Can AI interpretation increase inclusivity?

Heather Fair[1], Osmary A Medina-Báez[2], Barbara J Spiecker[3], Qingyu Gan[4], Yan Yin "Jenny" Cheung[5], Elvira D'Bastiani[6], and Gregory R Goldsmith[7]

[1]Department of Plant and Microbial Biology, University of Minnesota, St Paul, MN; [2]Department of Biology, Case Western Reserve University, Cleveland, OH; [3]Department of Biological Sciences, University of New Hampshire, Durham, NH; [4]Program in Ecology & Evolution, University of Wyoming, Laramie, WY; [5]Marine and Environmental Biology Section, Department of Biological Sciences, University of Southern California, Los Angeles, CA; [6]Department of Ecology and Evolutionary Biology, University of California–Los Angeles, Los Angeles, CA; [7]Schmid College of Science and Technology, Chapman University, Orange, CA

"*Presenting in my native language made me feel more comfortable, like I could be my authentic self. I believe that lowering the language barrier will help science hear a diversity of voices and help a diversity of people have their voices heard*".

This was expressed by a speaker at the 2024 Ecological Society of America's (ESA's) Annual Meeting, where we conducted an experimental special session (*SS 31 - Towards Equity in the Communication of Science: Harnessing the Power of AI for an Inclusive Tomorrow*) to evaluate artificial intelligence (AI) interpretation of scientific talks by native speakers of Cantonese, Mandarin Chinese, Portuguese, and Spanish into real-time running English captions.

One fundamental value of the ESA is to "[provide] the community of ecologists of diverse backgrounds, heritage, and career paths with a supportive home that advances their aspirations". Currently, the lingua franca at international science meetings is English, which certainly discourages or prevents many talented scientists from sharing their science, networking with colleagues, and otherwise fully participating in international conferences. Indeed, many avoid attending annual conferences altogether due to language barriers. Emerging AI technologies may soon make it possible to attract a diversity of new voices into science. ESA members are from all over the world: representing 88 countries as of 2023. By giving individuals the option to deliver presentations in the language they are most comfortable with through real-time AI interpretation, we might honor our collective ancestral voices and diversify the messages of science.

To assess the quality of AI interpretation through real-time subtitles for spoken content, we used a common presentation software and a paid AI interpretation service, in which we paired an Asian language with a Latin language of similar interpretation difficulty. After the talks, we conducted a survey and held an audience discussion assessing our experiences with the AI interpretation. Attendees of the session spoke several languages, and most were early-career individuals, both of which may not be reflective of ESA membership as a whole.

There was much to be excited about. Of the attendees who completed the survey, 84% (27/32) indicated they were likely to attend future sessions with interpretation subtitles, 84% (27/32) rated the continued exploration of AI interpretation as important, and 97% (29/30) agreed that "AI interpretation technologies make scientific conferences more inclusive for non-native English speakers". These results demonstrate support for AI interpretation as a means to enhance inclusion at scientific conferences, but there is more work to do to ensure that adequate consideration is given to how diverse individuals obtain and communicate information.

From our collective experience and discussion, we learned that AI interpretation is advancing rapidly, but is still not ready for widespread implementation: (1) The audience experienced cognitive overload with the lines of scrolling subtitles, which moved too quickly even with rehearsed pauses by the speakers; (2) As AI switched the grammatical structure from Asian languages to English, the text rearranged itself, which was difficult to follow. Configuring the AI to wait until after the speaker completes a sentence to display that sentence would fix the issue; (3) Because AI currently deciphers only one language at a time, it struggled to translate English scientific terms embedded in non-English content. This limitation forced speakers to use scientific terms in their native languages, which they found unnatural due to their having long since learned and adopted the use of certain scientific terms in English to align with the academic world; (4) AI needs to recognize different patterns of speech and dialects. For example, AI "understood" Puerto Rican Spanish less well than Colombian Spanish; (5) A text-to-speech option would not only support individuals with dyslexia and visual impairment but also reduce subtitle reader fatigue; and (6) Longer subtitles displayed less frequently would improve reader comprehension.

To amplify the diverse voices of scientists and to reduce language barriers, we look forward to the continued development of AI interpretation—and we are calling for software developers to partner with scientists from diverse backgrounds and to collaborate with scientific conferences that feature a wide range of speakers and topics to help train their AI. We also call for heightened attention to the development of AI interpretation for neurodivergent communicators and for signed languages so that a greater number of communication styles are considered when developing AI tools.

While AI needs to improve, our survey results show that we can promote inclusivity right now by enabling the real-time AI-generated subtitles that are already available in many presentation software packages.

We hope that the current generation of scientists will benefit from AI interpretation—and that the next generation of scientists won't have to ask themselves, "Is my English good enough yet for me to become a scientist?"