

AI-Driven Support for People with Speech & Language Difficulties

Aayushi Dangol
adango@uw.edu
University of Washington
Seattle, WA, USA

Yun Huang
yunhuang@illinois.edu
University of Illinois
Urbana-Champaign
Seattle, WA, USA

Srirangaraj Setlur
setlur@buffalo.edu
University at Buffalo
Buffalo, NY, USA

Adele Smolansky
adeles@stanford.edu
Stanford University
Stanford, CA, USA

Hari Subramonyam
harihars@stanford.edu
Stanford University
Stanford, CA, USA

Hyewon Suh
hyewon25@uw.edu
University of Washington
Seattle, WA, USA

Jinjun Xiong
jinjun@buffalo.edu
University at Buffalo
Buffalo, NY, USA

Julie A. Kientz
jkientz@uw.edu
University of Washington
Seattle, WA, USA

ABSTRACT

Speech and language difficulties present significant challenges to effective communication, impacting individuals' ability to express themselves and engage in meaningful interactions. Recent advances in AI technologies, particularly in natural language processing (NLP) and machine learning, have the potential to assist individuals with speech and language difficulties in improving their communication outcomes. However, given the probabilistic nature of AI models, there is a need to adopt and advance human-centered AI design methodologies to support the prototyping of AI user experiences. This Special Interest Group (SIG) aims to bring together researchers, practitioners, and designers from the fields of AI, accessibility, speech pathology, AI ethics, and HCI to facilitate high-level discussions around designing and evaluating reliable, safe, and human-centered AI-driven support and interventions for supporting individuals with speech and language difficulties.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility technologies; Usability testing.**

KEYWORDS

Speech and language difficulties, AI prototyping, Natural language processing, Human-centered design

ACM Reference Format:

Aayushi Dangol, Yun Huang, Srirangaraj Setlur, Adele Smolansky, Hari Subramonyam, Hyewon Suh, Jinjun Xiong, and Julie A. Kientz. 2024. AI-Driven Support for People with Speech & Language Difficulties. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
CHI EA '24, May 11–16, 2024, Honolulu, HI, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0331-7/24/05
<https://doi.org/10.1145/3613905.3643984>

(CHI EA '24), May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3613905.3643984>

1 INTRODUCTION

Reaching an understanding in everyday communicative interactions hinges on the successful reception, understanding, acceptance, and reciprocity between the speaker and the listener [23, 29]. When the speaker or listener encounters difficulty achieving these elements, the ability to convey or understand thoughts and ideas may be hindered, leading to a communication breakdown [23]. Given the complex nature of human interaction, communication breakdowns are observed in events involving people with “fully developed speech and language functions” [23, p.1]. However, individuals with speech and language difficulties experience a significantly higher rate of communication breakdowns per communication unit [16].

Speech and language difficulties comprise a wide range of conditions, including difficulty in producing speech sounds correctly (e.g., stuttering), the inability to express or comprehend language, difficulties in social and cognitive communication, as well as feeding and swallowing difficulties [1]. These difficulties can make traditional communication methods challenging, causing frustration and isolation for both the affected individual and their interlocutors [17]. For instance, studies indicate that numerous individuals with speech and language difficulties experience depression, social isolation, and a lower quality of life [18]. It can also negatively impact a person's employment status [20] and their ability to receive proper healthcare [9]. The impact of speech and language development is also noticeable among young children. Currently, there exists a significant academic and socio-emotional disparity between children with speech and language difficulties and their matched peers [4].

Early identification and access to Speech Language Pathologists (SLPs) is crucial to improving speech, but the shortage of qualified professionals and the limited time available to provide ability-based interventions pose significant challenges to addressing these needs comprehensively [30]. For example, within the U.S. public school system, despite SLP vacancies being among the most prevalent types of vacancies, they do not receive the necessary attention or

consideration[21]. This shortage of SLPs is creating significant challenges for individuals with speech and language concerns to access timely and comprehensive therapeutic interventions, hindering their ability to overcome communication barriers.

