

"Buckets of rain!" - Effects of colloquial and formal speech style of a voice assistant on humanness, competence, trust, and intentions to use

Frehmann, Katharina

Erstveröffentlichung / Primary Publication

Konferenzbeitrag / conference paper

Empfohlene Zitierung / Suggested Citation:

Frehmann, K. (2023). "Buckets of rain!" - Effects of colloquial and formal speech style of a voice assistant on humanness, competence, trust, and intentions to use. In M. Ziegele, A. S. Kümpel, & T. Dienlin (Eds.), *Jahrestagung der DGPUK-Fachgruppe Rezeptions- und Wirkungsforschung 2022: Tagungsband* (pp. 41-57). Düsseldorf: Deutsche Gesellschaft für Publizistik- und Kommunikationswissenschaft e.V. <https://doi.org/10.21241/ssoar.87698>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>

“Buckets of rain!” - Effects of colloquial and formal speech style of a voice assistant on humanness, competence, trust, and intentions to use

Katharina Frehmann

Department of Social Sciences, Heinrich Heine University

Zusammenfassung

Sprachassistenten als Informationsintermediäre werden beliebter. Typisch für sie ist ihre anthropomorphe Erscheinung, wie ihre menschlichen und seit neustem auch umgangssprachlichen Stimmen. Laut der CASA-Forschung kann Anthropomorphismus die Wahrnehmung von Menschlichkeit steigern, was auch Vertrauen stärkt. Umgangssprache könnte also über wahrgenommene Menschlichkeit das Vertrauen von Nutzern erhöhen, welche jedoch oft intransparente, unvollständige Antworten erhalten. Andererseits könnte der Sprachstil ähnlich wie Dialekt und Akzent die Wahrnehmung von Kompetenz und damit das Vertrauen verringern. Die Studie untersucht, wie der Sprachstil eines fiktiven Sprachassistenten Vertrauen und Nutzungsabsichten über die wahrgenommene Menschlichkeit und Kompetenz beeinflusst. Eine experimentelle Online-Vignetten-Studie mit einem 2 (formelle vs. umgangssprachliche Sprache) x 2 (informelle Alltags- vs. formelle politische Themen) Between-Subjects-Design (N = 244) ermittelte, dass der Sprachstil keine Wirkung auf die wahrgenommene Menschlichkeit hat. Umgangssprache reduziert jedoch die Kompetenz, was einen negativen Effekt auf Vertrauen vermittelt. Ein umgangssprachlicher Sprachassistent scheint also keine positiven Einstellungen und Verhaltensweisen von Nutzenden zu fördern. Unabhängig von der experimentellen Manipulation wurde jedoch ein direkter Effekt der Menschlichkeit auf die Nutzungsabsicht beobachtet, was dennoch die Relevanz des anthropomorphen Designs für die Nutzung zeigt.

Keywords: Sprachassistent, Sprachstil, Umgangssprache, Menschlichkeit, Kompetenz, Vertrauen, Nutzungsabsicht

Summary

Voice assistants are gaining popularity as new information intermediaries. A typical feature is their anthropomorphic design reflected in their voices and, as of late, colloquial speech style. According to CASA, anthropomorphism leads to perceptions of humanness in technology which strengthens trust. Colloquialism could thus increase the trust of users via perceived humanness which the opaque assistants and their oftentimes deficient answers might not deserve. On the other hand, colloquialism, similar to the use of dialect and accents, could also decrease the perception of competence, and therefore diminish trust. The current study investigates how the speech style of a fictitious voice assistant affects people's trust and their intentions to use the assistant via perceived humanness and competence of the assistant. A vignette-based experimental online study with a 2 (formal vs. colloquial language) x 2 (informal everyday vs. formal political issues) between-subjects design was conducted (N = 244). Results showed that speech style had no effect on perceived humanness, but colloquial speech reduced perceived competence, which mediated a negative effect on trust. An informally speaking voice assistant thus does not seem to further users' positive attitudes and behaviors towards it. However, independent of the experimental manipulation, a direct effect of perceived humanness on intentions to use was observed, stressing the relevance of anthropomorphic design for users.

Keywords: Voice assistants, speech style, colloquialism, humanness, competence, trust, intentions to use

Introduction

Voice assistants are gaining popularity worldwide (Newman et al., 2019, 2020, 2021, 2022). They are digital software agents on mobile or stationary devices that respond to spoken requests with spoken responses and can be interpreted as artificial companions that rely on automation (Hepp, 2020; Hoy, 2018). They are being used for different tasks, such as entertainment (e.g., playing music), assistance (e.g., setting timers) and the search for information (e.g., Arnold et al. 2019). Their use for information purposes and news has been growing internationally in recent years (Newman et al., 2019, 2020) and in 2020, up to 20 percent of German adults asked their voice assistant about the COVID-19 crisis on a weekly basis (Viehmann et al., 2022). Thus, voice assistants increasingly act as gatekeepers and information intermediaries that select, process and present information and therefore gain relevance due to their possible influence on the opinion building of individuals (Frehmann et al., 2022). A special feature of voice assistants is their anthropomorphic design, reflected in the human-sounding voices, names like "Alexa" or "Siri", and their social reactions and behaviors. According to research regarding the CASA paradigm (Computers as social actors), these anthropomorphic features can be considered social cues that trigger social perceptions and reactions in users and increase trust and intentions to use the technology (Lee & Nass, 2010). Current developments of anthropomorphism in voice assistants involve an increasing use of dialect and colloquial speech of the assistants (Blutag, 2022; Bültermann, 2022). This changed speech style can be interpreted as an effort to make assistants appear more human and natural in their interactions with users. Though a colloquial speech style is assumed to support feelings of humanness and intimacy and therefore the building of trust (Grünenfelder et al., 2021), the effects of the use of colloquialism by voice assistants have not been investigated yet. The following study uses an experimental online survey to examine the indirect effects of colloquial speech by a fictitious voice assistant on users' trust and intentions to use, compared to formal language, via the variables of humanness and competence. Considering that voice assistants become more relevant for information purposes, the examination of effects of human features in digital agents prove to be of great importance. Credibility of voice assistants might be achieved via communication style and be transferred to the contents of their answers

which do not hold up to journalistic standards (Frehmann et al., 2022).

Voice assistants and information seeking

Voice assistants such as Amazon Alexa, Apple's Siri and Google Assistant are software programs in form of digital agents on both mobile devices and stationary smart speakers. These assistants use natural language processing and machine learning algorithms and can therefore be seen as a form of artificial intelligence. They understand commands in spoken language and execute them while responding verbally (Hepp 2020; Hoy 2018). These commands can range from simple requests such as setting an alarm to more complex tasks such as controlling smart home devices or asking for factual and current information (Arnold et al., 2019). The dissemination and use of voice assistants, for example in form of smart speakers, has been growing internationally over the last years (Newman et al., 2020). In the UK, almost a quarter of the population (24%) owns a smart speaker (Newman et al., 2022). Though smart speakers are less common in Germany, with only 15 percent of Germans owning one (Newman et al., 2022), the majority (85%) owns at least one device that has a voice assistant installed, such as smartphones, laptops, tablets or speakers (Arnold et al., 2019). In 2020, about 30 percent of German online users had used a voice assistant before and in younger groups a percentage of 58 percent did so (ARD & ZDF, 2020).

