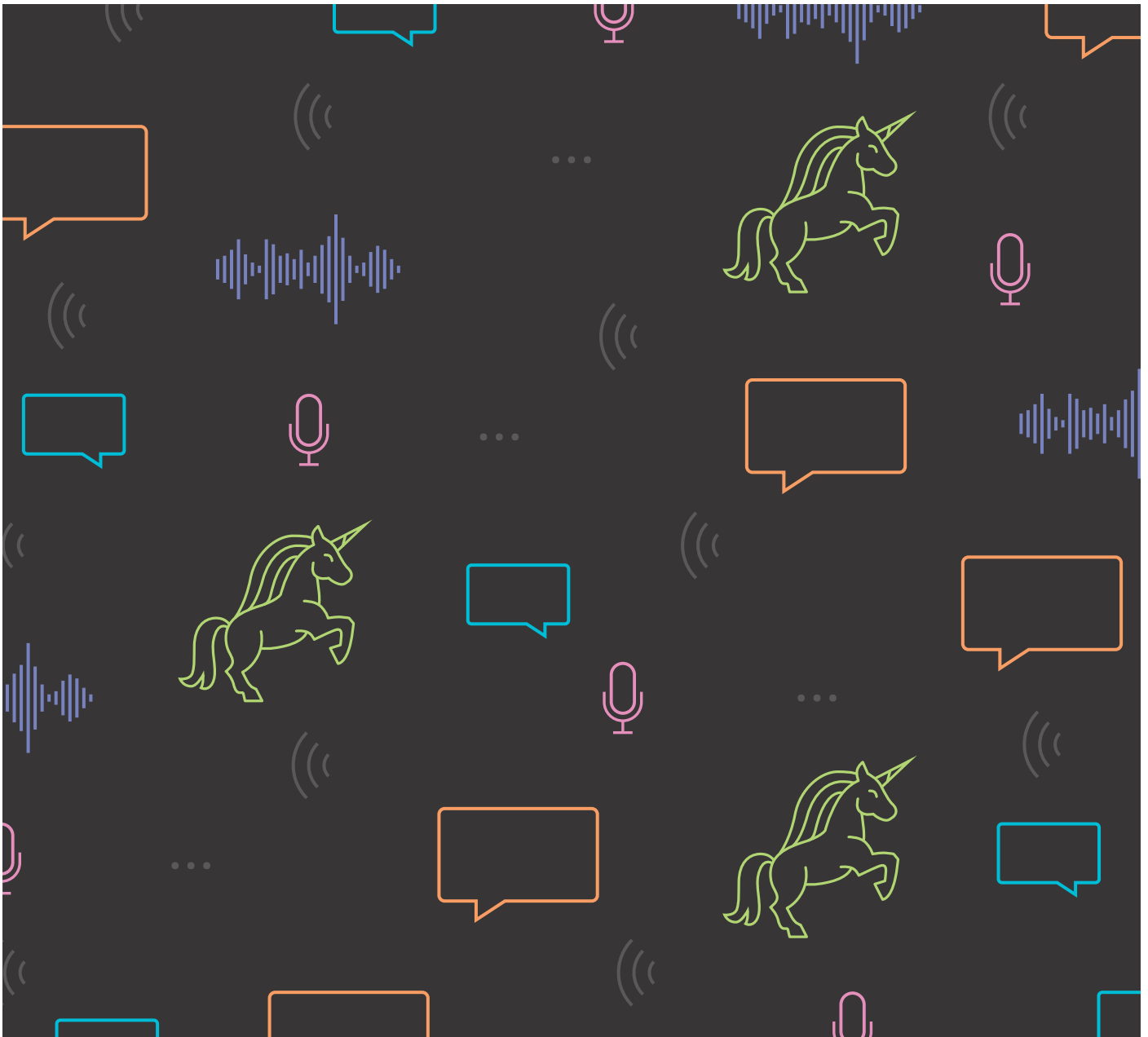# **Voice:** Welcoming the next generation of disruptors

"The best interface is no interface"
Golden Krishna

After many years of typing on keyboards and then tapping on mobile screens, we are now reverting to the original "user interface" – voice. While we are yet to feel the full impact of voice technology, it is gradually transforming how we interact with the digital world.

Importantly, this shift will go far beyond checking the weather forecast, requesting music or setting a timer. Voice represents both a platform and user interface (UI) shift comparable to the web and smartphones. And as the front door to artificial intelligence and machine learning, voice technology will find powerful new ways to serve us.

