

# Study shows people are not good at recognising voices

June 19 2017, by Carolyn Mcgettigan And Nadine Lavan



Credit: AI-generated image ([disclaimer](#))

*"Alexa, who am I?"* Amazon Echo's voice-controlled virtual assistant, Alexa, doesn't have an answer to that – yet. However, for other applications of speech technology, computer algorithms are increasingly able to discriminate, recognise and identify individuals from voice recordings.

Of course, these algorithms are far from perfect, as was recently shown when a BBC journalist broke into his own voice-controlled bank account [using his twin brother's voice](#). Is this a case of computers just failing at something humans can do perfectly? We decided to find out.

Each human being has a voice that is distinct and different from everyone else's. So it seems intuitive that we'd be able to identify someone from their voice fairly easily. But how well can you actually do this? When it comes to recognising your closest family and friends, you're probably quite good. But would you be able to recognise the voice of your first primary school teacher if you heard them again today? How about the guy on the train this morning who was shouting into his phone? What if you had to pick him out, not from his talking voice, but from samples of his laughter, or singing?

To date, research has only explored voice identity perception using a limited set of vocalisations, for example sentences that have been read aloud or snippets of conversational speech. These studies have found that we can actually recognise voices of familiar people's speech [quite well](#). But they have also shown that there are problems: ear-witness testimonies are notoriously [unreliable and inaccurate](#).

It's important to keep in mind that these studies have not captured much of the flexibility of the sounds we can make with our voices. This is bound to have an effect on how we process the identity of the person behind the voice we are listening to. Therefore, we are currently missing a very large and important piece of the puzzle.

Recognising voices requires two broad processes to operate together: we need to distinguish between the voices of different people (telling people apart) and we need to be able to attribute a single identity to all the different sounds (talking, laughing, shouting) that can come from the same person ("telling people together"). We set out to investigate the

limits of these abilities in humans.

## Voice experiment

Our recent study, [published in the Journal of Experimental Psychology: General](#), confirms that voice identity perception can be extremely challenging. Capitalising on how variable a single person's voice can be, we presented 46 listeners with laughter and vowels produced by five people. Listeners were asked to make a very simple judgement about pairs of sounds: were they made by the same person, or by two different people? As long as they could compare vowels to vowels or laughter to laughter respectively, discriminating between speakers was relatively successful.

But when we asked our listeners to make this judgement based on a mixed pair of sounds, such as directly comparing vowels to laughter in a pair, they couldn't discriminate between speakers at all – especially if they were not familiar with the speaker. However, even though a sub-group of people who knew the speakers performed better overall, they still struggled significantly with the challenge of "telling people together".

Similar effects have been reported by studies showing, for example, that it is [difficult to recognise a bilingual speaker](#) across their two languages. What's surprising about these findings is how bad voice perception can be once listeners are exposed to natural variation in the sounds that a voice can produce. So, it's intriguing to consider that while we each have a unique voice, we don't yet know how useful that uniqueness is.

But why have we evolved to have unique voices if we can't even recognise them? That's really an open question so far. We don't actually know whether we have evolved to have unique voices – we also all have different and largely unique fingerprints, but there's no evolutionary

advantage to that as far as we can tell. It just so happens that based on differences in anatomy and, probably most importantly, how we use our voice, that we all sound different to each other.

Luckily computer algorithms are still able to make the most of the individuality of the human voice. They have probably already outdone humans in some cases – and they will keep on improving. The way these [machine-learning algorithms](#) recognise speakers is based on mathematical solutions to create "voice prints" – unique representations picking up the specific acoustic features of each individual voice.

In contrast to computers, humans might not know what they are listening out for, or [how to separate out these acoustic features](#). So, the way that [voice](#) prints are created for the algorithms is not closely modelled on what human listeners appear to do – we're still working on this. In the long term, it will be interesting to see if there is any overlap in the way human listeners and machine-learning algorithms recognise voices. While human listeners are unlikely to glean any insights from how computers solve this problem, conversely we might be able to build machines that emulate effective aspects of human performance.

It is rumoured that Amazon is currently working on teaching Alexa how to [identify specific users by their voice](#). If this works, it will be a truly impressive feat and may put a stop to [further unwanted orders of dollhouses](#). But, do be patient if Alexa makes mistakes – you may not be able to do it any better yourself.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: Study shows people are not good at recognising voices (2017, June 19) retrieved 10 April 2024 from <https://medicalxpress.com/news/2017-06-people-good-recognising-voices.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.