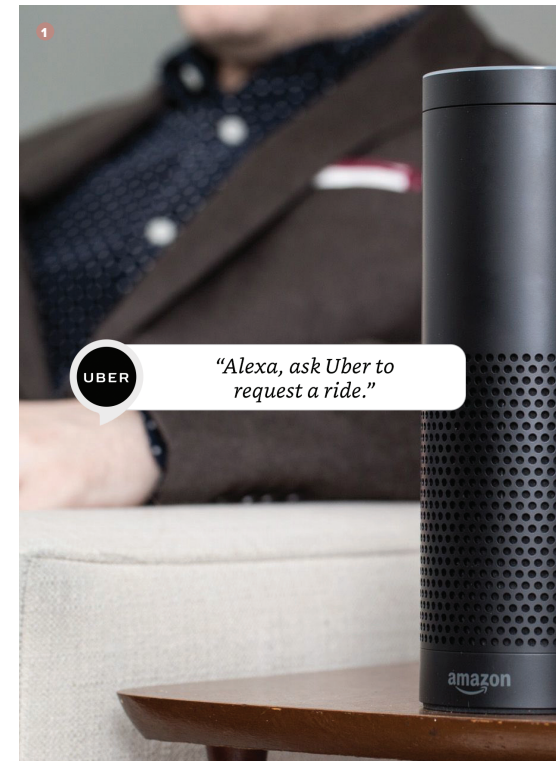
There will be some Skills that stand out, such as the Tide stain-removal one in the

United States, which allows users to get cleaning tips for specific stains, or the Uber Skill , which allows users to order and cancel rides or check on their status. But most people don’t want a particularly branded solution. Joseph Evans of Enders Analysis predicts, “You’ll have a layer of Skills by developers and they will present themselves to the Alexa central level, saying, ‘I can do X, I can do Y, I can do Z.’ And then when you have a user request, Alexa will do some kind of machine learning and say, ‘Okay, what they said means that they want Y,’ right?”

With the interruptive ad model in question, and Skills and Conversation Actions likely to recede into the background of the overall voice-assistant experience, brands should consider the following opportunities if they want a share in the more than \$2 billion of sales that were forecast to be driven by digital assistants in 2016 alone.⁸

“There’s a pretty good argument ... that the people who put these devices in your home seek to own the market because they own the habit.”

Nir Eyal, bestselling author and lecturer at Stanford University




⁸Finbarr Toesland, “Voice Search and Chatbots Are Transforming Commerce”, Raconteur, September 8, 2016, <http://bit.ly/2mlu4tT>

Recommendation and algorithm optimization

A key challenge in a world intermediated by voice assistants will be ensuring your brand or content is chosen by the assistant. Algorithm optimization will become the new SEO

The lack of a visual interface with a VUI makes it hard to deliver more than one outcome as audio lists of results are hard to follow. Voice technology works best when there's only one answer to your question or command according to 83% of regular voice technology users. As the market for voice technology matures, brands are likely to face three key options for ensuring the voice assistant delivers their brand as the single result: paid recommendations, the affiliate model and algorithm optimization.

In some circumstances it may be possible for a voice assistant to offer a clearly

identified "paid recommendation" to a user. For instance, asking Alexa  to buy more washing powder could lead to the (paid) suggestion of an alternative offer. At low frequency, and if the paid recommendation is genuinely of value to the user, paid interventions could be acceptable. When powered by all of the understanding Amazon has of its customers' shopping habits, it is plausible that paid recommendations could be sophisticated enough to be of value. Voice shopping on Alexa may reinforce brand loyalty, but it could also help the retailer in many fast-moving consumer goods (FMCG) categories. The key to this

model's success will be user trust that the suggestion is right for them and not just lucrative for Amazon or Google.

An alternative is an affiliate model, similar to price comparison services today. The voice assistant could deliver a single recommendation to a query and then take a commission on a subsequent lead or sale. Just as finance brands need to pay for a presence on the finance aggregator Moneysupermarket, so all brands may need to pay a commission to be accessible through voice assistants. The paid search costs of today could become the distribution (commission) costs of tomorrow. As of now, this approach does appear to be palatable to consumers, with 63% of regular voice technology users saying, "I don't mind if Google or Amazon take a commission from a purchase made by voice as long as the deal is good for me."

