# Aye, Robot: What Happens When Robots Speak Like Real People?

Mary Ellen Foster
School of Computing Science
University of Glasgow
Glasgow, United Kingdom
maryellen.foster@glasgow.ac.uk

Jane Stuart-Smith
English Language & Linguistics
University of Glasgow
Glasgow, United Kingdom
jane.stuart-smith@glasgow.ac.uk

Muneeb I. Ahmad
Department of Computer Science
Swansea University
Swansea, United Kingdom
m.i.ahmad@swansea.ac.uk

Julian Hough
Computer Science
Swansea University
Swansea, United Kingdom
julian.hough@swansea.ac.uk

## Abstract

In daily life, we interact with each other using the social, regional, and ethnic communication styles typical of our local communities. Successful communication further rests on our ability to seamlessly adjust to our interlocutors following the norms and expectations of our local social setting as well as conversational context and goals. However, despite significant advances in speech technology, most artificial speech systems—particularly, most social robots—still use a single, "standard", non-local communication style for all users, social settings and interaction goals. Recent research has shown that when they interact with digital agents, humans transfer and adapt their sociolinguistic behaviours, including communication bias. Despite this, the barriers set up by this inherent communication bias have never been systematically studied for HRI; and the potential benefits to user engagement from socially inclusive, diverse communication styles have not been explored. We argue that social robotics researchers should also consider sociolinguistic factors constraining human interaction. To explore the implications, we describe two hypothetical robots designed to support the local communication style of two regions of the United Kingdom, and we consider the potential sociolinguistic impact each robot might have on its conversational partners and the wider society.

## CCS Concepts

• **Human-centered computing** → **Natural language interfaces**; • **Social and professional topics** → **Cultural characteristics**; • **Computer systems organization** → *Robotics*.

## Keywords

Social robotics, Communication bias, Robots in public spaces

**ACM Reference Format:**

Mary Ellen Foster, Jane Stuart-Smith, Muneeb I. Ahmad, and Julian Hough. 2025. Aye, Robot: What Happens When Robots Speak Like Real People?. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces (CUI '25), July 08–10, 2025, Waterloo, ON, Canada.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3719160.3737615



*CUI '25, Waterloo, ON, Canada*



ACM ISBN 979-8-4007-1527-3/25/07
https://doi.org/10.1145/3719160.3737615

## 1 Introduction

The world is home to substantial sociolinguistic diversity. For example, whilst the UK generally has "English" as its common language, and there is a supralocal, codified "standard" form of (Southern) English recognized as socially-prestigious and typically used in media, broadcasting and education [30, 39], most people use a wide variety of local non-standard dialects (verbal and non-verbal ways of interacting) for daily communication. Importantly, all dialects are loaded with attitudinal and cognitive biases: standard dialects are considered authoritative and competent but less socially attractive; non-standard dialects are rated as warm and friendly but lower on dimensions of competence and intelligence [36]. However, irrespective of use case or user, standard English (either General American or RP/Southern Standard British English) is generally selected for speech-based technology. This systematically-encoded communication bias is not just an inconvenience, but a potential barrier to communication, participation and inclusion [18, 37]. Furthermore, digital technologies play an increasingly central role in our lives, with speech and speech-generating interfaces at the forefront: such technologies are common in all contexts and have become lifelines for individuals with disabilities who may rely on them for communication. Ensuring that such interfaces are inclusive and adaptable to diverse communication styles is not just a matter of convenience, but arguably one of fundamental human rights.

We particularly focus on user interactions with social robots: a robot, as a physically-embodied agent, is unavoidably situated in a specific physical and sociolinguistic context, which makes the appropriateness and relevance of its regional communication style (or lack thereof) even more impactful. Interestingly, whilst non-embodied digital agents have started to use a broader range of communicative styles, at least in the form of more potential accent choices [44, 60], social robotics has remained largely thoroughly conservative, with many researchers choosing voices for reasons such as convenience and availability, or because they had been used in previous studies [35]. We argue that researchers and system developers must challenge this prevailing "one size fits all" notion of communication for social robotics/HRI which does not cater to the diverse sociolinguistic characteristics and needs of users. Working towards this goal has the potential to address communication bias,

break down barriers, foster understanding, and empower individuals of all backgrounds to fully engage with technology. Ultimately, the goal is a global community where social inclusion is the norm for digital communication.