Voice assistants' relevance for society and communication science has been growing recently due to a new development in the usage scenarios: Voice assistants are increasingly being used as information sources (Newman et al., 2019, 2020). According to different studies, during the pandemic, people used them to stay updated about general news and information, as well as a source for specific facts about the current crisis (Brocks, 2020; NPR & Edison Research, 2020; Strathmann et al., 2022). In the first year of the pandemic, from mid to end of 2020, up to 20 percent of German adults asked a voice assistant at least once per week about the current situation (Viehmann et al., 2022). However, content analyses of typical answers of voice assistants to questions concerning the pandemic showed information of fluctuating quality: Although no false information was detected, voice assistants often could not answer a question or gave incomplete information on a topic (Frehmann et al., 2022). Furthermore, they did not

always disclose their sources (Frehmann et al., 2022, Goh et al., 2021). Similar results had been found before regarding different health-centered topics (e.g., Miner et al., 2016).

By selecting, processing, and presenting certain information, voice assistants act as new information intermediaries and gatekeepers (Frehmann et al., 2022). Considering this new role and the accompanying influence on users, the observed lack of precision and transparency is concerning. Additionally, verification of their statements is difficult because the assistants do not provide any insights into how they find and compile information (Natale & Cooke, 2021). If people base their knowledge and opinions on facts retrieved from voice search, they are at risk of believing in incomplete, untrustworthy or, in the worst case, wrong information. To date, it is unclear why people seem to rely on answers given by assistants whose ways of working they cannot observe. A possible explanation for the use of voice assistants despite this lack of transparency could be the concept of trust.

Trust in voice assistants

Trust is an important component of everyday life. It takes effect whenever a person expects the performance of another party without being able to control their actions. By trusting and relying on the other party called "trustee", the person called "trustor" accepts their own vulnerability and receives a reduction of uncertainty and social complexity in turn. This makes them more willing to take the risk of trusting the trustee (Luhmann, 2014; Mayer et al., 1995). Several studies have identified three central determinants of trust, also known as "trusting beliefs": The trustor should perceive the trustee as competent, meaning capable of meeting the expectations of the trustor. Furthermore, a trustee should appear benevolent, that is showing positive intentions towards the trustor and having their best interest in mind. Lastly, the integrity of the trustee should show a set of shared values with the trustor that justifies the perception of reliability (Mayer et al., 1995; McKnight et al., 1998). Though these dimensions have hence been adapted to better fit technological trustees, they remain relevant for human-seeming applications (Lankton et al., 2015; McKnight et al., 2011). If trust is given, trusting intentions and behaviors can be the consequences of the interaction. In the case of technology being the trustee, trust mainly leads to intentions and behaviors concerning the (continued) use of this technology (e.g., Lankton et al.,

2015; Zhou, 2013).

Applying this definition to voice assistants, trust could be an explanation as to why users turn to them for information despite their lack of transparency. Trusting the assistants and their work to be competent, benevolent, and having integrity could balance out the uncertainty about how and why they chose the given information and reduce the perception of risk. Dealing with the complexity of questioning the methods of the assistants can be avoided by simply trusting that they are capable and do their job well. Therefore, continued trust in a voice assistant may explain its ongoing use.

Studies on voice assistant use support these assumptions. In accordance with research concerning other technologies, trust has already been shown to positively affect the use (intentions) of voice assistants (e.g., Chérif & Lemoine, 2019; Nasirian et al., 2017, Scheuer, 2020). Nevertheless, how and why trust in voice assistants is formed, especially when they act as information intermediaries, is still not fully investigated. However, a recent study does confirm that the three dimensions of trust, that is competence, benevolence, and integrity, influence trust in voice assistants as well (Weidmüller, 2022). Additionally, the perception of voice assistants as an anthropomorphic and social actor, opposed to an inanimate object, is of relevance and increases trust (Weidmüller et al. 2022).

Voice assistants as social actors

Voice assistants are designed to be anthropomorphic, that is human-seeming. They often have human names like "Alexa" or "Siri" and use voices that resemble human speech (Wagner & Schramm-Klein, 2019). The effects of such anthropomorphism in technology have been studied in context of the "Computers as social actors" paradigm (CASA) in the field of Human-Computer Interaction (HCI) and Human-Machine Communication (HMC). CASA postulates that people behave socially towards technology that possesses human features or behaves in a human way. These features are interpreted as social cues and influence the perception of technology as social actors (Nass & Brave, 2005; Nass et al., 1994). Communicating via voice rather than text can already trigger this social perception of "humanness" in technology, as well as a human voice rather than a synthetic one (Chérif & Leomine, 2019; Cho et al., 2019). The consequences of the perception of technologies as human and social concern cognitions, affections and behaviors of users: Human-seeming agents are evaluated better than non-

human-seeming ones, they have greater influence on users' attitudes and evoke the appliance of social rules and prejudices, for example, politeness and gender stereotypes (Nass & Brave, 2005; Nass et al., 1994). For voice assistants, different effects of perceived humanness have been found: Simple cues like a human name in an assistant can impact how persuasive its statements appear (Voorveld & Araujo, 2020). If an assistant shows social support for its user and discloses "personal" facts about itself, users can develop feelings of friendship towards them (Ki et al., 2020). By and large, perceived humanness of an application increases trust in different technologies (Lee & Nass, 2010; Rheu et al., 2021) as well as in voice assistants specifically which in turn impacts the intentions to use it positively (Chérif & Leomine, 2019; Cho et al., 2019).

Voice assistants are constantly being updated in hard- and software and further developed in their design. Current advances in the anthropomorphic appearance of different assistants include, for example, that Apple added new voice options for different male or female voices that sound more natural than the original female Siri voice (Panzarino, 2021). Amazon now offers a new creative function for its assistant Alexa who tells co-created stories about characters that users imagine together with the assistant (Amazon, 2022). The progress in anthropomorphic design does not only affect the voice itself or the contents of the answer, but also the style the assistants answer in: When encountered with personal questions and statements, for example, "I love you", Siri has long been answering with informal expressions such as "Oh, I bet you say that to all your Apple products" (Bellegarda, 2014). Recently, an increasing use of dialect and colloquial expressions can be observed in various assistants. Even typical commands and factual questions evoke colloquial answers using modern expressions and informal idioms (Blutag, 2022; Bültermann, 2022). This colloquialism seems to be another effort to humanize the assistants and align their speech style with that of their users. It could then strengthen the perception of humanness in voice assistants and increase intimacy and trust between the user and the assistant through the informal interaction (Grünenfelder et al., 2021). So far, however, if and how colloquialism affects the perceived humanness of voice assistants is unknown.

Speech Style and its effects on trust

When it comes to perceived humanness due to anthropomorphism and social cues in voice assistants, especially in their *voices*, most studies compare either different modalities like text versus voice (e.g., Cho et al., 2019), different levels of authenticity of the voice itself (e.g., Chérif & Lemoine, 2019) or different genders of the voice (e.g., Tolmeijer et al., 2021). Up until now, the effects of different speech styles a voice assistant can use, namely a colloquial or a formal one, have not been studied. It can be assumed that colloquial language, when compared directly with formal language, is interpreted as a stronger sign of humanness: Humans regularly use colloquial language when conversing with each other, while technology such as voice assistants only uses "human language" when conversing with humans and is just beginning to integrate colloquialism into its speech style (Blutag, 2022). According to various CASA studies, the perception of humanness in technology is associated with increased trust in technology (e.g., Lee & Nass, 2010; Rheu et al., 2021). In fact, the modality of voice compared to text can already increase perceived humanness of a technology and, through that, indirectly increase trust in technology (Cho et al., 2019). Designing the voice modality to be even more anthropomorphic, in our case via a colloquial speech style, should increase trust via perceived humanness further. Transferring these assumptions, we hypothesize:

H1a: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will increase people's perceived humanness of the assistant.