At Mangrove we believe voice will be one of the defining themes of the next decade. As well as impacting every business with a digital presence, it will spawn a new generation of technology companies in every category. And together with AI, it will radically change the relationship between human and machine as well as make the next generation of software the most valuable yet.

The shift to voice

Voice technology has come a long way since Apple's launch of Siri in 2011. Thanks to developments in machine learning, we've seen considerable advances in both speech recognition (transcribing of audio to text) and natural language processing (identifying the command within the text). Indeed speech systems have surpassed 95% word accuracy rate, meaning their ability to understand language is equivalent to humans. With voice offering a far more natural, convenient and efficient form of communication than typing or texting, it is now destined to become the primary interface between human and machine.

There are already 2.5 billion voice assistants in existence according to Juniper Research and, with tech giants jostling for position, more and more households are becoming voice-enabled. The number of smart speakers in US homes grew a staggering 78% in 2018, from 66.7 million to 188.5 million with Amazon holding 64.6% of the market. Canalys expects the smart speaker installed base to grow to 500m worldwide by 2023.

Voice is also emerging as the gateway to the smart home and internet of things (IoT). Amazon already claims to have 28K smart home devices enabled with Alexa across 4,500 unique brands from smart TVs and cookers to robot vacuums.

As well as entering our homes, voice assistants are also entering hotel rooms, such as those of Marriott, and are proving especially popular in cars where they are being integrated into infotainment systems. Amazon is in talks with a number of car manufacturers and is reported to have over 1 million pre-orders for Echo Auto, which provides a hands-free experience for everything from making calls and directions to controlling smart home devices.

Voice technology is also transcending language barriers. Google now has capabilities for 30 languages – including the ability to understand users who speak more than one language at a time. Interestingly, voice adoption is growing even faster in China – where Alibaba takes the lead in smart speaker shipments.

The colourful array of applications enjoying traction is reminiscent of the early days of the App Store. Amazon Alexa is already providing breastfeeding support to new mums, encouraging kids to brush their teeth, guiding users through workouts, coaching football teams and helping people connect with God through daily prayers.

The number of voice assitants will surpass the human population by 2023
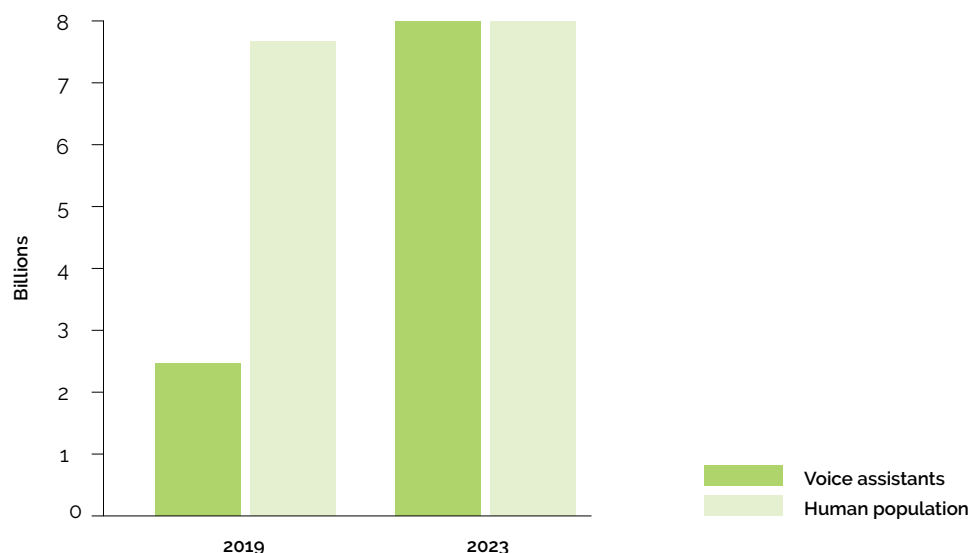
The rise of voice-commerce

Voice shopping is also growing with Alexa directly connected to one of the world's largest online stores. Amazon saw e-commerce on Alexa triple over the 2018 holidays compared to the year before. No wonder then that Walmart recently introduced voice ordering for its online grocery service through Google Assistant, allowing shoppers to add products to their list by talking to their Google Home, smartphones, or smartwatches.

Voice shopping is expected to jump from $2 billion in 2018 to $40 billion by 2022 across the US and UK according to Juniper. Interestingly, the three most commonly shopped categories through voice are all commoditized: grocery (20%), entertainment (19%) and electronics (17%), so having "Amazon Choice" status will be more important than ever for retailers. According to OC&C Strategy Consultants, 85% of Amazon customers end up selecting the product Amazon suggests, although how specific products get the designation remains a mystery.