A third option gaining recognition is "assistant optimization," also known as "algorithm optimization." Much as with search engine optimization (SEO), businesses will be able to affect the likelihood of the voice assistant recommending their brand. Understanding the criteria through which assistants will choose recommendations will be vital. In this context, the power of endorsements may grow, and being able to generate content, supported by key experts, media brands or consumer opinion, will be increasingly important.

In a voice-enabled world, ensuring your content is chosen by the assistant will become ever more critical.



"A voice-based digital assistant wouldn't be able to describe a map to you, so things which are visually represented are best seen rather than described in language."

Nick Ryan, composer, sound designer, artist and audio specialist

Join the Internet of Things

As voice capabilities spread to new contexts, brands should consider adding voice interactivity to their touchpoints

Voice technology is moving rapidly beyond mobile phones and smart speakers. Already Alexa has been built into dozens of smart devices.⁹ A key pillar of Amazon's strategy is the licensing of Alexa Voice Service to hardware manufacturers to help drive the voice ecosystem.

There is also a growing appetite for voice capabilities. General-purpose social robots such as Olly **1** or Jibo are an alternative to smart speakers. They behave more like Rosie, the robot in popular cartoon series *The Jetsons*, and attract a devoted audience. Voice-activated products such as the Hello Barbie doll and

Dreamhouse **2** or the kids' digital storybook *The Snow Fox* **3** offer a more interactive and enriching experience than their dumb equivalents. Meanwhile, the number of commercial settings integrating voice-activated services is growing. In hospitality, there is Connie the robot concierge at the Hilton McLean Hotel in Virginia, while teams of robots at the Henn-na hotel in Japan check in guests and offer concierge services.

Brands need to think about how they can make their own physical assets voice enabled, whether that be through using Alexa or their own voice technology. For



⁹Leslie Hook, Richard Waters and Tim Bradshaw, "Amazon Pours Resources into Voice Assistant Alexa," Financial Times, January 17, 2017, <http://on.ft.com/2mZUFd1>



retailers, this could include embedding voice assistants in display units or changing rooms to help customers with a focused set of needs. We found that 38% of smartphone users already want "a voice assistant that could communicate with the store (e.g. to help you navigate while there and help you find products more easily)."

Connected packaging also presents the possibility of being able to talk directly to physical products and access relevant information or content. Almost half (49%) of smartphone users agree with the statement "I like the idea of being able to ask my products questions about their provenance." At CES 2017, we saw the launch of talking packaging in the form of Cambridge Consultants' AudioPack concept for drugs and medical devices.

Media itself is increasingly likely to become voice enabled. Already we are seeing interactive audio or radio ads that allow users to talk to the brand and request further content or even purchase. Instreamatic.AI is an all-in-one interactive audio platform that allows advertisers to create voice-enabled audio ads, which companies as diverse as McDonald's, Jaguar, Sony and Starbucks have already explored. As the TV becomes voice enabled (Alexa is now built into the latest Amazon Fire stick), people may also use

their voice assistant to interact with TV advertisers. Sky has recently integrated voice recognition into its remote. There will also be opportunities to add voice interactivity to one of the oldest media of all, poster sites.

As brands make more of their touchpoints voice enabled, they will start to amass a voice dataset that will present opportunities for voice analytics. What can you learn about your customers from their tone of voice and how they talk to you? Can you apply sentiment analysis to the voice recordings to understand how customers feel?

What's more, brands that build their own voice-enabled products and content can position themselves outside the voice gatekeeper ecosystem and take a step towards winning back control over how they engage with their consumers.

"You have to be able to offer people something which makes it worth their while to actually interact with you ... and those who are able to do that will actually do it well."

*Joseph Evans, senior research analyst,
Enders Analysis*

Are you being served?

Brands that develop truly useful content have a great opportunity to engage in direct communication with their consumers

“The question for brands should be, if tomorrow you had a seamless way to talk to your product ... would that save the customer steps?”