H1b: There is a positive correlation between the degree of perceived humanness of a voice assistant and people's trust in it.

H1c: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will indirectly increase people's trust in the voice assistant, mediated via an increase of the assistant's perceived humanness.

However, taking studies concerning the effects of speech style of humans into consideration, colloquial speech could also be detrimental to the evaluation of a speaker: A meta-analysis found that humans who do not communicate in "standard" language but rather in dialect and non-standard accents are evaluated as of lesser status, as less lively and as less trustworthy (Fuertes et al., 2012; see also Vögele & Bachl, 2017).

This phenomenon could be explained by the observation that listeners attribute less competence to individuals speaking in accents or dialect than to individuals speaking standard language, partly due to stereotypes and prejudices (e.g., Blair & Connor, 1978; Cuddy et al., 2008). This observation might be transferable to colloquialism which also differs from standard formal language and might evoke prejudice in listeners as well. Furthermore, perceived competence is a strong determinant of trust in human-like technologies such as voice assistants (Lankton et al., 2015; McKnight et al., 2011; Weidmüller, 2022), meaning that a negative evaluation of competence of voice assistants would also negatively impact trust. Applied to this study, we hypothesize:

H2a: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will reduce people's perceived competence of the assistant.

H2b: There is a positive correlation between the degree of perceived competence of a voice assistant and people's trust in it.

H2c: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will indirectly decrease people's trust in the voice assistant, mediated via a decrease of the assistant's perceived competence.

Trust is a relevant predictor when it comes to intentions and behaviors of the trustor. In the case of technology being the trustee, trust determines intentions and behaviors concerning the (continued) use of it. The effect of trust in a technology on intentions to use it has been proven for different applications (e.g.; Lankton et al., 2015; Zhou, 2013) as well as voice assistants (e.g., Chérif & Lemoine, 2019). If trust is indirectly influenced by speech style as assumed in H1c and H2c, speech style will also have an indirect effect on intentions to use the voice assistant, mediated via trust. We therefore hypothesize:

H3a: Trust in the voice assistant will be positively related to people's intentions to use the assistant.

H3b: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will indirectly increase people's intentions to use the voice assistant, serially mediated via an increase of the assistant's perceived humanness and trust.

H3c: Compared to a voice assistant that uses a formal speech style, a colloquial speech style will indirectly

decrease people's intentions to use the voice assistant, serially mediated via a decrease of the assistant's perceived competence and trust.

Method

Design and participants

To test the hypotheses, we conducted a vignette-based experimental online survey in a 2 (speech style: formal language vs. colloquial language) x 2 (context: informal everyday issues vs. formal political issues) between-subjects design in July 2021. The manipulation of context was included to assess the generalizability of the findings across two different topics. For this purpose, we computed single-group models (see Results section). A vignette-design is a demonstrational approach typical for research in the field of HMC and describes a scenario with the help of prerecorded materials which participants imagine themselves in. It is especially useful for manipulating specific features of a digital agent while keeping all other conditions equal. It is also the most manageable version of an experiment for larger online samples (Greussing et al., 2022). The created scenarios described an interaction between a newly developed voice assistant and the participant as a user. Either the context was an informal everyday interaction on the weather and a musician, or the context was more formal, containing questions on the political program of three German parties. The voice assistant then answered either in a colloquial or in a formal speech style. The answers were prerecorded using the freely available demo version of the text-to-speech solution *Readspeaker* (<https://www.readspeaker.com/de/>). We adjusted the frequency of the voice in the program *Audacity*. Thereby we tried to create a genderless and unknown voice to avoid any effects of prior attitudes or stereotypes. Participants were asked to evaluate the gender of the voice in the main questionnaire (see Discussion section).

We recruited the participants for the study via the contacts of students of an undergraduate course at Heinrich-Heine-University Dusseldorf, Germany and the German commercial online access panel *Respondi*. A sample of 275 completed questionnaires was obtained. Participants that spent too much or not enough time on the stimulus pages ($n = 17$), failed quality checks ($n = 9$) or had technical issues ($n = 5$) were excluded. The final sample consisted of 244 participants with an average age of 42 years. 50 percent identified as male, 49 percent as female and 1 percent

as diverse. Regarding education, 31 percent of participants presented a low education from no degree up to a middle school degree, 32 percent held a high school diploma and 35 percent a university degree.

Procedure and stimulus

Participants first answered general questions on demographics and use of and attitudes towards voice assistants. Afterwards they were asked to imagine an interaction with a supposedly new voice assistant named "Q" which is being developed but, at the current time, cannot interact with users directly. After randomization by age, gender and education, participants were sorted into one of the four experimental groups. Depending on their assigned group, participants were prompted to imagine asking the voice assistant either three pre-formulated questions on the topic of everyday life (e.g., "What will the weather be like?") or politics (e.g., "What is the name of the Green Party candidate for chancellor?") which were written out on screen (see Table 1 in Appendix for all questions). After each question, participants could access the voice assistant's answer via a clickable audio file. Either the answers were provided in a colloquial speech style or a formal speech style. After hitting the play button and listening to all three answers, participants were asked to evaluate Q and the imagined interaction.

We manipulated the speech style of the voice assistant in each answer in used idioms and wordings. The online website of the German spelling dictionary "Duden" provided some colloquial alternatives to formal expressions that we expanded on. For example, the German equivalent of the formal "raining heavily" was replaced by the informal "raining buckets" (see Table 1 in Appendix for answers in both speech styles). The speech style was successfully tested in a pretest ($N = 30$) via a colloquialism scale containing the items "conventional", "formal", "everyday language", "casual", "colloquial", "informal", "educational language" 1 = "Strongly disagree", 5 = "Strongly agree", $\alpha = 0.87$). Participants interpreted the speech style as intended ($M_{\text{formal}} = 2.51$, $M_{\text{colloquial}} = 3.64$, $t(28) = -4.680$, $p < .001$). All audio files of Q's answers (in German) and the written-out stimulus texts both in German and English are also available in an OSF repository: <https://osf.io/bvwe6/>

Measures

Dependent variables were measured using items and scales of previous work in the HMC and HCI field (for items, cronbach's alpha, CFA, general means and means per experimental group including CI see Table 2): We measured perceived humanness on a semantic differential scale in accordance to Ho and MacDorman (2010). Participants had to evaluate four contrary attribute-pairs differentiating between human and technical characteristics on a seven-point scale. For competence, we chose four reoccurring items in several studies of the HCI/HMC-field revolving around the ability to fulfil given tasks (Burgoon et al., 2000; Carolus et al., 2019; Lee, 2003). Participants evaluated the competence items on seven-point-scales (1 = "Strongly disagree", 7 = "Strongly agree") which were later compiled to a competence index. Similarly, trust was assessed through five items regarding trustworthiness and reliability (Burgoon et al., 2000; Lee, 2003) later compiled to an index. For intentions to use, we contrived four typical items that were tailored to our scenario of the newly developed voice assistant Q.

Additional control variables were measured before the stimulus was shown. These variables included gender, age and education as well as the use frequency of voice assistants (1 = "never", 2 = "seldomly", 3 = "at least once per month", 4 = "at least once per week", 5 = "at least once per day") and, in accordance to Arnold et al. (2019), prior attitudes regarding voice assistants on a seven-point-scale ("user-friendly", "time-saving", "useful", "independent", "needs-centered", "innovative", $\alpha = .83$, $M = 4.35$, 95% CI [4.20; 4.49]). The treatment check was conducted via the same items used in the pretest to measure colloquialism on a seven-point-scale ("conventional", "formal", "everyday language", "casual", "colloquial", "informal", "educational language", $\alpha = 0.87$, $M = 4.08$, 95% CI [3.94; 4.23]).