To address challenges related to the effectiveness of interventions and accessibility, many researchers have proposed AI-based automated speech therapy tools for individuals with speech and language difficulties [7, 9, 25, 27, 28]. In the past, CHI has been a platform for hosting numerous papers around this very topic [5–7, 10, 22, 31]. With the recent surge in popularity of generative AI, there is a growing recognition not just among researchers but also among care providers to leverage AI-powered capabilities like speech recognition, alongside quantitative data such as facial expressions and intonations for automating some aspects of care, such as analyzing learning patterns, identifying speech difficulties, and dynamically adjusting instructional strategies [8, 15]. Not only can AI-driven technology alleviate SLPs' burden on delivering and executing interventions, but its scalability can also help reach out to broader populations who have been underserved [19, 24, 26].

At the same time, the distribution of these benefits may vary among stakeholders. For example, bias and fairness issues, potentially embedded in AI systems through training data [3, 34], can create disparities in intervention effectiveness among different demographic groups [14]. Another concern is that these systems may not be as empathetic [13] or unable to provide emotional support to individuals going through speech and language therapy in the same way direct communication with a human would. Thus, there is a need to adopt and advance human centered AI design methodologies in the development of reliable, safe, and trustworthy AI based technologies.

2 GOALS OF THE SIG

The goal of this SIG is to bring together interested researchers to start a conversation around hard questions in designing and evaluating AI-driven support for mediating communication breakdowns with individuals who have speech and language difficulties. By providing a dedicated space for the interdisciplinary exchange of ideas, methodologies, and findings at the intersection of AI, healthcare, education, accessibility, and HCI, this SIG seeks to connect people from different areas of CHI who are working in this domain and provide a forum for them to plan ways to collaboratively advance the development and implementation of AI-driven interventions to support individuals with speech and language difficulties. A few of our goals and action items for the SIG are as follows:

- (1) Establish a collaborative research network to facilitate knowledge exchange and interdisciplinary collaboration among researchers, practitioners, and experts in the fields of AI, healthcare, accessibility, and HCI.
- (2) Discuss and document ethical considerations and user acceptance challenges associated with development and deployment of AI applications in supporting individuals with speech and language difficulties.
- (3) Contribute to the broader conversation on the societal impact of AI technologies in healthcare, education, and accessibility domains.

- (4) Identify directions for research collaborations to address key topics discussed below (see §3).

3 SIG TOPICS AND THEIR RELEVANCE TO THE HCI COMMUNITY

From an HCI standpoint, it is crucial to design and assess AI systems that prioritize the experiences and needs of all stakeholders to ensure a human-centered approach throughout the development lifecycle. For example, recognizing that children often perceive technologies as "creepy" when they involve deception, exhibit unpredictability, or reduce their sense of control [36] is important as frustration induced by poorly designed AI systems could lead children to swiftly abandon these technologies. While actively incorporating feedback from stakeholders throughout the development process can help mitigate initial design challenges, developing prototypes for AI user experiences poses additional complexities due to the probabilistic nature of AI's outputs [11, 35]. For example, evaluating design concepts through conventional low or medium-fidelity prototyping methods, such as Wizard of Oz prototyping, can pose challenges due to uncertainty surrounding AI's capabilities and AI's output complexity [35]. Moreover, both the intended and unintended consequences of AI's outputs such as recognition errors and algorithmic bias, need to be understood which can prevent teams from engaging in rapid iterative prototyping in the first place[35]. For example, speech recognition bias is evident in automated speech recognition systems, which show less proficiency for Black adult speakers of African American English (AAE) compared to white speakers of General American English (GAE) [14]. These disparities are often more pronounced in children, who may have limited exposure to speakers outside their community and consequently display more features of their dialect [32, 33]. While the HCI and design communities have proposed numerous guidelines and recommendations for designing AI user experiences [2, 12, 35], these recommendations often provide guidance at a high-level and lack actionable specifics about how to prototype likely interaction scenarios and failures [11]. At the same time, it is important design and build on general advances in AI ethics, fairness, explainability as a foundation for creating responsible and safe AI systems. Taking these into consideration, the list of potential discussion topics are:

- Challenges in Prototyping AI User Experiences
- Synergies between HCI and AI for Responsible Speech-Language AI Systems
- How AI can complement the role of Speech-Language Pathologists
- Integration of AI with existing speech therapy practices and protocols
- Evaluation of AI's effectiveness in real-world speech therapy scenarios
- Addressing biases in AI algorithms for speech recognition and therapy
- Learner modeling and synthetic data for personalized learning design