Nir Eyal, bestselling author and lecturer at Stanford University

There are two routes a brand can take when developing a chat experience. One is to build a chatbot for a messaging platform such as Facebook Messenger or Kik. Users who are interacting with a brand’s chatbot spend on average 32% more time on Kik than those who don’t.¹⁰ Text-based messaging will continue to be valued by consumers for its speed and simplicity, but brands should be thinking

about the circumstances in which chatbots could or should be voice-enabled. It is possible to envisage successful text-based chatbots, such as 1-800-Flowers ² or Skyscanner, delivering an even more flexible experience for customers by adding voice capabilities.

In addition to chatting on a messaging platform, people will increasingly expect to carry on their conversations through voice on a VUI. For example, someone might arrange a delivery in Messenger one day and want to check up on it via voice through Alexa the next. Creating symbiotic Skills or Conversation Actions could be one way to address consumers’ desire for a seamless and integrated experience.

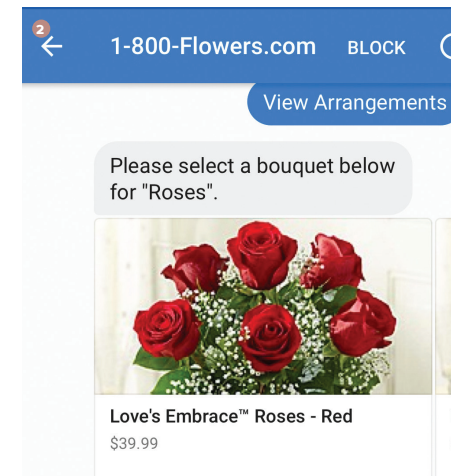


¹⁰BI Intelligence, “Chatbots are thriving on the Kik chat app”, Business Insider UK, August 4, 2016, <http://read.bi/2nCHdiH>

The second route to consider is creating a Skill or Conversation Action. When contemplating this investment, brands must first decide what the service will accomplish. “Amazon made us boil down our original ideas to the simplest possible concept. Which is what you’ve got to do when you’re designing for voice,” says Michael Hill, founder and managing director of Radioplayer, about his experience of designing the Radioplayer Skill.

The characteristics of some of the early examples of successful Skills are rooted in their simplicity and ability to offer real value in a natural, straightforward way. A user who activates the BMW Skill ¹ can

ask Alexa about a scheduled trip, what time to leave, and send the destination to the vehicle. The Just Eat Skill allows users to reorder their last takeaway with one command, significantly reducing the number of steps required to place an order—something likely to appeal to the close to one quarter of regular voice users who cite fewer steps as one of the main reasons for using voice technology. Ultimately, enabling interaction with bots through voice, regardless of location, comes back to the core promise of the technology—understanding the needs of the user and delivering on that need through a process that is simple and intuitive.



“Human language is the new user interface, bots are the new apps.”

*Satya Nadella, CEO, Microsoft*¹¹

¹¹James Titcomb, “Microsoft Build 2016: Windows 10, HoloLens and Xbox Updates Revealed,” The Telegraph, March 30, 2016, <http://bit.ly/1qjCkbo>

Find your voice

Speaking directly to a brand appears to strengthen emotional relationships. Brands have a great opportunity to define their own voice

Voice is a hugely significant opportunity for brands. Our neuroscience research gives early indication that speaking to a brand delivers a deeper emotional connection than interacting with it through type or touch. When people asked a question involving a brand name, their brain activity showed a significantly stronger emotional response compared to people typing that same brand question (see Fig 9). The act of saying a brand name appears to strengthen the pre-existing emotional associations to a greater degree than typing it.

“People respond to sound with profound depth emotionally and physically, but they don’t necessarily understand that what they are being moved by is audio.”

Nick Ryan, composer, sound designer, artist and audio specialist

This emphasizes the need for brands to really craft the sound of their own voice, as it shows that voice interactions can have a deeper impact on brand perceptions. While the voice assistant will act as a critical gatekeeper for many purposes, such as discovery, there will be instances

where consumers will expect to talk to a brand directly, especially where they already have a strong relationship with the brand, such as in banking or grocery.

“Companies will now need to think about the actual voice of their brand ... They have to think about how their brand sounds, and the words and language that their brand uses when communicating with customers ... the personality of their brand as it’s presented to users.”

Martin Reddy, cofounder and chief technology officer, PullString

In addition, consumers are already looking for greater variety in voices. Just under three quarters (72%) of regular voice technology users say “I think brands should have unique voices and personalities for their apps/Skills and not just use the assistant on my phone.”