Table 2: Summary of scale for humanness, competence, trust and intentions to use and individual items

	α	CFA	M	M _{form} [95% CI]	M _{coll} [95% CI]
1. Humanness	.94		2.99	2.99 [2.74; 3.24]	2.99 [2.73; 3.24]
H1: artificial – natural	.91	.91	2.93	2.97 [2.70; 3.24]	2.88 [2.61; 3.15]
H2: human-made – human-like	.89	.89	2.89	2.92 [2.64; 3.20]	2.86 [2.59; 3.13]
H3: inanimate – living	.82	.82	3.14	3.09 [2.82; 3.36]	3.19 [2.89; 3.49]
H4: synthetic – real	.92	.92	3.00	2.98 [2.71; 3.26]	3.01 [2.73; 3.28]
2. Competence	.93		4.43	4.63 [4.37; 4.88]	4.22 [3.95; 4.49]
C1: competent	.91	.91	4.67	4.85 [4.56; 5.15]	4.48 [4.16; 4.80]
C2: intelligent	.87	.87	4.22	4.39 [4.08; 4.70]	4.05 [3.73; 4.37]
C3: experienced	.85	.85	4.04	4.20 [3.91; 4.48]	3.88 [3.58; 4.17]
C4: professional	.85	.85	4.49	4.85 [4.55; 5.15]	4.12 [3.80; 4.43]
3. Trust	.92		4.20	4.31 [4.07; 4.55]	4.07 [3.80; 4.35]
T1: trustworthy	.84	.84	3.99	4.16 [3.88; 4.45]	3.81 [3.49; 4.13]
T2: reliable	.85	.85	4.44	4.53 [4.26; 4.8]	4.35 [4.04; 4.66]
T3: responsible	.80	.80	3.72	3.83 [3.54; 4.12]	3.60 [3.31; 3.89]
T4: sincere	.80	.80	4.20	4.21 [3.92; 4.50]	4.18 [3.88; 4.49]
T5: credible	.88	.88	4.64	4.84 [4.57; 5.10]	4.43 [4.10; 4.48]
4. Intentions to use	.90		2.84	2.75 [2.50; 3.01]	2.94 [2.63; 3.23]
U1: I could imagine using Q.	.92	.92	2.90	2.93 [2.60; 3.26]	2.87 [2.54; 3.19]
U2: I would like to have a "real" conversation with Q.	.68	.68	2.93	2.64 [2.30; 2.99]	3.22 [2.85; 3.60]

Notes: $n = 244$, individual items measured on seven-point-scales, 1 = "Strongly disagree", 7 = "Strongly agree"

Table 3: Zero-order correlations

	1	2	3	4
1. Perceived humanness	1			
2. Competence	.476**	1		
3. Trust	.501**	.786**	1	
4. Intentions to use	.577**	.569**	.565**	1

Notes: $n = 244$, ** $p < .01$, scales compiled of items measured on seven-point-scales, 1 = "Strongly disagree", 7 = "Strongly agree"

Data analysis and results

Randomization and treatment check

The four experimental groups were randomized by age, gender, and education and are roughly of the same size ($n_{\text{everyday/colloquial}} = 56$, $n_{\text{everyday/formal}} = 65$, $n_{\text{politics/colloquial}} = 65$, $n_{\text{politics/formal}} = 56$). No significant differences between the groups were found regarding the sociodemographic variables age ($F(3,240) = 0.31$, $p = .816$), gender ($\chi^2 = 1.62$, $df = 3$, $p = .655$), education ($\chi^2 = 7.84$, $df = 9$, $p = .55$) nor in variables of use frequencies ($\chi^2 = 11.53$, $df = 12$, $p = .484$) or prior attitudes ($F(3,240) = 0.53$, $p = .662$) regarding voice assistants. The treatment check was successful, participants that listened to the colloquial assistant perceived it as more colloquial than the formal speaking one ($t(220.541) = -7.84$, $p < .001$, $M_{\text{colloquial}} = 4.62$, $M_{\text{formal}} = 3.56$).

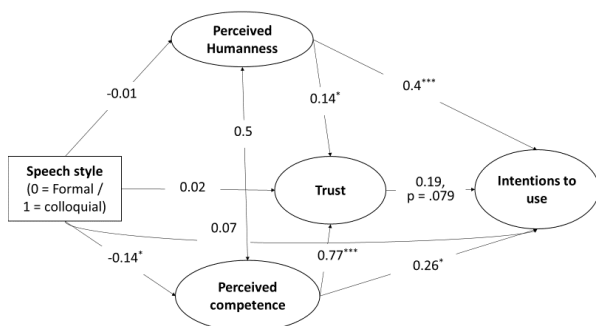
Data Analysis

To analyze the data, we applied a structural equation model (SEM) using AMOS 28. The experimental manipulation of the *speech style* of Q was used as an indicator, and the modelled latent variables of perceptions of *humanness* and *competence* as mediators. Perceived *trust* and *intentions to use* of Q acted as outcomes. We also added the correlation between perceived humanness and competence (Table 2). Indirect effects were assessed via bootstrapping ($m = 5,000$, bias-corrected 95% confidence intervals).

For the single-group analyses, context (formal political vs. informal everyday issues) was used as a grouping variable. Since there were no significant differences between the experimental groups regarding our measured control variables, we did not include them in the model. The model fit the data well: $\chi^2 (126) = 242.58$, $df = 126$, $p = .000$, $\chi^2/df = 1.93$, CFI = .968, RMSEA = .062, SRMR = .042 (Hu & Bentler, 1999). Figure 1 shows the structural equation model.

Hypothesis Testing

Figure 1: Structural equation model of effects of speech style manipulation



Notes: $n = 244$, maximum likelihood estimation, calculated with AMOS 28, $\chi^2(126) = 242.58$, $df = 126$, $p = .000$, $\chi^2/df = 1.93$, $CFI = .968$, $RMSEA = .062$, $SRMR = .042$; * $p > .05$, ** $p < .01$, *** $p < .001$, total effect of speech style on trust: $B = -0.265$; $p = .169$; 95% CI [-0.660; 0.099], total effect of speech style on intentions to use: $B = 0.048$; $p = .856$; 95% CI [-0.407; 0.495]

Hypotheses 1a-c concerned the relationship of speech style, perceived humanness and trust. Contrary to our hypothesis, colloquial speech did not influence perceived humanness ($\beta = -.006$, $p = .926$), H1a had to be rejected. Although perceived humanness correlated positively with participants' trust in the voice assistant ($\beta = .143$, $p = .005$) and H1b was accepted, no mediation of speech style via humanness on trust was observed ($B = -.003$, $p = .881$, 95% CI [-0.061; 0.057]). Hence, H1c had to be rejected as well.

Hypotheses 2a-c predicted an influence of speech style on trust via perceived competence. In line with H2a, a colloquial speech style negatively affected the voice assistant's perceived competence ($\beta = -.143$, $p = .033$). When Q used a colloquial speech style, it was perceived as less competent than when it used a formal speech style. Competence, in turn, was positively related to participants' trust in the voice assistant ($\beta = .768$, $p < .001$) as H2b assumed. The indirect effect of speech style on trust via competence was significant and negative ($B = -.315$, $p = .034$, 95% CI [-0.600; -0.025]). H2a-c are therefore accepted.

