

ARTS & CULTURE

Sprachassistenten wie Siri oder Alexa antworten mit zunehmend menschlich klingenderen Stimmen. Das jedoch könnte uns dazu verleiten, zu viel von unseren digitalen Assistenten zu erwarten. Ist doch ein Gespräch zwischen Menschen weit mehr als nur der Austausch von Informationen und Befehlen.

SOUNDING TOO HUMAN

Voice assistants like Siri or Alexa are sounding more and more human-like. But this aim for perfection might be deceiving and might lead us to expect too much from our digital assistants. A human conversation is, after all, much more than just giving orders and exchanging information.

More than 200 million homes now have a smart speaker providing voice-controlled *access* to the internet, according to one global *estimate*. Add this to the talking virtual assistants installed on many smartphones, not to mention *kitchen appliances* and cars, and that's a lot of Alexas and Siris.

Because talking is a fundamental part of being human, it is *tempting* to think these assistants should be designed to talk and behave like us. While this would give us a *relatable* way to interact with our devices, replicating *genuinely* realistic human conversations is incredibly difficult. What's more, research *suggests* making a machine sound human may be unnecessary and even *dishonest*. Instead, we might need to rethink how and why we interact with these assistants and learn to *embrace* the *benefits* of them being a machine.

Speech technology designers often talk about the concept of "humanness". Recent developments in artificial voice development have resulted in these systems' voices *blurring* the line between human and machine, sounding *increasingly* human-like. There have also been *efforts* to make the language of these interfaces appear more human.

Perhaps the most famous is Google Duplex, a service that can book appointments over the phone. To add to the human-like nature of the system, Google included *utterances* like "hmm" and "uh" to its assistant's speech output – sounds we *commonly* use to signal we are listening to the conversation or that we intend to start speaking soon. In the case of Google Duplex, these were used with the *aim* of sounding natural. But why is sounding natural or more human-like so important?

Chasing this *goal* of making systems sound and behave like us perhaps *stems* from pop culture inspirations we use to *fuel* the design of these systems. The idea of talking to machines has fascinated us in literature, television and film

for decades, through characters such as HAL 9000 in 2001: A Space Odyssey or Samantha in Her. These characters *portray seamless* conversations with machines. In the case of Her, there is even a love story between an operating system and its user. Critically, all these machines sound and respond the way we think humans would.

There are interesting technological challenges in trying to *achieve* something *resembling* conversations between us and machines. To this end, Amazon has recently *launched* the Alexa Prize, looking to "create socialbots that can *converse coherently* and *engagingly* with humans on a range of current events and popular *topics* such as entertainment, sports, politics, technology, and fashion". The current round of competition asks teams to produce a 20-minute conversation between one of these bots and a human *interactor*.

These grand challenges, like others across science, clearly *advance* the state of the art, bringing planned and unplanned benefits. Yet when *striving* to give machines the ability to truly converse with us like other human beings, we need to think about what our spoken interactions with people are actually for and whether this is the same as the type of conversation we want to have with machines.

MORE THAN GETTING STUFF DONE

We converse with other people to get stuff done and to build and maintain relationships with one another – and often these two *purposes intertwine*. Yet people see machines as tools serving limited purposes and hold little appetite for building the kind of relationships with machines that we do every day with other people.

Pursuing natural conversations with machines that sound like us can become an unnecessary and *burdensome objective*. It creates unrealistic *expectations* of systems that

can actually communicate and understand like us. Anyone who has interacted with an Amazon Echo or Google Home knows this is not possible with existing systems.

This matters as people need to have an idea of how to get a system to do things which, because voice-only interfaces have limited *buttons* and *visuals*, are guided significantly by what the system says and how it says it. The importance of interface design means humanness itself may not only be questionable but *deceptive*, especially if used to *fool* people into thinking they are interacting with another person. Even if their *intent* may be to create *intelligible* voices, tech companies need to *consider* the potential *impact* on users.

Rather than consistently embracing humanness, we can accept that there may be fundamental limits, both technological and philosophical, to the types of interactions we can and want to have with machines.

We should be inspired by human conversations rather than using them as a *perceived* gold standard for interaction. For instance, looking at these systems as *performers* rather than human-like *conversationalists* may be one way to help to create more *engaging* and expressive interfaces. *Incorporating* specific elements of conversation may be necessary for some contexts, but we need to think about whether human-like conversational interaction is necessary, rather than using it as a *default* design goal.

It is hard to *predict* what technology will be like in the future and how social *perceptions* will change and develop around our devices. Maybe people will be ok with having conversations with machines, becoming friends with robots and seeking their *advice*.