4 EXPECTED OUTCOMES

During the SIG itself and in related asynchronous discussions, we will actively encourage the formation of partnerships, including research collaborations or larger efforts. Specifically, participants will be encouraged to engage in discussions with those they have not yet had the opportunity to meet or work with. The SIG could serve as a meeting point for researchers, educators, and practitioners from multiple backgrounds, encompassing HCI, NLP, accessibility, speech-language pathology, AI ethics, and beyond, to come together. Post-SIG, we will create a LinkedIn group to continue coordinating on these issues. All attendees will be welcome to sustain discussions and maintain their presence on the LinkedIn group channel. Any guidelines or findings from this SIG will be written up for publication in *Interactions* or *Communications of the ACM*. Finally, we aspire to continue discussions about the developed research agendas and SIG themes in the future by organizing bi-monthly conversations with invited speakers, which the SIG organizers will facilitate.

ACKNOWLEDGMENTS

This SIG is supported under the AI Research Institutes program by National Science Foundation and the Institute of Education Sciences, U.S. Department of Education, through Award 2229873 - AI Institute for Transforming Education for Children with Speech and Language Processing Challenges or National AI Institute for Exceptional Education. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the U.S. Department of Education.

REFERENCES

- [1] American Speech-Language-Hearing Association. 2022. *Speech-Language Pathologists*. <https://www.asha.org/students/speech-language-pathologists/>
- [2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fournery, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. 2019. Guidelines for human-AI interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–13.
- [3] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the dangers of stochastic parrots: Can language models be too big?. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623.
- [4] Dorothy VM Bishop. 2014. Ten questions about terminology for children with unexplained language problems. *International journal of language & communication disorders* 49, 4 (2014), 381–415.
- [5] Humphrey Curtis, Zihao You, William Deary, Miruna-Ioana Tudoreanu, and Timothy Neate. 2023. Envisioning the (In) Visibility of Discreet and Wearable AAC Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [6] Jiamin Dai, Karyn Moffatt, Jinglan Lin, and Khai Truong. 2022. Designing for Relational Maintenance: New Directions for AAC Research. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [7] Giuseppe Desolda, Rosa Lanzilotti, Antonio Piccinno, and Veronica Rossano. 2021. A system to support children in speech therapies at home. In *CHIItaly 2021: 14th Biannual Conference of the Italian SIGCHI Chapter*. 1–5.
- [8] Joseph R Duffy. 2016. Motor speech disorders: where will we be in 10 years?. In *Seminars in speech and language*, Vol. 37. Thieme Medical Publishers, 219–224.
- [9] Jared Duval, Zachary Rubin, Elena Márquez Segura, Natalie Friedman, Milla Zlatanov, Louise Yang, and Sri Kurniawan. 2018. SpokeIt: building a mobile speech therapy experience. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–12.
- [10] Mauricio Fontana de Vargas, Jiamin Dai, and Karyn Moffatt. 2022. AAC with Automated Vocabulary from Photographs: Insights from School and Speech-Language Therapy Settings. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–18.
- [11] Matthew K Hong, Adam Fournery, Derek DeBellis, and Saleema Amershi. 2021. Planning for natural language failures with the ai playbook. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [12] Eric Horvitz. 1999. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 159–166.
- [13] Angeliki Kerasidou. 2020. Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare. *Bulletin of the World Health Organization* 98, 4 (2020), 245.
- [14] 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.
- [15] Bridget Murray Law. 2020. AI: A New Window Into Communication Disorders? *Leader Live* (2020).
- [16] Barbara G MacLachlan and Robin S Chapman. 1988. Communication breakdowns in normal and language learning-disabled children's conversation and narration. *Journal of speech and hearing disorders* 53, 1 (1988), 2–7.