H3a-c assumed a mediation of speech style via humanness or competence and trust on intentions to use. The model showed only an effect of speech style on competence, not on humanness (see H1a), which is

why H3b is rejected. A positive correlation was observed between trust and intentions to use, which, however, slightly missed the conventional level of significance ($\beta = .191$, $p = .079$). Similarly, the indirect effect of speech style on intentions to use via competence and trust only approached the conventional level of significance ($B = -.070$, $p = .085$; 95% CI [-0.261; 0.008]). Still, H3a and H3c are rejected. Interestingly, intentions to use the voice assistant correlated positively both with perceived competence ($\beta = .263$, $p = .032$) and humanness of the assistant ($\beta = .404$, $p = .001$).

Regarding the direct effects of a colloquial speech style, no such effects were observed on participants' trust in the assistant ($\beta = .019$, $p = .659$) and on participants' intentions to use it ($\beta = .072$, $p = .137$). There were also no significant total effects of the colloquial speech style on participants' trust ($B = -0.265$; $p = .169$; 95% CI [-0.660; 0.099]) and on their intentions to use ($B = 0.048$; $p = .856$; 95% CI [-0.407; 0.495]).

Generalizability across two topics

To assess the generalizability of our findings, we computed single grouped SEMs for each level of the context variable, with the "formal political" versus "informal everyday" topic conditions as a grouping variable. We then restricted the models by iteratively setting equal the strength of one path of one model with the respective path of the other model. A moderation effect is significant when this restriction significantly impedes the quality of the models (measured by chi-square tests of difference) because then, it cannot be assumed that the path strengths are equal for both groups (Byrne, 2010). Table 2 shows the results of this analysis. As can be seen, the models in which single paths were set equal as well as the fully restricted model did not differ significantly from the unrestricted model. We therefore conclude that the effects and relations reported in the previous section occur across different topics.

Table 4: Moderation effects of context on the effects and correlations assumed

	Path(s) set equal	χ^2 (df)	χ^2_{diff}	<i>p</i>
Unrestricted model		427.945 (252)		
H1a	coll → hum	427.945 (253)	0.00	1.000
H1b	hum → trust	428.106 (253)	0.161	0.688
H2a	coll → comp	428.563 (253)	0.618	0.432
H2b	comp → trust	429.151 (253)	1.206	0.272
H3a	trust → use	428.126 (253)	0.181	0.671
Direct 1	coll → trust	428.393 (253)	0.448	0.503
Direct 2	coll → use	429.693 (253)	1.748	0.186
Direct 3	hum → use	427.964 (253)	0.019	0.890
Direct 4	comp → use	428.280 (253)	0.335	0.563
Fully restricted model	all the above	432.533 (261)	4.588	0.869

Notes: coll = colloquial speech style; hum = perceived humanness; trust = trust in the assistant; comp = perceived competence; use = intentions to use

Discussion and limitations

In this paper, we reacted on current observations of colloquialism in the speech style of voice assistants and checked its possible effects on trust and intentions to use via mediations of humanness and competence. Colloquialism could act as a human-seeming feature and trigger perceived humanness which, according to research concerning technology and the CASA paradigm, can positively influence trust. However, according to research regarding human speech, dialect and accents evoke a negative influence on competence and later trust in comparison to standard language. Colloquialism could have a similar effect. Trust in turn usually has a positive influence on intentions to use and might be indirectly influenced by speech style. The findings of our experimental online study only confirmed that colloquial speech reduced perceived competence which mediated a negative effect on trust. No further significant effects were found in a structural equation model (SEM).

Even though the SEM revealed no total effect of speech style on trust or intentions to use, the findings still contain interesting insights into the perception of voice assistants: The current study did not find support for the assumptions based on the CASA paradigm. H1a and H1c and consequently H3b had to be rejected, because no influence of colloquialism was found on perceptions of humanness. This observation might point to a problem with our small sample and consequently low power in the mediation analysis which we discuss as a limitation (see below).

Nevertheless, this came as a surprise, as informal expressions and idioms are a characteristic way of human communication. Possibly, the expressions in the stimulus material either were not strong enough in their colloquialism to have an effect on humanness or trying too hard to feel like “natural” human speech. However, the treatment check confirmed that colloquialism was perceived as intended in comparison to the formal speech style. An alternative explanation might be that an intervening factor prevented any effects of colloquialism on humanness: Possibly, the used artificial and genderless voice of Q felt too unfamiliar for participants and deviated too much from usual voices of current assistants to be perceived as human. Checking with the data, the ambiguous gender of Q did not seem to confuse people too much: Only 17 percent of all participants were not able to assign Q a gender, while 29 percent felt the voice was male and 54 percent perceived it as female. Nevertheless, some participants communicated their frustration with the voice of Q via a general comment field at the end of the questionnaire, describing its voice as “choppy sounding”, “monotonous” and “slow-motion”, “annoying” and “unnatural” which made one participant “aggressive”. This dissatisfaction with Q’s voice could have been reflected by participants in the evaluation of humanness of Q. It is striking that the mode of the variable is “1”: 15 percent of all participants perceived Q as clearly artificial. To check for possible floor-effects and ask for perceptions of humanness on a more finely graded measurement, we also included a scale of humanness-evaluation with a slider between 1 = “thing” and 100 = “person” (Etzrodt et al., 2021). Again, we found the clustering trend on the left extreme point: 16 percent of participants clearly stated that they perceived Q as a definite “thing” by answering with the extreme value and mode of “1”, and even more than half (51%) placed the slider in the first quartile of this scale (M = 32,2, SD = 27,51). It seems the use of the freely available voice that was additionally edited was a poor decision to measure effects of speech style on the perceived humanness, because Q and its voice were not perceived as pleasant and human-like to begin with and were therefore not susceptible to an influence by speech style. Furthermore, the frequent use of the extreme scale points could be a sign of reactance or even defiance in participants: In the current time of fake news and a growing body of convincing artificial agents and chatbots (Verma, 2022), participants may have wanted to clearly show that they can identify artificial entities and are not fooled by designs that are meant to be anthropomorphic.

H2a-c, however, were accepted, meaning that the assumptions based on prior studies regarding the perception of human speech styles were confirmed. A colloquial speaking voice assistant is perceived as less competent (H2a), competence and trust are positively related (H2b) and therefore, colloquial speech has a negative indirect effect on trust via competence (H2c). This observation confirms two presuppositions: One, colloquialism and its effects on the evaluation of a speaker are similar to those of dialect and accents. Prior studies concerned themselves with various degrees of non-standard-language but mainly focused on accents and dialects and not exclusively on colloquialism (e.g., Fuertes et al., 2012). The current study now shows that colloquialism suffices to reduce interpersonal evaluations as accents and dialects do. Two, prior findings on *human* speakers are transferable on an *artificial* speaker, in this case a voice assistant which could be a subtle proof for the CASA paradigm, after all. The social rules of human prejudice due to non-standard language and their effects on evaluations of human speakers were applied to a digital agent. This social treatment of the assistant was possibly triggered due to perceptions of humanness in the assistant on a subconscious level which could not be measured by the questionnaire (see above). It remains uncertain whether this perception of humanness was caused by the colloquialism or the assistant itself. Several features of Q that were not experimentally manipulated could have had this influence, for example the ability to speak in a voice, the knowledge on topics, the level of interactivity and so on.