To open discussions around expanding the communicative styles of social robots, we consider two hypothetical robot use cases, each explicitly designed for a region of the United Kingdom where there is a strong, distinctive local accent and culture: a guide robot designed for use in a Welsh tourist attraction, and a companion robot designed to work with senior citizens in a Glasgow museum. For each robot, we briefly consider the technical feasibility of such a robot using state-of-the-art technology, discuss some of the sociolinguistic issues and implications in terms of design and user reception, and then outline the possible impacts that such a robot might have for its local users, who in both cases might never before have encountered a robot communicating like a local.

## 2 Background

### 2.1 Social robots in public spaces

In social robotics, the goal is to develop robots – physically embodied, autonomous agents – that allow humans to engage in face-to-face interactions using the same norms and conventions they use to interact with each other. A crucial factor that makes a robot "social" is the ability to engage in two-way interactions, where the robot both recognises the social signals of its human partners and generates appropriate social signals in response [21]. Social signals may be expressed verbally and non-verbally, including linguistic content, speech prosody, body language and posture, gestures, and facial expressions. Because this degree of social interaction presents significant technical challenges, much initial work on social robots was lab-based, but these robots are increasingly being deployed in public [41]. By far the most common use case for a public-space social robot is as a receptionist or guide; the Pepper robot has been successfully used in this capacity in shopping malls [17], museums [12], restaurants [53], libraries [42], and train stations [57], and similar deployments have also been successful for other robots including Ari, Furhat and Robovie [2, 4, 33]. While social robots are becoming common in public, there has been little investigation into how location, robot role, and target audience may affect the design choices for these robots, including how they might best interact with their users.

### 2.2 Communication bias in human-human interaction

A substantial body of sociolinguistic research has demonstrated highly structured relationships between speakers' social attributes – their age, gender identity, social, regional and ethnic background – and all kinds of linguistic variation, from core aspects of pronunciation, grammar, vocabulary [29, 55], and interactional features like turn timing [52], to paralinguistic features such as how close people stand to each other, how much eye contact they make (and when), how expressive their faces and gestures are, and in the use of body movements to acknowledge, modify, amend, prompt, interrupt or adapt what is said [e.g., 14, 16, 26]. These dialectal differences between communities and speakers are also highly context-sensitive, varying stylistically in response to a range of factors, including formality, interlocutor, locality, and interaction function [9]. Specifically, sociolinguists have demonstrated that humans acquire communicative competence [7, 20] which governs the appropriateness of verbal and non-verbal language use in specific social situations. This includes knowledge of "when to speak, when not and … what to talk about with whom, when, were, in what manner" [23].

Dialects also carry complex intersecting sets of attitudes and prejudices, constituting inherent communication bias. This means that speakers not only acquire knowledge of communication style, so which variety to use in which context, they also acquire shared ideologies and stereotypes about how to evaluate different dialects and their speakers. Research on language attitudes has shown that dialects – and their speakers – are consistently evaluated in terms of dimensions relating to status/superiority (competence, intelligence, education), solidarity/social attractiveness (warmth, likeability, friendliness), and dynamism (e.g. enthusiasm, energy) [19, 62]. Recent work has focussed on the social impacts arising from a particular facet of communicative competence, "accent bias" [31, 45], i.e. positive or negative evaluation of speakers solely based on their accent. Surveys and matched guise studies (using the same speaker with different accents) have demonstrated that speakers of Standard Southern British English (SSBE) are rated as more intelligent, competent and educated than those speaking regional and/or urban UK accents, but at the same time, the same speakers are considered to be less likeable or friendly; and the reverse evaluations are found for regional, urban, and minority ethnic accents [36, 50]. At the same time, ratings are also influenced by listener's own dialect background, local context, and social setting of language use. Communication bias is an issue for human interaction because it impacts listener effort and comprehension [59], as well as emotional and social responses to the person, professionalism and fairness of the speaker [6]. In other words, communication bias has real-world impacts on speakers' lives, just because of how they talk.

### 2.3 Communication bias in spoken language technology

Since the goal for social robots is to leverage human communication mechanisms, and it is known that humans extend and adapt human-human interactional behaviours when interacting with non-human agents [34, 43, 51], we hypothesise that many aspects of HRI may similarly be likely to be negatively impacted by communication bias because of dialect choice and communication styles which are inappropriate for the users, robot role and deployment context. Or, put positively, it might be possible to improve the efficacy of HRI by exploiting known aspects of communication bias around dialects, in social robot design.

