
How Speech Recognition Is Set to Disrupt



By , INSEAD Senior Affiliate Professor of Strategy

In the digital age, nothing stays the same for very long.

Back in 2012, only 10 percent of web pages were accessed via mobile devices. Who would have thought that less than six years later, this would be more than **50 percent** and we would be consuming news and video *en masse* on our smartphones?

The digital world is characterised by seismic shifts – great leaps in innovation that deeply impact the digital ecosystem and recalibrate user behaviour. And all the signs are that we’re approaching a new tipping point: voice control.

Until recently, voice technology was the butt of a plethora of jokes. A 2011 **viral video** from the BBC has two Scots trapped in a voice-activated lift with the digital assistant unable to interpret their regional accent. But such is the pace of innovation that today the shoe is very much on the other foot.

Try asking Siri, Apple’s iPhone voice interface, if it is a robot and one possible answer is: *Well, I don't want to brag, but I got a B+ on the Turing test!*

The joke, it would seem, is now on us.

Speaking up

Voice technology has reached maturity. And it has happened fast. Early in 2018, the quality of speech recognition exceeded the magical **95 percent accuracy** mark. That's the same quality as normal communication between human beings. And the impact is already being felt.

The most recent Microsoft Windows update actively pushes its voice feature – a mechanism that enables you to dictate messages at the speed of normal speech, which is four times faster than typing. In 2016, the company claimed that **25 percent** of the searches through its Windows 10 taskbar were voice searches. Market research firm **comScore** expects that 50 percent of all searches will use voice by 2020.

It's the same story with hardware and apps. The voice ecosystem is developing so fast that it's estimated a staggering **75 percent** of households in the United States will own a voice-activated smart speaker within the next two years. Meanwhile there are more than 2,600 voice apps (called “skills”) available for download on Amazon's Alexa Appstore.

Keeping pace with innovation

The speed of adoption points to the next major shift in digital, with voice poised to take its place alongside text and video. For companies, the implications of this are likely to be greater than we can currently imagine.

Recent history shows us that businesses often struggle to anticipate the full impact of these digital fluctuations. When the internet disrupted the newspaper sector, incumbents initially placed their content online in PDF format, essentially failing to understand that behavioural dynamics had changed more radically than format and appearance. Only over time did papers start to exploit the visual and interactive possibilities.

The same is true of mobile. Companies drastically underestimated the implications of mobile. Many of them focused on website optimisation only, oblivious to the opportunities to leverage geo-localisation or the easy upload of user-generated pictures and video. Even more significantly, with mobile innovation came a shift in preference from websites to apps – making it significantly more difficult for companies to reach the consumer, as we no longer explore the internet but refer time and again to a few trusted apps, especially in **Asia**.

Mobile has also transformed advertising revenue models. Mobile ad formats leverage speedier, impulse-purchasing behaviours rather than exploration. The exponential growth of Facebook is at least partly due to its timely transition to a mobile-first platform. This move enabled it to replace the traditional banner ad with more effective timeline advertising.

Voice has the potential to change the playing field just as radically as mobile.

The sound of the future

Take the search function, for instance.

With voice, we can expect users to spend less time conducting lengthy or cumbersome searches themselves. They are more likely to leave search tasks to a “voice butler” – AI apps that can source the best flight, order the cheapest products, find the right song or book the most romantic table in a fraction of the time it might take a human being to type words into the search bar. These butlers will become the internet’s new gatekeepers and it is for this reason that the tech giants – Amazon’s Echo and Alexa, Google Home – are vying for the prime position in the smart speaker market.

Then there’s ad revenue. Naturally, voice will make it harder to earn money with visual ads and consumers will have limited patience with audio ads. Revenue will thus shift away from advertising to sales commission and subscription models. With regard to brand building, **sound** and storytelling are likely to become just as important as the visual dynamics.

With these major shifts in digital, there are inevitably winners and losers.

The advent of voice could make incumbents like Google and Facebook more vulnerable given their reliance on advertising revenue. Amazon’s business model is better geared to the voice world, especially because the company will be able to promote its own branded products. Microsoft also stands a good chance to expand its subscription model to voice services.

Whatever the future, astute businesses would do well to start thinking through how the next tipping point is going to impact user behaviour and their bottom line. And plan their strategies to leverage voice accordingly.

You heard it here first.

Annet Aris is a Senior Affiliate Professor of Strategy at INSEAD. She is also a board member of Thomas Cook PLC in London, ASML N.V. in Veldhoven,

Randstad N.V. in Diemen, ASR Nederland N.V. in Utrecht and Jungheinrich AG in Hamburg.

Annet was named one of the 50 most inspirational women in the European technology sector for 2016 by [Inspiring Fifty](#).

Follow INSEAD Knowledge on [Twitter](#) and [Facebook](#).

Find article at

<https://knowledge.insead.edu/strategy/how-speech-recognition-set-disrupt>