Finally, the variable of intentions to use offers some new insights, though the corresponding results slightly missed the conventional level of significance and H3a and H3c were hence rejected. Possibly, in a bigger sample containing more statistical power, a significant direct effect of trust as well as an indirect effect of speech style via competence and trust on intentions to use could be found that we only observed by trend. Interestingly though, significant direct relations of intentions to use with competence and humanness were found. Competence is often considered a subdimension of trust or a "trusting belief" that, in combination with benevolence and integrity, forms the concept of trust as a whole (e.g., Lankton et al., 2015; Weidmüller 2022). The findings of this study confirm that competence can clearly be seen as a closely related concept to trust, as they are strongly correlated ($r = .786, p < .01$) and the different relationships with intentions to use and speech style show that competence is not completely equal to trust. It is more of a subdimension or

determinant of it and is additionally more relevant to intentions to use than the broad concept of trust. Further, perceptions of humanness had a direct positive relation with intentions to use. This relation could be a sign of users' shifting demands when it comes to information requests: Information intermediation by voice assistants might be attractive for the very reason of interactive and spoken out answers that mirror human conversations. Hearing the answer to a specific question in a friendly and personal voice could be preferable compared to general, impersonal news content in "classical" media. Thus, the process of information retrieval becomes more dialogic, personalized and individual. Therefore, humanness in an information intermediary might not only increase trust in them but be also a key selling point for information search via voice assistants and influence intentions to use directly.

This study has to be interpreted in light of a few limitations: First, some expected effects were not found or only observed as a trend that slightly missed the significance level which could be due to our sample: We did not perform an a-priori-power analysis to assess the needed sample size to find expected effects. Therefore, our sample might have been too small to obtain the needed power for a reliable detection of all existent effects and relations. Future studies should hence include such prior analyses and aim for larger samples to perform mediation analyses with more power. Furthermore, we recruited participants via student contacts and a commercial panel and cannot rule out that these groups had distorting effects on our results. Therefore, representative samples could be helpful to review our findings on a firmer basis. Especially the direct effects of humanness and competence on intentions to use had no connection to our experimental manipulation and should be verified for a representative sample.

Second, the used voice of Q has to be criticized. We purposely designed the voice of Q as gender-ambiguous and different to current options on the market to avoid prior attitudes regarding voice assistants. However, the sound and way of speaking of Q did not appeal to some participants as some comments revealed which might have had an intervening effect on the relation between colloquialism and perceived humanness. Future studies should implement more typical and smoothly speaking voices into experimental designs and check again for effects of colloquialism as proof for the CASA paradigm. One option is to record real voice assistants

like Alexa or Siri that can read out loud written stimulus texts. These studies have to control for attitudes and experiences concerning the specific assistants and their providers though. Another option is to use other highly developed artificial voices that are lesser-known, for example Microsoft's text-to-speech assistant Zira (Greussing et al., 2022; Yuan et al., 2019). This method allows to avoid intervening effects due to prior experiences.

Third, the vignette-design of the study may have had a central disadvantage. Participants were asked to imagine an interaction with a newly developed voice assistant but did not actually speak with the assistant. The lack of action from participants and real reactions of Q could have negatively influenced how interactive the situation was perceived. This missing feeling of interactivity could have had effects on measured attitudes towards the assistant. Future study designs should avoid this problem, for example by using stimulus videos from the first-person perspective (e.g., Whang & Im, 2021) or by switching to laboratory experiments to enhance external validity. A wizard-of-Oz-design is helpful to simulate an interaction between user and voice assistant in the lab while the given answers are actually controlled by a hidden researcher. In doing so, researchers can better control the experimental conditions and avoid confounding variables (Greussing et al., 2022).

Lastly, due to prior literature connecting the speech style of human speakers with competence and trust, we only focused on one of the three typical determinants of trust and omitted two further trusting beliefs. Additional measurements of benevolence and integrity as well as the determinants of technical trust – namely functionality, helpfulness and reliability (Lankton et al., 2015; McKnight et al., 2011) – could have been added to our questionnaire. Technical trust has recently been proven to be relevant to voice assistants as well as human trust (Weidmüller, 2022). Future studies should look into effects of colloquialism on these different determinants of trust and how they play together to build trust in voice assistant. Possible additional dimensions of trust specific to the information intermediation of voice assistants should be considered as well. When users turn to voice assistants for current information, aspects of their methods of information research and editing may also be relevant in forming trust. In fact, recent findings do point to the fact that the influence of anthropomorphic design on trust has been overestimated due to the undetected impact of credibility of the statements a voice assistant makes

(Weidmüller et al., 2022). Future studies should additionally include measures regarding the credibility of the given information.

Taken together, the current study showed that independent of the topic a voice assistant talks about, its colloquial speech style – compared to a formal one – has a negative effect on competence evaluations, which also mediates a negative effect of speech style on trust. In contrast, humanness is not influenced by a colloquial speech style. Based on these findings, a voice assistant will therefore not benefit from a colloquial speech style regarding the perceptions and behaviors of its users. However, future studies should carefully review these findings with a bigger sample containing more statistical power and use a more common voice for the assistant. Possibly the found harmful relation of colloquialism and trust via competency can be balanced out by an additional positive influence on trust via perceived humanness which we could not observe. Furthermore, the found direct relationship of perceived humanness and intentions to use the assistant reveals more research potential: If an anthropomorphic and human-seeming design encourages users to turn to voice assistants, research should look into the significance of this feature for the use of voice assistants, especially for information purposes. Considering the ambiguous quality of voice assistants' factual answers (e.g., Frehmann et al., 2022; Goh et al., 2021), there might be a risk that a natural human sounding voice masks flaws in factuality and reliability of the content. Future studies should determine specific motives to use voice assistants as an information source and investigate how users deal with answers, assess their credibility and whether they factcheck specific information which they received from an assistant.

Acknowledgements

The author would like to express her gratitude and appreciation to the students who co-designed this study in the context of an undergraduate course at Heinrich-Heine-University Duesseldorf as well as to her advisor Prof. Marc Ziegele who co-taught the course. Specifically, the efforts of students Anna Bäβ, Laurids Blume, Carolin Boot, Alexandra Hellmann, Isabel Hellmann, Sara Hermes, Leonard Kern-Wagner, Yasmin Kirstein, Anja Lukassek, Leonie Panusch, Laura Stapkevica and Anna Teichmann in designing the questionnaire, recruiting participants and executing the study should be recognized.