Recent research has confirmed that human communication bias also applies with speech-based systems. Automated speech recognition (ASR) has been shown to perform less well on accents such as African American Vernacular English (AAVE) [27], with word error rates up to twice those for other accents, and this has a disproportionate impact on the self-esteem of Black users [37, 61]. Perhaps due to such experiences, speakers of non-standard dialects often expect to have to change the way they speak for a speech-based system [18]. Similar considerations also hold for system output: for example, an animated "virtual peer" was able to engage better with

schoolchildren when it used the same AAVE style that they did [8]; participants in New Zealand preferred a robot that spoke with a local accent over one that used a General American accent in a health-care setting [56]; and proficient dialect speakers in Berlin trusted a video-presented robot more when it used the local dialect than when it spoke Standard German [28].

Speech technology developers are creating a more diverse range of synthesised voices, and ASR researchers are actively working to fine-tune their models on a wider range of more diverse accents. While these innovations are being deployed in the context of non-embodied agents such as Siri, studies on the role of accent for physically-embodied robot agents are extremely rare. In a survey carried out in a UK university [58] of mainly young, Southern, British speaker preferences for a generic, hypothetical robot accent—with no embodiment, role or use case specified—only a third of respondents chose SSBE, with a large range of other preferred accents being selected.

## 3 Representative robot use cases

Based on the factors above, our position is that robots should be tailored whenever possible to the local speaking style. As this is not something that people have generally experienced, we now consider two hypothetical robot systems, each designed for a specific use case in a part of the United Kingdom that has a strong and distinctive regional dialect. We briefly discuss the sociolinguistic and technical considerations involved in developing each robot, and then examine the likely sociological implications of deploying such a robot.

### 3.1 A tourist information robot at Cardiff Castle

We consider a tour guide robot designed to work in Cardiff Castle,[1] a historic castle located in the city centre in Cardiff. The castle's history goes back 2000 years, including Roman, Norman, and mediaeval buildings, and a significant feature is the neo-Gothic Victorian apartments within the Castle. The Castle allows visitors to explore freely, and also offers a range of guided tours including costumed mediaeval tours.

The role of the robot would to be work alongside the human staff and tour guides within the castle grounds. Since the site is not suitable for a mobile robot, the robot would be designed to be situated in indoor areas such as the visitor information centre, the Castle Apartments, and the Clock Tower. The persona of the robot would be based on the staff at the Castle itself: a knowledgeable, friendly, and clearly Welsh person who is proud of the Castle and happy to share it with visitors from around the UK and abroad.

*3.1.1 Sociolinguistic considerations.* Welsh English has been the subject of a large number of attitudinal studies [e.g., 46]. Welsh people themselves have tended to evaluate Welsh English lower in terms of status, but higher for solidarity traits of trustworthiness, honesty, and good humour. Indeed, on a recent UK reality series, a contestant deliberately chose to use a Welsh accent with the hope of seeming trustworthy [5]. Responses from English-English study participants have been mixed, ranging from fairly negative (in comparison with RP) in the 1970s and 80s, but improved to third most positive in a simple YouGov ratings survey of British and Irish accents, and was also considered "gentle" and "friendly" in [40]. In terms of regional Welsh accent, Coupland et al. [11] found that the Southwestern accent was considered by Welsh participants as "Truly Welsh", and scored better in terms of prestige, pleasantness and dynamism. Were the tour guide to be a human, a likely expected accent would be a South Welsh accent, either Southwestern or urban Southeastern (e.g. Cardiff), and so could be selected for the robot.

*3.1.2 Technical considerations.* From a technical standpoint, the required behaviour would be straightforward to implement: current off-the-shelf text-to-speech technology [e.g., 15] can very easily produce high-quality, realistic, Welsh-accented English. Providing visitor information would also be relatively straightforward to implement through techniques such as Retrieval-Augmented Generation [32] using existing data sources. And managing the conversation in this sort of informative context would also be straightforward using a combination of slot-filling dialogue techniques alongside an LLM to provide persona responses and social chat.

For speech recognition, given the interactions we expect, it is possible that off-the-shelf speech recognition techniques could be suitable. However, as the robot would be expected to interact with visitors from everywhere in the world, it will be crucial to train and test the speech recogniser on a corpus such as the Edinburgh International Accents of English Corpus [48]: current speech recognition technology still has variable performance depending on the accent (with notably bad performance on Welsh English [25]), and it would be completely inappropriate to transfer this bias to the guide robot.

