

Making Room for Speech Diversity

A 50 Year Retrospective of Speech Science and Technology through a Neurodivergent Lens

Rebecca Lietz
University of California, Santa Cruz

Jingjin Li
AImpower.org

Peiyao Liu
University of California, Santa Cruz

Jennifer Chien
Stanford University

Norman Makoto Su
University of California, Santa Cruz

Shaomei Wu
AImpower.org

AImpower.org

UC SANTA CRUZ | Baskin Engineering

HAI Stanford University Human-Centered Artificial Intelligence

Background

Models of Disability

Medical model

Social model

Models of Disability

	Medical model	Social model
Disability is …	an individual pathology	
Meaningful Interventions:	diagnosis, treatment, cure, prevention	
Advocate for …	individual correction	

Models of Disability

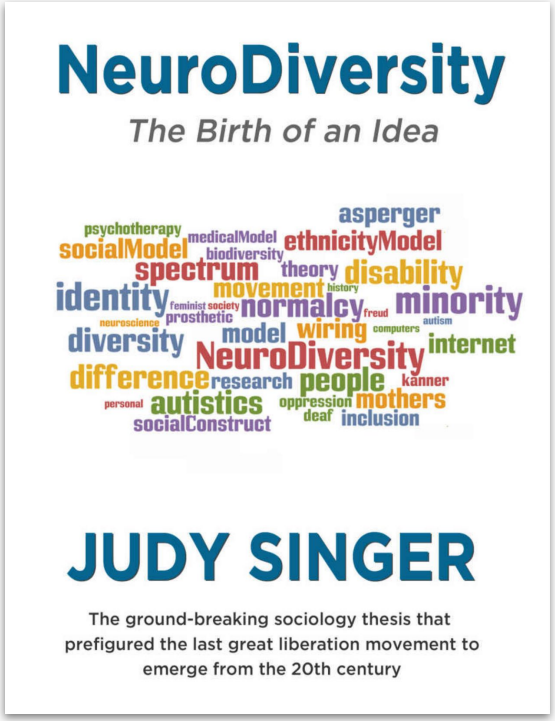
	Medical model	Social model
Disability is …	an individual pathology	incompatibility between body and the environment
Meaningful Interventions:	diagnosis, treatment, cure, prevention	making the environment more accessible
Advocate for …	individual correction	social and structural inclusion

Neurodiversity

“ Why not propose that just as biodiversity is essential to ecosystem stability, so neurodiversity may be essential for cultural stability? ”

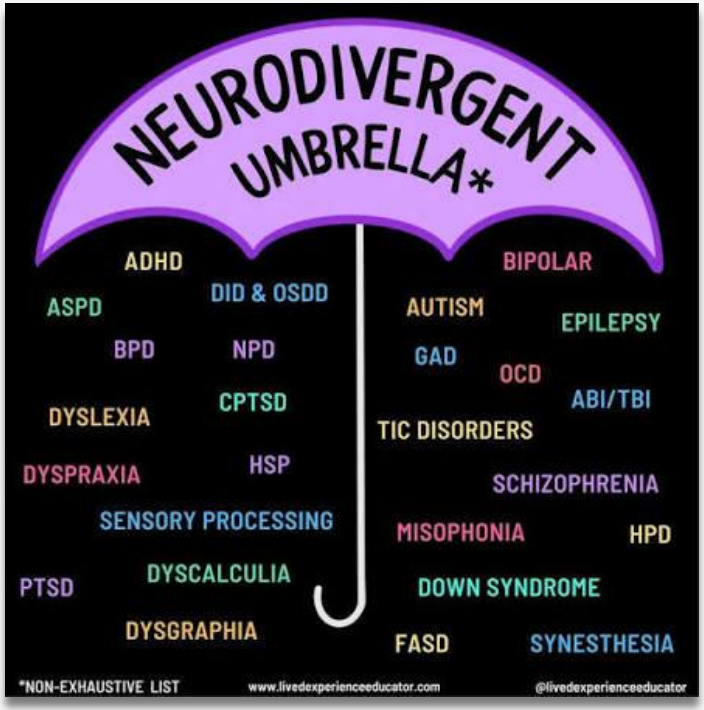
– Judy Singer





Neurodiversity

- Umbrella term
- Estimated to affect 10-15% population
- Neurodivergent speech traits are often non-degenerative and non-reversible