Voice technology is already guiding consumers through purchasing, whether online or in physical stores. And FMCG manufacturers are exploring how voice-based experiences, such as cocktail making and whiskey tastings, can enhance their product positioning and drive loyalty.
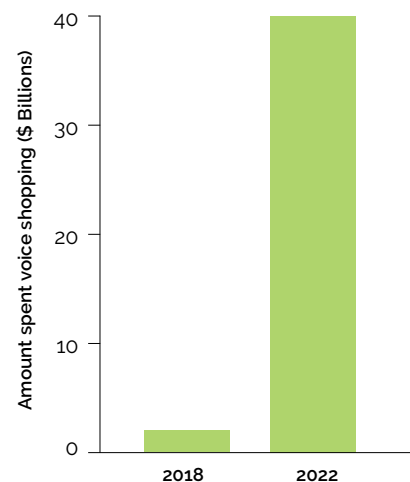
The importance of Amazon Choice status
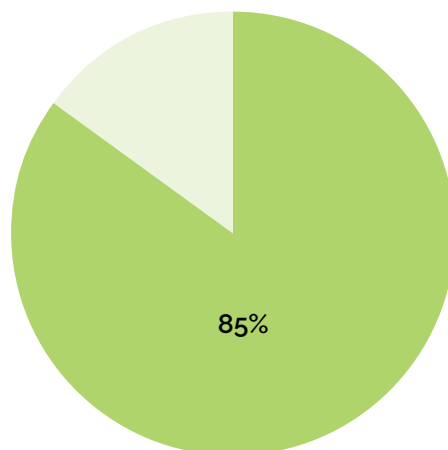
85% of consumers end up selecting the product Amazon suggests

The rise of voice shopping

Voice shopping is expected to jump from $2 billion in 2018 to $40 billion by 2022 across the US and UK

You will never be alone

Crucially, voice is moving beyond the transactional and becoming conversational. The ramifications of this are difficult to fully comprehend but consider this: *you will never be alone again, whether you talk to Alexa, another human being or another AI.*

The first wave of AI was about using algorithms to build recommendation engines and provide shortcuts on the internet, such as those by Google, Amazon and Facebook. Then came the second wave of AI, which was about helping businesses take advantage of all their structured data to improve decision-making. Companies such as Palantir have emerged to mine vast databases of structured data for hidden correlations that escape the naked eye and human brain.

The third wave of AI, which is being implemented now, sees AI getting upgraded with eyes, ears, and myriad other senses, collecting new data that was never before captured and using it to automate more complex processes. This is also where we move from command-based static processes to dynamic processes. This represents a tectonic shift in computing and it is going to revolutionize how we experience and interact with software, blurring the lines between the digital and physical worlds.

As AI systems become more adept in dialogue and collect more data about people and the world around them, they will acquire context - the many interconnected layers of accumulated knowledge that humans effortlessly acquire and apply in conversation. Context will allow humans and machines to interact and collaborate in a more natural way. And as the experience improves and becomes more useful, they will interact more regularly and machines will in turn become more knowledgeable.

This virtuous circle means that within three to five years, the conversational capabilities of computers will be vastly more sophisticated and transformative. Rather than spending our time typing on keyboards or tapping on phones, we will simply talk to an omnipresent AI. Importantly, the way we interact with software will change forever – forms will be dead and touchpoints will become 'listening points'.

This explains why the world's largest technology companies are investing so heavily in voice technology. Voice represents both a platform and user interface (UI) shift comparable to the web and smartphones, except of course that the shift to voice doesn't require any training. Having built dominant and seemingly untouchable platforms, these companies suddenly have everything to lose. No wonder Amazon already has over 10,000 employees working on Alexa and is one of the most prolific investors in voice-based startups.

The dawn of voice-first interfaces

The emergence of screenless phones: The Lightphone

The implications of this are profound – every company will need to rethink how it interacts with its audience. It is not just about adapting SEO efforts and owning Alexa 'Skills' or Google 'Actions'.

Companies will ultimately need to provide context aware voice-first interfaces. Those that are able to create the illusion of awareness and a sense of connection will be best placed to engage audiences. Those that respond predictably with scripted responses will be at a considerable disadvantage.

Technology companies heavily reliant on notifications for driving app usage may need to review their strategy. Consumers will have little patience for companies that breach the 'sonic threshold' with too many jingles (LinkedIn take note).