References

- Amazon (2022, September 28). Do more with these new Alexa features at home and on-the-go. *About Amazon*.
<https://www.aboutamazon.com/news/devices/alexa-news-2022>
- ARD & ZDF (2020). *Ergebnisse der ARD/ZDF-Onlinestudie 2020* [Results of the ARD/ZDF online study 2020]. https://www.ard-zdf-onlinestudie.de/files/2020/2020-10-12_Onlinestudie2020_Publikationschart_s.pdf
- Arnold, R., Taş, S., Hildebrandt, C., & Schneider, A. (2019). Any Sirious concerns yet? – An empirical analysis of voice assistants’ impact on consumer behavior and assessment of emerging policy challenges. *TPRC47: The 47th Research Conference on Communication, Information and Internet Policy 2019*. <https://doi.org/10.2139/ssrn.3426809>
- Bellegarda, J. R. (2014). Spoken language understanding for natural interaction: The siri experience. In J. Mariani, S. Rosset, M. Garnier-Rizet & L. Devillers (eds.), *Natural Interaction with Robots, Knowbots and Smartphones* (pp. 3-14). Springer. https://doi.org/10.1007/978-1-4614-8280-2_1
- Blair, L. M., & Connor, H. S. (1978). Black and rural accents found to lessen job opportunities. *Monthly Labor Review*, 101, 35–36.
- Blutag (2022, October 4). The personification of voice assistants. *Blutag*. <https://blu.ai/blog/the-personification-of-voice-assistants>
- Brocks, L. (2020, April 17). „Ok, Corona.“ – Welchen Einfluss das Virus auf die Nutzung von Voice in Deutschland hat [„Ok, Corona.“ – The influence of the virus on the use of voice in Germany]. *Let’s make lemonade*.
<https://lets-make-lemonade.de/ok-corona-welchen-einfluss-das-virus-auf-die-nutzung-und-nutzungsabsicht-von-voice-in-deutschland-hat/>
- Bültermann, S. (2022, November 24). Amazon Alexa kann nun mit Akzent sprechen [Amazon Alexa is able to speak with an accent now]. *Computerbild*. <https://www.computerbild.de/artikel/cb-News-Smart-Home-Alexa-kann-nun-mit-Akzent-sprechen-34874271.html>
- Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human–computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(6), 553–574.
[https://doi.org/10.1016/S0747-5632\(00\)00029-7](https://doi.org/10.1016/S0747-5632(00)00029-7)
- Byrne, B. M. (2010). *Structural equation modeling with Amos: Basic concepts, applications, and programming* (2nd ed.). Routledge/Taylor & Francis Group.
- Carolus, A., Muench, R., Schmidt, C., & Schneider, F. (2019). Impertinent mobiles - Effects of politeness and impoliteness in human-smartphone interaction. *Computers in Human Behavior*, 93(3), 290–300.
<https://doi.org/10.1016/j.chb.2018.12.030>
- Chérif, E., & Lemoine, J.-F. (2019). Anthropomorphic virtual assistants and the reactions of Internet users: An experiment on the assistant’s voice. *Recherche Et Applications En Marketing (English Edition)*, 34(1), 28–47. <https://doi.org/10.1177/2051570719829432>
- Cho, E., Molina, M. D., & Wang, J. (2019). The effects of modality, device, and task differences on perceived human likeness of voice-activated virtual assistants. *Cyberpsychology, Behavior and Social Networking*, 22(8), 515–520.
<https://doi.org/10.1089/cyber.2018.0571>
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal trait dimensions of interpersonal and intergroup perceptions: The stereotype content model and the BIAS map. In M. P. Zanna (ed.), *Advances in experimental psychology* (Vol. 40, pp. 61–149). Academic Press.
<https://doi.org/10.1016/S0065-2601%2807%2900002-0>
- Easwara Moorthy, A., & Vu, K.-P. L. (2015). Privacy concerns for use of voice activated personal assistant in the public space. *International Journal of Human-Computer Interaction*, 31(4), 307–335.
<https://doi.org/10.1080/10447318.2014.986642>
- Etzrodt, K. (2021). The ontological classification of conversational agents. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, E. Luger, M. Goodwin & P. B. Brandtzaeg (eds.) *Chatbot Research and Design* (pp. 48–63). Springer, Cham.
https://doi.org/10.1007/978-3-030-68288-0_4

- Ford, M., & Palmer, W. (2019). Alexa, are you listening to me? An analysis of Alexa voice service network traffic. *Personal and Ubiquitous Computing*, 23(1), 67–79. <https://doi.org/10.1007/s00779-018-1174-x>
- Frehmann, K., Ziegele, M., & Rosar, U. (2022). "Alexa, Siri, Google, what do you know about corona?" A quantitative survey of voice assistants and content analysis of their answers on questions about the COVID-19 pandemic. *Studies in Communication and Media*, 11(2), 278–303. <https://doi.org/10.5771/2192-4007-2022-2-278>
- Fuentes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*, 42(1), 120–133. <https://doi.org/10.1002/ejsp.862>
- Goh, A. S. Y., Wong, L. L., & Yap, K. Y.-L. (2021). Evaluation of COVID-19 information provided by digital voice assistants. *International Journal of Digital Health*, 1(1), Article 3. <http://doi.org/10.29337/ijdh.25>
- Greussing, E., Gaiser, F., Klein, S., Straßmann, C., Ischen, C., Eimler, S., Frehmann, K., Gieselmann, M., Knorr, C., Henestrosa, A. L., Räder, A., & Utz, S. (2022). Researching interactions between humans and machines: methodological challenges. *Publizistik*, 67(5), 531–554. <https://doi.org/10.1007/s11616-022-00759-3>
- Grünenfelder, J. I., Zierau, N., & Janson, A. (2021). Alexa, are you still there? Understanding the habitual use of AI-based voice assistants. *ICIS: Forty-Second International Conference on Information Systems, Austin 2021*.
- Hepp, A. (2020). Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies. *Media, Culture & Society*, 42(7–8), 1410–1426. <https://doi.org/10.1177/0163443720916412>
- Ho, C. C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, 26(6), 1508–1518. <https://doi.org/10.1016/j.chb.2010.05.015>
- Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, 37(1), 81–88. <https://doi.org/10.1080/19312458.2018.1506021>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Ki, C.-W., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Parafriendship development mechanism between IPAs and their users. *Computers in Human Behavior*, 111(4), Article 106412. <https://doi.org/10.1016/j.chb.2020.106412>
- Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, 16(10), 880–918. <https://doi.org/10.17705/1jais.00411>
- Lee, E. (2003). Effects of "gender" of the computer on informational social influence: The moderating role of task type. *International Journal Human-Computer Studies*, 58(4), 347–362. [https://doi.org/10.1016/S1071-5819\(03\)00009-0](https://doi.org/10.1016/S1071-5819(03)00009-0)
- Lee, J., & Nass, C. (2010). Trust in computers: The Computers-Are-Social-Actors (CASA) paradigm and trustworthiness perception in human-computer communication. In D. Latusek & A. Gerbasi (eds.), *Trust and technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives* (pp. 1–15). Information Science Reference. <https://doi.org/10.4018/978-1-61520-901-9.ch001>
- Liao, Y., Vitak, J., Kumar, P., Zimmer, M., & Kritikos, K. (2019). Understanding the role of privacy and trust in intelligent personal assistant adoption. In N. G. Taylor, C. Christian-Lamb, M. H. Martin, & B. Nardi (eds.), *Information in Contemporary Society* (pp. 102–113). Springer International Publishing. https://doi.org/10.1007/978-3-030-15742-5_9
- Luhmann, N. (2014). *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität [Trust: A mechanism to reduce social complexity]*. UVK.

- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709. <https://doi.org/10.2307/258792>
- McKnight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology. An investigation of its components and measures. *ACM Transactions on Management Information Systems*, 2(2), 1–25. <https://doi.org/10.1145/1985347.1985353>
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review*, 23(3), 473–490. <https://doi.org/10.5465/amr.1998.926622>
- Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Internal Medicine*, 176(5), 619–625. <https://doi.org/10.1001/jamainternmed.2016.0400>
- Nasirian, F., Ahmadian, M., & Lee, O. K. D. (2017). AI-based voice assistant systems: Evaluating from the interaction and trust perspectives. *Americas Conference on Information Systems*.
- Nass, C. I., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. MIT press.
- Nass, C. I., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 72–78).
- Natale, S., & Cooke, H. (2021). Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture & Society*, 43(6), 1000–1016. <https://doi.org/10.1177/0163443720983295>
- Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019). *Reuters Institute Digital News Report 2019*. Reuters Institute for the Study of Journalism, Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf
- Newman, N., Fletcher, R., Schulz, A., Andi, S., & Nielsen, R. K. (2020). *Reuters Institute Digital News Report 2020*. Reuters Institute for the Study of Journalism, Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf
- Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2021). *Reuters Institute Digital News Report 2021. 10th Edition*. Reuters Institute for the Study of Journalism, Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf
- Newman, N., Fletcher, R., Robertson, C. T., Eddy, K., & Nielsen, R. K. (2022). *Reuters Institute Digital News Report 2022*. Reuters Institute for the Study of Journalism, Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf
- NPR, & Edison Research (2020). *The smart audio research*. https://www.nationalpublicmedia.com/uploads/2020/04/The-Smart-Audio-Report_Spring-2020.pdf
- Pal, D., Arpikanondt, C., Razzaque, M. A., & Funilkul, S. (2020). To trust or not-trust: Privacy issues with voice assistants. *IT Professional*, 22(5), 46–53. <https://doi.org/10.1109/MITP.2019.2958914>
- Panzarino, M. (2021, March 31). Apple adds two brand new Siri voices and will no longer default to a female or male voice in iOS. *Techcrunch*. <https://techcrunch.com/2021/03/31/apple-adds-two-siri-voices/>
- Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human-Computer Interaction*, 37(1), 81–96. <https://doi.org/10.1080/10447318.2020.1807710>
- Scheuer, D. (2020). *Akzeptanz von Künstlicher Intelligenz: Grundlagen intelligenter KI-Assistenten und deren vertrauensvolle Nutzung [Acceptance of artificial intelligence: Foundations of intelligent AI assistants and their trust-based use]*. Springer. <https://doi.org/10.1007/978-3-658-29526-4>

- Strathmann, C., Szymczyk, N., Szczuka, J. M., & Krämer, N. C. (2021). Hey voice assistant, can you help me through the pandemic? A study on the use of voice assistants regarding loneliness, homeschooling and home office activities during the Covid-19 pandemic. In A. M. Rosenthal-von der Pütten, M. Elson, & S. Schiffer (eds.), *Proceedings of the 12th Media Psychology Conference 2021 (MediaPsych 2021)* (pp. 226–229). RWTH Aachen University. <https://doi.org/10.18154/RWTH-CONV-245969>
- Tolmeijer, S., Zierau, N., Janson, A., Wahdatehagh, J. S., Leimeister, J. M. M., & Berstein, A. (2021). Female by default? – Exploring the effect of voice assistant gender and pitch on trait and trust attribution. In Y. Kitamura, A. Quigley, K. Isbister & T. Igarashi (eds.), *CHI EA '21: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–7). <https://doi.org/10.1145/3411763.3451623>
- Verma, P. (2022, December 28). The year AI became eerily human. *The Washington Post*. <https://www.washingtonpost.com/technology/2022/12/28/ai-chatgpt-dalle-year-in-review/>
- Viehmann, C., Ziegele, M., & Quiring, O. (2022). *Informationsnutzung in der Corona-Krise. Report zu den Befunden der Panelbefragung. Befunde aus dem Juli & November 2020*. [Information use in the corona crisis. Report on the findings of the panel survey. Findings from July & November 2020]. <https://www.kowi.ifp.uni-mainz.de/juli-november-2020-befunde-aus-der-zwischenphase-und-zum-zweiten-lockdown/>
- Vögele, C., & Bachl, M. (2017). Der Einfluss des Dialekts auf die Bewertung von Politikern [The influence of dialect on evaluations of politicians]. *Studies in Communication and Media*, 6(2), 196–215. <http://doi.org/10.5771/2192-4007-2017-2-196>
- Voorveld, H. A. M., & Araujo, T. (2020). How social cues in virtual assistants influence concerns and persuasion: The role of voice and a human name. *Cyberpsychology, Behavior, and Social Networking*, 23(10), 689–696. <https://doi.org/10.1089/cyber.2019.0205>
- Wagner, K., & Schramm-Klein, H. (2019). Alexa, are you human? Investigating anthropomorphism of digital voice assistants – A qualitative approach. In H. Krcmar, J. Fedorowicz, W. F. Boh, J. M. Leimeister & S. Wattal (eds.), *Proceedings of the 40th International Conference on Information Systems, ICIS 2019, Munich*.
- Weidmüller, L. (2022). Human, hybrid, or machine? Exploring the trustworthiness of voice-based assistant. *Human-Machine Communication*, 4, 85–110. <https://doi.org/10.30658/hmc.4.5>
- Weidmüller, L., Etzrodt, K., & Engesser, S. Trustworthiness of voice-based assistants: integrating interlocutor and intermediary predictors. *Publizistik* 67, 625–651. <https://doi.org/10.1007/s11616-022-00763-7>
- Whang, C., & Im, H. (2021). “I like your suggestion!” The role of humanlikeness and parasocial relationship on the website versus voice shopper’s perception of recommendations. *Psychology & Marketing*, 38(4), 581–595. <https://doi.org/10.1002/mar.21437>
- Yuan, Y., Thompson, S., Watson, K., Chase, A., Senthilkumar, A., Bernheim Brush, A. J., & Yarosh, S. (2019). Speech interface reformulations and voice assistant personification preferences of children and parents. *International Journal of Child-Computer Interaction*, 21, 77–88. <https://doi.org/10.1016/j.ijcci.2019.04.005>
- Zhou, T. (2013). An empirical examination of continuance intention of mobile payment services. *Decision Support Systems*, 54(2), 1085–1091. <https://doi.org/10.1016/j.dss.2012.10.034>

Question in informal everyday context	Formal answer	Colloquial answer
What will the weather be like tonight?	Tonight, the temperature will be 10°C and it will be raining heavily. Tomorrow the weather will be better with 22°C and sunshine.	Tonight, the temperature will be 10°C and it will be <u>raining buckets</u> . Tomorrow the weather will be better with 22°C and sunshine.
What is the name of Amelia Taylor's new album?	Amelia Taylor's new album is called "Keep my heart around". It is selling very well, about 5 million copies to date.	Amelia Taylor's new <u>record</u> is called "Keep my heart around". It is selling <u>like hot cakes</u> , about 5 million copies to date.
How old is Amelia Taylor?	Amelia Taylor is 36 years old. She was born on 01.02.1985 in London. She learned to play the piano at the age of 5.	Amelia Taylor has 36 <u>on her hump</u> . She was born on 01.02.1985 in London. She was already <u>jamming</u> on the piano at the age of 5.
Question in formal political context	Formal answer	Colloquial answer
What does the Green Party say about the speed limit on highways?	According to their election program, the Green Party wants to limit the speed on highways to 130 km/h in general. Faster driving would no longer be possible anywhere.	According to their election program, the Green Party wants to limit the speed on highways to 130 " <u>things</u> " in general. Driving <u>with more "Karacho"</u> would no longer be possible anywhere.
What does the SPD say about retirement?	In their election programme, the SPD demands that it be possible to enter retirement at 67. In the SPD's view, the Riester pension model is outdated in its current form.	In its election programme, the SPD demands that people be able to <u>retire</u> at 67. In the SPD's view, the Riester pension model <u>has had its day</u> .
Where can I find out more about the positions of the FDP?	More information on the positions of the FDP can be found at fdp.de. The most important points and further facts on politics can also be found at bpb.de.	More <u>info</u> on the positions of the FDP can be found at fdp.de. The most important <u>benchmarks</u> and other facts from the <u>world of politics</u> can also be found at bpb.de.

Table 1: Questions in informal everyday and formal political context and answers in formal and colloquial speech style (colloquial wordings underlined)