Some of the early Skill pioneers have started to use different voices to reflect the service they offer. The BBC News Skill broadcasts a pre-recorded presenter, while the Grand Tour Skill uses the

prerecorded voices of Jeremy Clarkson et al. In time, we will see the development of increased personality, whether through the use of more celebrity voices or the design of a unique tone of voice, once the AI capabilities allow the response to be truly interactive.

62% of regular voice technology users agree with the statement “I like the idea of being able to give my voice assistant a celebrity’s voice.”

Our research has shown that people have strong preconceived notions about what a brand should sound like, so developing the brand’s voice will be an increasingly important strategic consideration. Hit the right notes and you can be music to your customers’ ears.

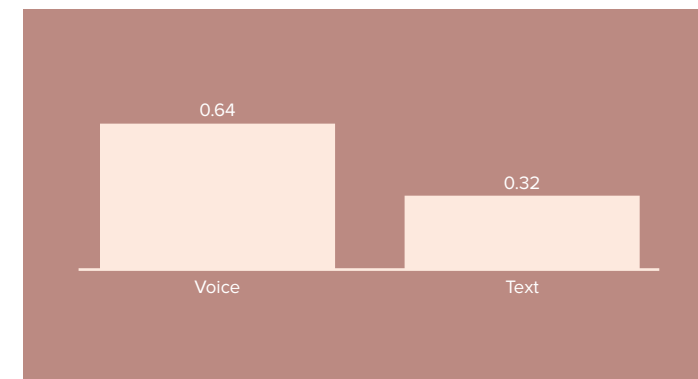


Figure 9
Emotional response for branded questions*

Source: Neuro-Insight study Feb 2017; n = 102 UK smartphone users. Brain activity measured using SST headsets; unit of measurement is radians, which equates to positivity of brain response

**Specific requests for brands*

Short-term brand implications

What marketers can do to prepare for a voice-activated world in the short term

1. Consider how voice could genuinely augment the touchpoints on your consumer journey.

How could a voice interaction add value to or remove friction from the consumer experience?

2. Learn the rules of engagement in conversational commerce.

Build a chatbot and explore the type of conversations you can start with your customer.

3. Become a voice native yourself.

Understand how relationships with technology and brands change when you can speak to them, and find out what works and what doesn't.

4. Experiment with voice-user interfaces.

Test and learn through an Alexa Skill or Google Conversation Actions trial. Explore how you can provide utility to your customers or drive new behaviors when there's no screen.

5. Use radio or interactive audio ads to develop your brand's voice.

What do people want to hear when they speak directly to your brand?

Long-term brand implications

What marketers can do to prepare for a voice-activated world in the medium to long term

1. Dial up your privacy credentials.

Review and develop your personal data privacy policies. Position yourself as the brand to trust.

2. Review whether your search activity is prepared for voice.

Does your keyword strategy capitalize on the long tail of conversational search terms?

3. See what you can learn from how customers talk about your brand.

Try using voice-analytics software to detect how your customers really feel.

4. Re-evaluate your PR efforts.

Getting your brands recommended by respected journalists and other thought leaders could be your best bet to staying relevant in the era of voice and the affiliate model.

5. Forge strategic partnerships for voice integration.

Work with retailers and service brands to be the brand they recommend.

Thank You

The Innovation Group and Mindshare Futures would like to thank the experts who generously gave their time and shared their ideas with us for this project.



DUNCAN ANDERSON

*Chief technology officer,
IBM Watson Europe*

Duncan Anderson is the European CTO for IBM Watson, a role he has held for the past three years. In this role he leads the European technical team for one of the world's foremost AI/cognitive computing efforts. He works with customers and technologists of all types and backgrounds to help bring the promise of cognitive computing to reality. Over the past 15 years, Duncan has held a series of senior technical leadership positions within IBM.



HEATHER ANDREW

*CEO,
Neuro-Insight UK*

With a background in marketing and communication, Heather Andrew has held marketing director roles at Nestlé and a television channel, as well as working as a consultant at PwC and OxfordSM. She is CEO of Neuro-Insight UK, working alongside the neuroscientists to run the client-facing part of the business.