Just as the most compelling mobile apps have become the dominant platforms of today, the most engaging voice-first interfaces could become the dominant services of tomorrow. It is difficult to predict the ideal interface but it will likely only take 3 – 5 years for the next Airbnb, Monzo and Uber to emerge.

The emergence of screenless phones such as Light Phone provide a glimpse into a future without mobile apps we know and love. Given voice has a low learning curve and significantly reduces friction, it is reasonable to assume the voice economy will grow very quickly once we begin to realize its full potential.

Work in the age of voice

It is at work that we spend the most time interacting with software, so it is a logical assumption that it will be at work that we see the biggest change. While we may still swipe or tap screens, the keyboard becomes redundant. Afterall the average person can type at least 40 words per minute while the average person can speak about 150 words per minute.

The omnipresent AI will be able to do much of the tasks and processes we do in our daily work routines – respond to emails, gather information from meetings, circulate actions with team members, update the relevant systems with data and so on. It may also mean that we are no longer shackled to desks and are able to live much less sedentary lives.

Taking this a step further, the omnipresent AI will be able to spot high performers and inform others on the techniques that generate optimal outcomes. AI learning loops that continuously identify specific strategies from outliers and distribute the best performing advice to the rest of the network are already emerging. Far from replacing humans, these optimisation networks depend on humans to learn and improve the process – the human stays in the middle of the loop.

One example attracting significant interest is Chorus.ai, which joins calls and provides prompts for effective responses. Rather than giving a written script for a sales call, it spots how different approaches by team members lead to positive outcomes and encourages creativity.

By making humans smarter and more effective, this new generation of SaaS (software-as-a-service) offerings will be able to command higher subscription fees and look set to become the most valuable generation of business software yet.

From virtual beings to virtual doctors

Virtual influencer Lil Miquela is proving popular among fashion houses.

Already we are seeing virtual influencers gain momentum – Lil Miquela, a 19-year-old virtual influencer from Los Angeles, has clocked up with 1.5 million followers on Instagram and is now being featured in campaigns by major fashion brands such as Calvin Klein.

Fable, a creative studio in San Francisco, is taking this a step further with virtual beings - fully interactive characters that have their own personalities. Soon we will be able to create AI hologram companions - such as Blade Runner's Joi, which was manufactured by Wallace Corp to be 'everything you want to see'.

Virtual beings will likely change the face of entertainment and customer service, but could also play a vital role in supporting the elderly and dependent; after all those who identify as lonely have a 59% greater risk of health decline and social isolation is one of the most common causes of depression.

Barcelona-based Safe365 has demonstrated the appetite among the elderly for a tech-driven community and support network. Voice offers a very user-friendly way of monitoring health and cognitive function, as well as providing useful reminders for exercises and medication.

Indeed, voice and AI could ultimately replace doctors entirely. Already K Health has built an AI which combines the experiences of thousands of doctors and is capable of accurate diagnosis. With voice technology, it could also provide the comforting and reassuring tones of a human doctor.

Mangrove analysis of voice funding

Investment in voice startups is now accelerating according to our analysis. Voice startups have raised $786m already this year, significantly exceeding the $581m raised in 2018 and $298m raised in 2017. This rise in funding reflects a growing belief that voice technology will be transformative. The size of fundraisings has also increased markedly – with an average of $30m so far in 2019 versus $18m in 2018 and $17.5m in 2017.

Entertainment and podcasting is one area already attracting significant interest. Spotify acquired podcasting startups Gimlet and Anchor for a combined $337m. Other notable startups in this space include Sybel, which makes premium voice-first content.

Meanwhile Amazon's Alexa Fund has made 72 investments since its launch in 2015. Over the last 18 months it has invested in startups covering entertainment, healthcare, education, transport, e-commerce, hospitality and SaaS. This demonstrates Amazon's conviction that voice technology will change a wide range of industries.

Our analysis shows the dominance of the U.S. in voice startup activity – with over $1bn invested in voice startups based in the U.S. versus $118m in Europe. This is particularly concerning given voice technology represents a paradigm shift. European entrepreneurs and investors will need to respond quickly if the region wants to have a major stake in this next key phase of technology evolution.