- [17] Jane McCormack, Sharynne McLeod, Lindy McAllister, and Linda J Harrison. 2010. My speech problem, your listening problem, and my frustration: The experience of living with childhood speech impairment. (2010).
- [18] Ray M Merrill, Nelson Roy, and Jessica Lowe. 2013. Voice-related symptoms and their effects on quality of life. *Annals of Otolaryngology, Rhinology & Laryngology* 122, 6 (2013), 404–411.
- [19] Paul L Morgan, George Farkas, Marianne M Hillemeier, Richard Mattison, Steve Maczuga, Hui Li, and Michael Cook. 2015. Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher* 44, 5 (2015), 278–292.
- [20] Megan A Morris, Sarah K Meier, Joan M Griffin, Megan E Branda, and Sean M Phelan. 2016. Prevalence and etiologies of adult communication disabilities in the United States: Results from the 2012 National Health Interview Survey. *Disability and health journal* 9, 1 (2016), 140–144.
- [21] U.S. News. 2022. *New Federal Data Shows Pandemic's Effects on Teaching Profession*. <https://www.usnews.com/news/education-news/articles/2022-03-02/new-federal-data-shows-pandemics-effects-on-teaching-profession> Accessed: 2022-04-08.
- [22] Mmachi God'sglory Obiorah, Anne Marie Marie Piper, and Michael Horn. 2021. Designing AACs for people with aphasia dining in restaurants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [23] Emily A Ondo. 2015. Acquired language disorders as barriers to effective communication. (2015).
- [24] Rhea Paul. 1996. Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology* 5, 2 (1996), 5–21.
- [25] Pavithra Ramamurthy and Tingyu Li. 2018. Buddy: a speech therapy robot companion for children with cleft lip and palate (cl/p) disorder. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. 359–360.
- [26] Gregory C Robinson and Pamela C Norton. 2019. A decade of disproportionality: A state-level analysis of African American students enrolled in the primary disability category of speech or language impairment. *Language, Speech, and Hearing Services in Schools* 50, 2 (2019), 267–282.
- [27] V Robles-Bykbaev, M Guamán-Heredia, Y Robles-Bykbaev, J Lozano-Redrován, F Pesántez-Aviles, D Quisi-Peralta, M López-Nores, and J Pazos-Arias. 2017. Ontospeltra: A robotic assistant based on ontologies and agglomerative clustering to support speech-language therapy for children with disabilities. In *Advances in Computing: 12th Colombian Conference, CCC 2017, Cali, Colombia, September 19-22, 2017, Proceedings 12*. Springer, 343–357.
- [28] Ahmed Farag Seddik, Mohamed El Adawy, and Ahmed Ismail Shatin. 2013. A computer-aided speech disorders correction system for Arabic language. In *2013 2nd International Conference on Advances in Biomedical Engineering*. IEEE, 18–21.
- [29] Doreen Spilton and Lee C Lee. 1977. Some determinants of effective communication in four-year-olds. *Child Development* (1977), 968–977.
- [30] Jonathan M Sykes and Travis T Tollefson. 2005. Management of the cleft lip deformity. *Facial Plastic Surgery Clinics* 13, 1 (2005), 157–167.
- [31] Stephanie Valencia, Richard Cave, Krystal Kallarackal, Katie Seaver, Michael Terry, and Shaun K Kane. 2023. "The less I type, the better": How AI Language Models can Enhance or Impede Communication for AAC Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [32] Laura Wagner, Cynthia G Clopper, and John K Pate. 2014. Children's perception of dialect variation. *Journal of child language* 41, 5 (2014), 1062–1084.
- [33] Julie A Washington, Lee Branum-Martin, Congying Sun, and Ryan Lee-James. 2018. The impact of dialect density on the growth of language and reading in African American children. *Language, speech, and hearing services in schools* 49, 2 (2018), 232–247.
- [34] Robert Wolfe and Aylin Caliskan. 2022. Detecting Emerging Associations and Behaviors With Regional and Diachronic Word Embeddings. In *2022 IEEE 16th International Conference on Semantic Computing (ICSC)*. IEEE, 91–98.

- [35] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. 2020. Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*. 1–13.
- [36] Jason C Yip, Kiley Sobel, Xin Gao, Allison Marie Hishikawa, Alexis Lim, Laura Meng, Romaine Flor Ofiana, Justin Park, and Alexis Hiniker. 2019. Laughing is scary, but farting is cute: A conceptual model of children's perspectives of creepy technologies. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.