But we are currently sceptical of this. In our view it is all to do with context. Not all interactions and interfaces are the same. Some speech technology may be required to *establish* and *foster* some form of social or emotional *bond*, such as in specific *healthcare* applications. If that is the aim, then it makes sense to have machines converse more *appropriately* for that purpose – perhaps sounding human so the user gets the right type of expectations.

Yet this is not universally needed. *Crucially*, this human-likeness should link to what the systems can actually do with conversation. Making systems that do not have the ability to converse like a human sound human may do far more *harm* than good.

<<< Leigh Clark & Benjamin Cowan >>>

Dieser Beitrag ist ursprünglich auf „The Conversation“ erschienen. Die Veröffentlichung erfolgt im Rahmen seiner Creative Commons Lizenz. Der Original-Beitrag findet sich unter <https://theconversation.com/voice-assistant-technology-is-in-danger-of-trying-to-be-too-human-125856>

<https://lmhclark.com>
<http://www.benjaminclark.com>

Die privaten Webseiten der beiden Autoren Leigh Clark und Benjamin Cowan.

<i>access</i>	Zugang, Zugriff
<i>achieve, to</i> // ə'tʃi:v	erreichen, schaffen
<i>advance, to</i>	vorantreiben, voranbringen
<i>advice</i>	Rat, Empfehlung
<i>aim</i>	Ziel, Absicht
<i>appropriately</i> // ə'prəʊpɪətli	angemessen, entsprechend
<i>benefit</i>	Vorteil, Vorzug, Nutzen
<i>blur, to</i> // blɜ:	verwischen, undeutlich machen
<i>bond</i>	Bindung, Beziehung
<i>burdensome</i> // 'bɜ:dnəm	beschwerlich, mühsam
<i>button</i>	Schalter, Knopf
<i>chase, to</i>	verfolgen, jagen
<i>coherently</i>	zusammenhängend, stimmig
<i>commonly</i>	im Allgemeinen, gemeinhin
<i>consider, to</i>	bedenken, erachten
<i>conversationalist</i>	Gesprächspartnerin, Plauderer
<i>converse, to</i>	sich unterhalten, sprechen
<i>crucially</i> // 'kru:ʃəli	maßgeblich, entscheidend
<i>deceive, to</i> // dɪ'si:v	täuschen, trügen
<i>deceptive</i> // dɪ'septɪv	trügerisch, irreführend
<i>default</i>	Standard-, Vorgabe
<i>dishonest</i>	unredlich, unehrlich
<i>effort</i>	Anstrengung, Bemühung
<i>embrace, to</i>	hier: annehmen, begrüßen
<i>engaging(ly)</i> // ɪn'geɪdʒɪŋ(li)	ansprechend, einnehmend
<i>establish, to</i>	aufbauen, schaffen
<i>estimate</i>	Schätzung, Annahme
<i>expectation</i>	Erwartung
<i>fool, to</i>	täuschen, irreführen
<i>foster, to</i>	pflügen, fördern
<i>fuel, to</i>	schüren, befeuern
<i>genuinely</i> // 'dʒenjuːli	echt, authentisch
<i>goal</i>	Ziel, Absicht
<i>harm</i>	Schaden, Unheil
<i>healthcare</i>	Gesundheitswesen
<i>impact</i>	Einfluss, Auswirkung
<i>incorporate, to</i>	einbeziehen, einbinden
<i>increasingly</i>	zunehmend
<i>intelligible</i> // ɪn'telɪdʒəbl	verständlich, erkennbar
<i>intent</i> // ɪn'tent	Absicht, Intention
<i>interactor</i>	Interagierende/r
<i>intertwine, to</i> // ɪntə'twaɪn	ineinandergreifen, verflechten
<i>kitchen appliance</i>	Küchengerät
<i>launch, to</i> // lɔ:ntʃ	einführen, starten
<i>objective</i>	Zielsetzung, Aufgabe
<i>perceive, to</i> // pə'si:v	auffassen, empfinden
<i>perception</i>	Auffassung, Wahrnehmung
<i>performer</i>	Darsteller/in, Interpret/in
<i>portray, to</i>	darstellen, schildern
<i>predict, to</i>	voraussagen, vorhersehen
<i>purpose</i> // 'pɜ:pəs	Zweck, Absicht
<i>pursue, to</i> // pə'sju:	streben nach, verfolgen
<i>relatable</i>	zuordenbar
<i>resemble, to</i>	ähneln, gleichen
<i>seamless</i>	nahtlos
<i>stem, to</i>	stammen, sich ableiten
<i>strive, to</i>	anstreben, versuchen
<i>suggest, to</i> // sə'dʒest	hier: nahelegen, andeuten
<i>tempt, to</i>	verlocken, verleiten
<i>topic</i>	Thema, Gegenstand
<i>utterance</i> // 'ʌtərəns	Äußerung, Ausspruch
<i>visual</i>	hier: grafische Anzeige