Funding in voice startups
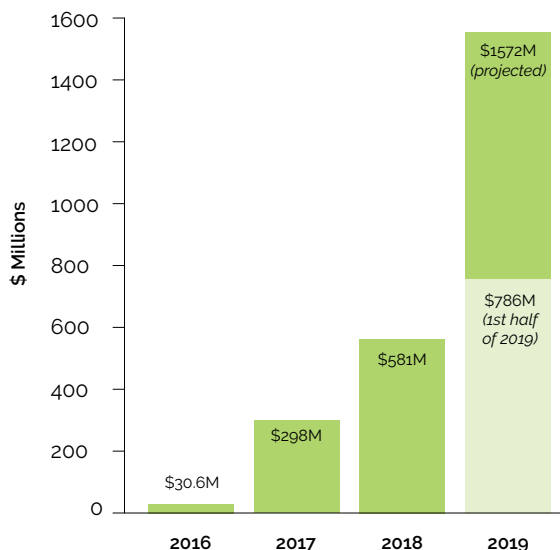
Funding in voice startups by country or region

USA: $1008.52M
Europe: $118M
Israel: $60.75M
Canada: $69M
China: $100M
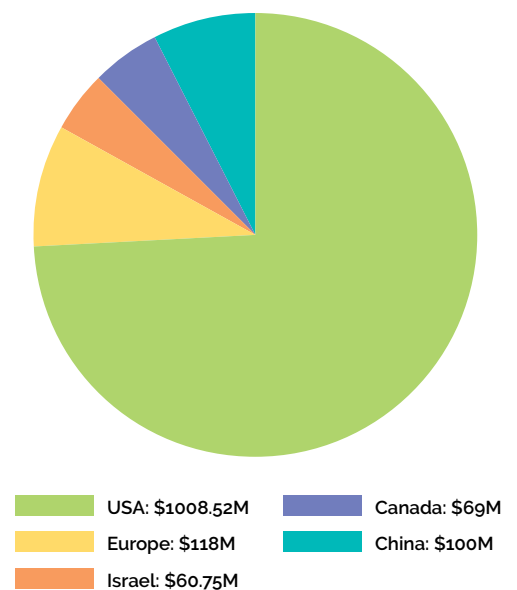
Challenges

One of the biggest challenges with voice is privacy. According to Microsoft, forty-one percent of voice assistant users are concerned about trust, privacy and passive listening. The discovery of microphones in Google's Nest devices probably didn't help, nor Amazon's filing of a patent to 'listen to all conversations in a room'.

It will be up to technology companies to respond to these concerns – something Apple has been doing particularly effectively. Its acquisition of Silk Labs, which develops on-device AI operating software that removes the need for cloud processing, was a shrewd move.

Apple has been less effective at building an ecosystem around voice. With its vision of the 'Knowledge Navigator', Apple's voice strategy was more clearly articulated in 1987 than it is now. The voice community expects Apple to release a SiriOS for its developer community at WWDC 2020 which would accelerate innovation and adoption.

"A SiriOS is desirable for enabling innovation and is viewed by many as required to match the progress made by Amazon and Google with their voice assistants," Bret Kinsella, Editor and Publisher at Voicebot.ai

One issue that has been popularized by movies, such as Her and Ex Machina, is the threat to humanity if we become so enamoured by robots that we fall in love with code instead of other humans. While it may sound far fetched to some, dating sims (romance simulation games) are already proving extremely popular. In China, a dating sim called Love and Producer was downloaded more than 7m times in its first month.

While challenging our notion of privacy and raising new questions about our relationship with technology, there are many reasons to be optimistic. It will drive a new wave of productivity gains which will benefit consumers, businesses and society. And it will help us to care for the elderly as well as make the internet accessible to 774m people around the world who can't read.

Predictions for the age of voice

- The death of the keyboard: instead of tapping
  and typing, we will be speaking and swiping.
  The keyboard will be largely redundant within
  5-10 years

- Voice will lead to a whole new product category
  including screenless smartphones

- Brands will soon be instantly recognisable
  through their 'sonic identity' and product search
  will change forever

- There will be a shift of power from software back
  to hardware because voice is application agnostic

- Virtual companions will be a reality for all your
  needs in life

- Apple's launch of SiriOS in 2020 will unleash huge
  innovation in the voice economy, which will be
  worth $1 trillion by 2025 (surpassing the growth
  of the mobile app economy)

About Mangrove Capital Partners

Mangrove Capital Partners (www.mangrove.vc) is Europe's leading early stage venture capital firm. It works with top entrepreneurial talent at the earliest stages of innovation, with the aim of being the first institutional investor: the firm has co- created projects and regularly injects funds prior to product launch, often in unproven, unusual or unfavoured technologies. Mangrove manages more than $1 billion in assets and is headquartered in Luxembourg with offices in Berlin and Tel Aviv.

Mangrove Capital Partners is Europe's leading early stage venture capital firm.

www.mangrove.vc