*3.1.3 Sociolinguistic HRI implications.* Selecting a favourably-rated Welsh accent for the tour guide robot could help promote a "Welsh" persona, in line with the goal of enhancing the experience of visiting a major Welsh site. Positive evaluations of the Welsh accent by both Welsh and British speakers in terms of warmth could also increase user engagement with the robot, and might be shared by international visitors. The increasing use of other regional/national accents for digital agents, e.g. Standard Scottish English and Standard Irish (Republic) English, might well mean that users would be less likely to be surprised - and so confused - by a social robot using this local accent for this particular role. Potential risks of voicing the robot with the Welsh accent could be: users might find the robot accessible, but also regard the robot as less authoritative/competent, and some non-native English speakers might find the robot more difficult to understand. In general, using a relatively standard variety of (Southern) Welsh English would seem to be relatively low risk and high gain, maximising inherent positive bias and enhance the user experience for many.

### 3.2 A companion robot museum guide for senior citizens in Glasgow

For this robot, we consider the "Ageing Well" Programme[2] offered by a group of public museums in Glasgow. This programme offers free sessions to groups of older adults and care homes, providing a relaxed visit to the museum at a slower pace, supported by a

[1] https://www.cardiffcastle.com/

[2] https://www.glasgowlife.org.uk/museums/glasgow-life-museums-ageing-well-programme/ageing-well-programme-groups

member of the museum team. Participants have the opportunity to handle real museum objects and to have a hot drink and a chat.

The role of the robot would be to work alongside the museum staff member in supporting the group: for example, it might engage in social chat with some participants, and would also have access to information about the museum exhibits and objects. Its persona would be a friendly and informed Glaswegian.

Note that, while this use case is similar in some ways to the Cardiff Castle example, the role and target audience are different. Unlike the castle robot, the role of this robot is to be a social companion rather than an authoritative tour guide; also, the target audience is elderly people living in the Glasgow area rather than international visitors.

*3.2.1 Sociolinguistic considerations.* The city of Glasgow is well known for its distinctive vernacular dialect, Glaswegian, which is infamous for being difficult to understand by English-English people, as well as more generally [1]. Attitudinal studies have consistently shown Glaswegian to be rated extremely low across UK participants in terms of prestige, intelligence and competence, but more highly for warmth and humour [10]. Glaswegian is also overtly stigmatized by the local population, who report their own dialect as "bad language" or "slang" [38]. At the same time, Glaswegian holds substantial "covert prestige" in terms of likeability, honesty, and friendliness for urban Scots and is widely spoken and strongly maintained by the majority of city dwellers. SSBE typically scores highly for status amongst Scottish and Glaswegian participants, but is at the same time regarded with some hostility [54]. Within Scotland, the regional national standard accent, Scottish Standard English, is used for education, professional life, and local national broadcasting. Scottish Standard English generally scores much higher than Glaswegian in UK attitudinal surveys, often positioned just below SSBE. Amongst Scottish and Glaswegians, Scottish Standard English is better liked than SSBE, but with lower ratings for warmth.

Unlike the previous use case, there is evidence from a recent survey of how Scottish users might respond to a (typical standard English-speaking) social robot [18]. The results showed a sociolinguistic asymmetry: in general participants expected to be able to understand the robot, whilst 70% expressed some doubt that communicating with the robot would be easy – and qualitative analysis showed that those who identified as working-class explicitly mentioned their own accent as a potential barrier for communication. There was also an effect of age, with older participants showing less confidence about being able to understand the robot.

Were the museum companion to be human, they could speak with a Standard Scottish English accent, but might be found more approachable and friendly by the target population of elderly users if they spoke Glaswegian.

*3.2.2 Technical considerations.* As in the previous case, some core behaviour would be straightforward to implement using current off-the-shelf text-to-speech technology for the Glaswegian voice, and Retrieval-Augmented Generation to generate descriptions of museum objects.

However, some other aspects of this robot could be challenging technically. Based on the goals of the Glasgow museum programme, it is likely that some of the participants might have dementia or similar challenges, and this is a population for whom research is still very much ongoing on how best to develop a conversational system [3, 24, 47]. Also, on the speech recognition side, off-the-shelf recognition might have issues handling either this local dialect or speech from elderly participants: to support this population would require not only training models on a broader dataset, but also likely integrating state-of-the-art research models into the overall system such as [13, 22].