NIR EYAL

*Bestselling author and lecturer
at Stanford University*

Nir Eyal is the author of the bestselling book *Hooked: How to Build Habit-Forming Products* about the intersection of psychology, technology and business. He runs Nirandfar.com, has founded two tech companies and taught at the Stanford Graduate School of Business and the Hasso Plattner Institute of Design at Stanford.



MICHAEL HILL

*Founder and managing director,
Radioplayer*

Michael Hill is founder and managing director of Radioplayer, a partnership between the BBC and UK commercial radio. He was previously at the BBC where he ran several national radio networks, developed new digital products and services, led major strategic transformation programs, and managed creative and journalistic teams.



MARTIN REDDY

*Cofounder and chief technology
officer, PullString*

Martin cofounded PullString and manages the company's engineering effort. He has published more than 40 peer-reviewed articles, authored half a dozen patents, and written two books. Martin spent five years in the Artificial Intelligence Center at SRI International and six years at Pixar, leading the development of its internal filmmaking software.



NICK RYAN

*Composer, sound designer, artist
and audio specialist*

Nick Ryan is a multi-award-winning sound artist and audio specialist. He is widely recognized as a leading thinker in relation to the future of sound and for creating unique and highly conceptual audio experiences that push the boundaries of audio.

We'd also like to call out the amazing talent that helped bring this report to life: Jane Barnes, Barneys New York, Claire Bischer, Vanessa Brown, Noel Bussey, Zac Cole, Delta Sound, Johanna Doyle, Fondation d'entreprise Hermès, Claire Gervat, Lucie Greene, Glenn Hatton, Hunger Magazine, Leif Johnson, Aleksandra Kingo, Katie Knott, Hester Lacey, Jeff Louviere, Peter Lowey, James Lucking, Florence Lujani, Philippe Medina, Mish Mane, Alan Masferrer, Rachel Moss, Guy Murphy, Vaclav Mlynar, Tom O'Donnell, Maria Pavlopoulos, Lucas Peon, Rankin, James Ratnarajah, Michael Reeder, Bryan Riddle, David Roth, Sophie Rhys Evans, Sennheiser, Daniel Thomas Smith, Jonathan Terry, Rhoneil Tiburcio, Igor Tolkachev, Leaf Troup, Nick Tsolkas, Jordan Walker Shuttlewood, Lauren Wahed, Davey Williamson and Will Wright.

About us

MINDSHARE FUTURES

Mindshare Futures is Mindshare's emerging media and technology research program. It focuses on the underlying consumer behavior behind future trends and advises marketers on the implications for communications. It produces annual trends publications, deep dive reports on specific tech trends, and offers consultancy services.

THE INNOVATION GROUP

The Innovation Group is J. Walter Thompson's futures, research and innovation unit. It charts emerging and future global trends, consumer change and innovation patterns—translating these into insight for brands. It offers a suite of consultancy services, including bespoke research, presentations, co-branded reports and workshops. It is also active in innovation, partnering with brands to activate future trends within their framework and execute new products and concepts.

ABOUT J. WALTER THOMPSON INTELLIGENCE

The Innovation Group is part of J. Walter Thompson Intelligence, a platform for global research, innovation and data analytics at J. Walter Thompson Company, housing three key in-house practices: SONAR™, Analytics and the Innovation Group. SONAR™, J. Walter Thompson's research unit, develops and exploits new quantitative and qualitative research techniques to understand cultures, brands and consumer motivation around the world. Analytics focuses on the innovative application of data and technology to inform and inspire new marketing solutions. It offers a suite of bespoke analytics tools.

REPORT CONTRIBUTORS

Contributing author:

Marie Stafford, J. Walter Thompson

Qual lead:

Tilly Flannery Ashworth, Mindshare

Quant lead:

Sarah Tilley, J. Walter Thompson

Additional research:

Jade Perry, J. Walter Thompson

Ella Britton, J. Walter Thompson

Josie Ung, Mindshare

Irina Lim, Mindshare

Contact

ELIZABETH CHERIAN

UK Director, The Innovation Group
J. Walter Thompson
elizabeth.cherian@jwt.com
@elizcherian

JEREMY POUNDER

Futures Director
Mindshare
Jeremy.Pounder@mindshareworld.com
@jezp76

MINDSHARE FUTURES



J. WALTER THOMPSON LONDON

#SpeakEasy2017