*3.2.3 Sociolinguistic HRI implications.* Voicing the companion museum guide robot with a Glaswegian accent would be genuinely novel. It would also carry the potential risks of eliciting surprise and concern from both the elderly users, as well as from the museum staff. It might be the case that the users would be less likely to want to engage with the robot and/or understand its information, whilst the staff might be less willing for the robot to be regarded as an authoritative information-bearing source. If so, these responses would indicate transfer of bias associated with a highly-stigmatized accent being used in a less familiar social setting, where – even with a strong social companion "friendship" role – minimally a standard accent would be expected. The modality of the robot itself, with the implicit expectation of a standard accent, could also cause communicative dissonance. Selection of a "halfway" option, e.g. a Standard Scottish English accent, might be a mitigating strategy, and act as a more acceptable option given that it is better regarded than SSBE, and a known voice for digital agents.

At the same time, we could also envisage a scenario in which the elderly users quickly accommodate to the apparently inconsistent modalities of a robot with their own local accent, and transfer existing positive solidarity traits to the interaction, thus improving their experience. It is also possible that elderly users might have a less well established expectation that the robot should speak with a standard accent, especially if they had little previous experience with digital agents [18]. A positive outcome with the users might then also persuade museum staff that a locally-voiced robot would enhance their resources for local city residents more generally.

Overall, the potential sociolinguistic outcome of this use case is less easy to predict given the low status of Glaswegian dialect, and the general reluctance of digital technology to embrace local regional accents [though see 60].

## 4 Discussion and Conclusion

We have argued that social robotics researchers should aim to broaden the communication styles of their robots to include local regional, social and ethnic dialects—that is, to make robots talk more like real people. To explore the implications, we have explored two potential cases of a social robot explicitly designed to follow local communication norms and have discussed the sociological implications of each. We predict that a Welsh-accented robot tour guide would likely be positively received; however, due to the differences in both the use case and the status of the Glasgow accent, it is harder to predict the outcome of the second use case.

Given this uncertainty, it would be easy to question the value: if "standard" voices appear adequate, and the technical issues remain, why take up the uncharted challenge of using local norms? One reason is that when users, more representative of the broader population, evaluate digital agents and social robots, they often indicate

dissatisfaction with the status quo, rejection of the standard, and a desire for their own dialects [8, 28, 56].

Another perhaps more compelling reason is that just because a sociolinguistic convention is in place does not mean that it should remain. A good analogy can be found in the recent history of British broadcasting: for many years, the BBC exclusively used RP, the socially-prestigious, non-regional accent of a tiny proportion of the British population, on the grounds that it would be the most comprehensible. The first departure from RP was in 1941, when Wilfred Pickles, a Northern English (Yorkshire) speaker, was selected to read the news on the National Service, as a tactic to avoid German enemy impersonation of British media. Initial reactions to hearing the news in a Northern English accent were mixed: some were delighted, but many were outraged [49]. The BBC returned to RP after the war, but later expanded its use of regional accents, sometimes with resistance: Susan Rae famously received hate mail for her Scottish accent on Radio 4 in the 1980s. Today, authority and informational broadcasting is delivered using a range of regional accents across the United Kingdom.

Ultimately, while shifting away from standard sociolinguistic norms for voicing social robots might seem counterintuitive and difficult, and initially potentially odd, once the steps have been taken, they could lead to a far wider community of users and much more effective and enjoyable HRI for all.

## References


[1] Patti Adank, Bronwen G Evans, Jane Stuart-Smith, and Sophie K Scott. 2009. Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of experimental psychology. Human perception and performance* 35 (2009), 520–529. Issue 2. doi:10.1037/a0013552

[2] Angus Addlesee, Neeraj Cherakara, Nivan Nelson, Daniel Hernandez Garcia, Nancie Gunson, Weronika Sieińska, Christian Dondrup, and Oliver Lemon. 2024. Multi-party Multimodal Conversations Between Patients, Their Companions, and a Social Robot in a Hospital Memory Clinic. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*. St. Julians, Malta, 62–70. https://aclanthology.org/2024.eacl-demo.8/

[3] Nada Altwala, Christopher Bull, and David Kirk. 2024. Exploring the Potential of AI Chatbots for Supporting Individuals with Dementia: Thematic Insights from User-Centered Workshops. In *Proceedings of the 12th International Conference on Human-Agent Interaction* (Swansea, United Kingdom) *(HAI '24)*. 432–434. doi:10.1145/3687272.3690907

[4] Dražen Brščić, Tetsushi Ikeda, and Takayuki Kanda. 2017. Do You Need Help? A Robot Providing Information to People Who Behave Atypically. *IEEE Transactions on Robotics* 33, 2 (2017), 500–506. doi:10.1109/TRO.2016.2645206

[5] Charlie Buckland. 2025. The Traitors: Is a Welsh accent really more trustworthy? *BBC News* (Jan. 2025). https://www.bbc.co.uk/news/articles/cql5xg6x0dko

[6] Kathryn Campbell-Kibler. 2009. The nature of sociolinguistic perception. *Language Variation and Change* 21, 1 (2009), 135–156. doi:10.1017/S0954394509000052

[7] Michael Canale and Merrill Swain. 1980. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics* I, 1 (March 1980), 1–47. doi:10.1093/applin/I.1.1

[8] Justine Cassell. 2009. Social Practice: Becoming Enculturated in Human-Computer Interaction. In *Universal Access in Human-Computer Interaction. Applications and Services*, Constantine Stephanidis (Ed.). 303–313. doi:10.1007/978-3-642-02713-0_32

[9] N Coupland. 2007. *Style: Language Variation and Identity*. CUP.

[10] Nikolas Coupland and Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics* 11, 1 (2007), 74–93.

[11] Nikolas Coupland, Angie Williams, and Peter Garrett. 1994. The social meanings of Welsh English: Teachers' stereotyped judgements. *Journal of Multilingual & Multicultural Development* 15, 6 (1994), 471–489.

[12] Laura-Dora Daczo, Lucie Kalova, Kresta Louise F. Bonita, Marc Domenech Lopez, and Matthias Rehm. 2021. Interaction Initiation with a Museum Guide Robot—From the Lab into the Field. In *Human-Computer Interaction – INTERACT 2021*. 438–447. doi:10.1007/978-3-030-85613-7_30

[13] Renchang Dong, Yijie Li, Dongxing Xu, and Yanhua Long. 2024. Cross-Modal Parallel Training for Improving end-to-end Accented Speech Recognition. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 10396–10400. doi:10.1109/ICASSP48485.2024.10447979

[14] David Efron. 1972. *Gesture, race and culture: A tentative study of the spatio-temporal and "linguistic" aspects of the gestural behavior of eastern Jews and southern Italians in New York City, living under similar as well as different environmental conditions.*

[15] ElevenLabs. 2025. Free Text to Speech & AI Voice Generator. https://elevenlabs.io/.

[16] Hillary Anger Elfenbein. 2013. Nonverbal Dialects and Accents in Facial Expressions of Emotion. *Emotion Review* 5, 1 (2013), 90–96. doi:10.1177/1754073912451332

[17] Mary Ellen Foster, Rachid Alami, Olli Gestranius, Oliver Lemon, Marketta Niemelä, Jean-Marc Odobez, and Amit Kumar Pandey. 2016. The MuMMER Project: Engaging Human-Robot Interaction in Real-World Public Spaces. In *Proceedings of the International Conference on Social Robotics*. 753–763. doi:10.1007/978-3-319-47437-3_74

[18] Mary Ellen Foster and Jane Stuart-Smith. 2023. Social Robotics meets Sociolinguistics: Investigating Accent Bias and Social Context in HRI. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. 156–160. doi:10.1145/3568294.3580063

[19] Peter Garrett. 2010. *Attitudes to language*. Cambridge University Press.

[20] John J. Gumperz. 1984. *Communicative competence revisited*. Technical Report. Cognitive Science Program, Institute of Cognitive Studies, University of California at Berkeley.

[21] Anna Henschel, Guy Laban, and Emily S. Cross. 2021. What Makes a Robot Social? A Review of Social Robots from Science Fiction to a Home or Hospital Near You. *Current Robotics Reports* 2, 1 (Feb. 2021), 9–19. doi:10.1007/s43154-020-00035-0

[22] Shujie Hu, Xurong Xie, Mengzhe Geng, Zengrui Jin, Jiajun Deng, Guinan Li, Yi Wang, Mingyu Cui, Tianzi Wang, Helen Meng, and Xunying Liu. 2024. Self-Supervised ASR Models and Features for Dysarthric and Elderly Speech Recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 32 (2024), 3561–3575. doi:10.1109/TASLP.2024.3422839

[23] Dell Hymes. 1972. On communicative competence. *Sociolinguistics* (1972).

[24] Bahar Irfan, Sanna Kuoppamäki, Aida Hosseini, and Gabriel Skantze. 2025. Between reality and delusion: challenges of applying large language models to companion robots for open-domain dialogues with older adults. *Autonomous Robots* 49, 1 (March 2025). doi:10.1007/s10514-025-10190-y

[25] Dewi Jones. 2022. Development and evaluation of speech recognition for the welsh language. In *Proceedings of the 4th Celtic Language Technology Workshop within LREC2022*. 52–59.

[26] Adam Kendon. 2004. *Gesture: Visible Action as Utterance*. Cambridge University Press.

[27] Allison Koenecke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Toups, John R. Rickford, Dan Jurafsky, and Sharad Goel. 2020. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences* 117, 14 (2020), 7684–7689. doi:10.1073/pnas.1915768117

[28] Katharina Kühne, Erika Herbold, Oliver Bendel, Yuefang Zhou, and Martin H. Fischer. 2024. "Ick bin een Berlina": dialect proficiency impacts a robot's trustworthiness and competence evaluation. *Frontiers in Robotics and AI* Volume 10 - 2023 (2024). doi:10.3389/frobt.2023.1241519

[29] William Labov. 2001. *Principles of Linguistic Change: Social Factors*. Vol. Volume 1. Blackwell.

[30] Stephen Levey. 2024. *Standard and Non-standard English*. Cambridge University Press, 48–69.

[31] Erez Levon, Devyani Sharma, Dominic J L Watt, Amanda Cardoso, and Yang Ye. 2021. Accent Bias and Perceptions of Professional Competence in England. *Journal of English Linguistics* 49 (2021), 355–388. Issue 4.

[32] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 9459–9474. https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf

[33] Mei Yii Lim, José David Aguas Lopes, David A. Robb, Bruce W. Wilson, Meriam Moujahid, and Helen Hastie. 2022. Demonstration of a Robo-Barista for In the Wild Interactions. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (Sapporo, Hokkaido, Japan) *(HRI '22)*. IEEE Press, 1200–1201.

[34] Birgit Lugrin. 2021. *Introduction to Socially Interactive Agents* (1 ed.). Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3477322.3477324

[35] Conor McGinn and Ilaria Torre. 2019. Can you Tell the Robot by the Voice? An Exploratory Study on the Role of Voice in the Perception of Robots. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 211–221. doi:10.1109/HRI.2019.8673305

[36] Robert M. McKenzie and Andrew McNeill. 2022. *Implicit and Explicit Language Attitudes: Mapping Linguistic Prejudice and Attitude Change in England*. Taylor &

Francis.
[37] Zion Mengesha, Courtney Heldreth, Michal Lahav, Juliana Sublewski, and Elyse Tuennerman. 2021. "I don't Think These Devices are Very Culturally Sensitive."—Impact of Automated Speech Recognition Errors on African Americans. *Frontiers in Artificial Intelligence* 4 (2021). doi:10.3389/frai.2021.725911
[38] Janet Menzies. 1991. An Investigation of Attitudes to Scots. http://www.arts.gla.ac.uk/STELLA/STARN/lang/MENZIES/menzie1.htm
[39] J Milroy and L Milroy. 1985. *Authority in language: Investigating language prescription and standardisation.* Routledge.
[40] Chris Montgomery. 2016. *The Perceptual Dialectology of Wales from the Border.* Palgrave Macmillan UK, London, 151–179. doi:10.1057/978-1-137-52897-1_6
[41] Omar Mubin, Muneeb Imtiaz Ahmad, Simranjit Kaur, Wen Shi, and Aila Khan. 2018. Social Robots in Public Spaces: A Meta-review. In *International Conference on Social Robotics*. 213–220. doi:10.1007/978-3-030-05204-1_21
[42] Omar Mubin, Isha Kharub, and Aila Khan. 2020. Pepper in the Library" Students' First Impressions. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA). 1–9. doi:10.1145/3334480.3382979
[43] Clifford Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)*. 72–78. doi:10.1145/191666.191703
[44] Matthew Panzarino. 2021. Apple adds two brand new Siri voices and will no longer default to a female or male voice in iOS. *TechCrunch* (2021). https://techcrunch.com/2021/03/31/apple-adds-two-siri-voices/
[45] Laura Paterson. 2019. Interview with Erin Carrie and Rob Drummond of The Accentism Project. *Journal of Language and Discrimination* 3, 1 (2019), 76–84.
[46] Heli Paulasto, Rob Penhallurick, and Benjamin Jones. 2020. *Welsh English.* Vol. 12. Walter de Gruyter GmbH & Co KG.
[47] Nicole Ruggiano, Ellen L Brown, Lisa Roberts, C Victoria Framil Suarez, Yan Luo, Zhichao Hao, and Vagelis Hristidis. 2021. Chatbots to Support People With Dementia and Their Caregivers: Systematic Review of Functions and Quality. *J Med Internet Res* 23, 6 (3 Jun 2021), e25006. doi:10.2196/25006
[48] Ramon Sanabria, Nikolay Bogoychev, Nina Markl, Andrea Carmantini, Ondrej Klejch, and Peter Bell. 2023. The Edinburgh International Accents of English Corpus: Towards the Democratization of English ASR. In *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 1–5. doi:10.1109/ICASSP49357.2023.10095057
[49] Robert Seatter. 2025. In Focus: Wilfred Pickles. *BBC Creative Diversity* (2025). https://www.bbc.co.uk/creativediversity/history/wilfredpickles
[50] Devyani Sharma, Erez Levon, and Yang Ye. 2022. 50 years of British accent bias : Stability and lifespan change in attitudes to accents. *English World-Wide* 43, 2 (2022), 135–166. doi:10.1075/eww.20010.sha
[51] Laura Staum Casasanto, Kyle Jasmin, and Daniel Casasanto. 2010. Virtually accommodating: Speech rate accommodation to a virtual interlocutor.. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 32.
[52] Tanya Stivers, N. J. Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, Federico Rossano, Jan Peter de Ruiter, Kyung-Eun Yoon, and Stephen C. Levinson. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106, 26 (2009), 10587–10592. doi:10.1073/pnas.0903616106
[53] Ruth Maria Stock and Moritz Merkle. 2018. Can humanoid service robots perform better than service employees? A comparison of innovative behavior cues. In *Proceedings of the 51st Hawaii international conference on system sciences*.
[54] J Stuart-Smith. 2006. The influence of the media on language. In *The Routledge Companion to Sociolinguistics*, C Llamas, P Stockwell, and L Mullany (Eds.). Routledge, 140–148.
[55] Sali Tagliamonte. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation.* Wiley-Blackwell.
[56] Rie Tamagawa, Catherine I. Watson, I. Han Kuo, Bruce A. MacDonald, and Elizabeth Broadbent. 2011. The Effects of Synthesized Voice Accents on User Perceptions of Robots. *International Journal of Social Robotics* 3, 3 (June 2011), 253–262. doi:10.1007/s12369-011-0100-4
[57] Sofia Thunberg and Tom Ziemke. 2020. Are People Ready for Social Robots in Public Spaces?. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (Cambridge, United Kingdom) *(HRI '20)*. 482–484. doi:10.1145/3371382.3378294
[58] Ilaria Torre and Sébastien Le Maguer. 2020. Should robots have accents?. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 208–214. doi:10.1109/RO-MAN47096.2020.9223599
[59] Kristin J. Van Engen and Jonathan E. Peelle. 2014. Listening effort and accented speech. *Frontiers in Human Neuroscience* 8 (2014). doi:10.3389/fnhum.2014.00577
[60] Jim Waterson. 2020. 'Hey Beeb': new BBC digital assistant gets northern male accent. *The Guardian* (2020). https://www.theguardian.com/media/2020/jun/03/hey-beeb-new-bbc-digital-assistant-gets-northern-male-accent
[61] Kimi Wenzel, Nitya Devireddy, Cam Davison, and Geoff Kaufman. 2023. Can Voice Assistants Be Microaggressors? Cross-Race Psychological Responses to Failures of Automatic Speech Recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Article 109. doi:10.1145/3544548.3581357
[62] Christopher J Zahn and Robert Hopper. 1985. Measuring language attitudes: The speech evaluation instrument. *Journal of language and social psychology* 4, 2 (1985), 113–123.