Method

Method: scoping review

Scope

InterSpeech (including ICSLP, EuroSpeech) and ICASSP

1976–2025

Partial match in the title, abstract, or full text for at least one of:

stutter , disfluen , dysfluen , disab , disord , atyp

1170 papers

Method: scoping review

Scope	Inclusion Criteria
InterSpeech (including ICSLP, EuroSpeech) and ICASSP **1976–2025** Partial match in the title, abstract, or full text for at least one of: **stutter , disfluen , dysfluen , disab , disord , atyp**	1. Studying human speech 2. Studying one or more of the following conditions: Stuttering, Attention Deficit Hyperactivity Disorder (ADHD/ADD), Autism Spectrum Disorder (ASD), Bipolar Disorder (BD), or Down Syndrome (DS)

1170 papers

117 papers

75	33	5	4
Interspeech	ICASSP	ICSLP	EuroSpeech

Method: scoping review

Scope	Inclusion Criteria	Manual categorization
InterSpeech (including ICSLP, EuroSpeech) and ICASSP **1976–2025** Partial match in the title, abstract, or full text for at least one of: **stutter , disfluen , dysfluen , disab , disord , atyp**	1. Studying human speech 2. Studying one or more of the following conditions: Stuttering, Attention Deficit Hyperactivity Disorder (ADHD/ADD), Autism Spectrum Disorder (ASD), Bipolar Disorder (BD), or Down Syndrome (DS)	• Main contribution • Intended use case • Data practice • Evaluation method • Language use

1170 papers

117 papers

75	33	5	4
Interspeech	ICASSP	ICSLP	EuroSpeech

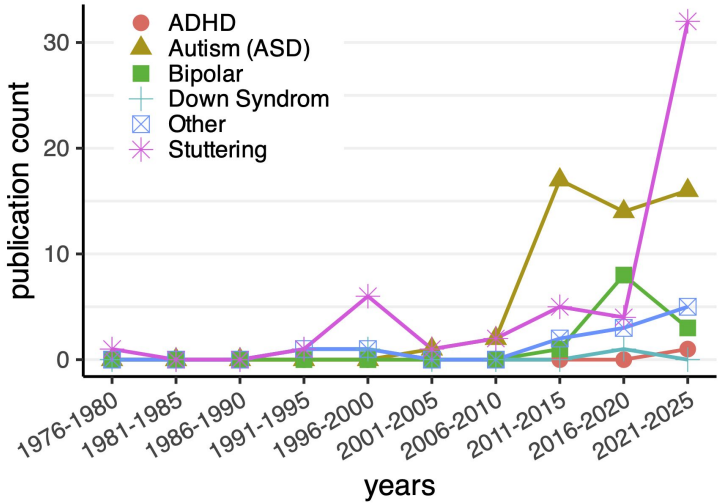
Access full dataset

zenodo.org/records/20754795

Key Findings

Key Findings #1

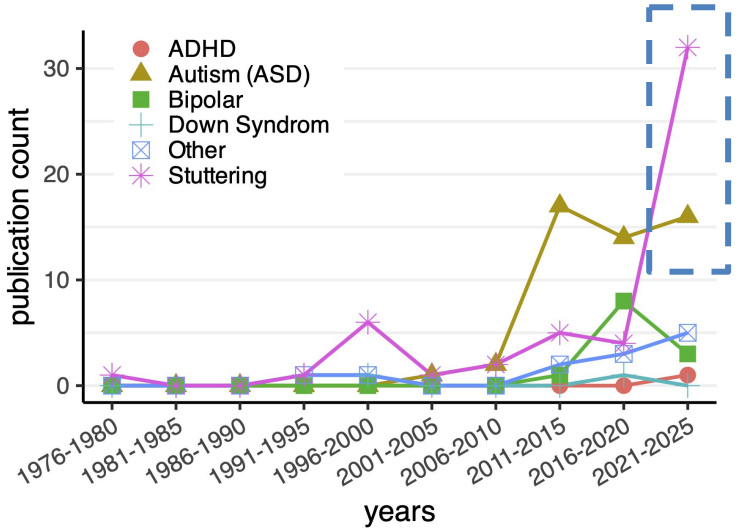
Increased Research Interests over Years



Number of publications covered in scoping review

Key Findings #1

Increased Research Interests over Years



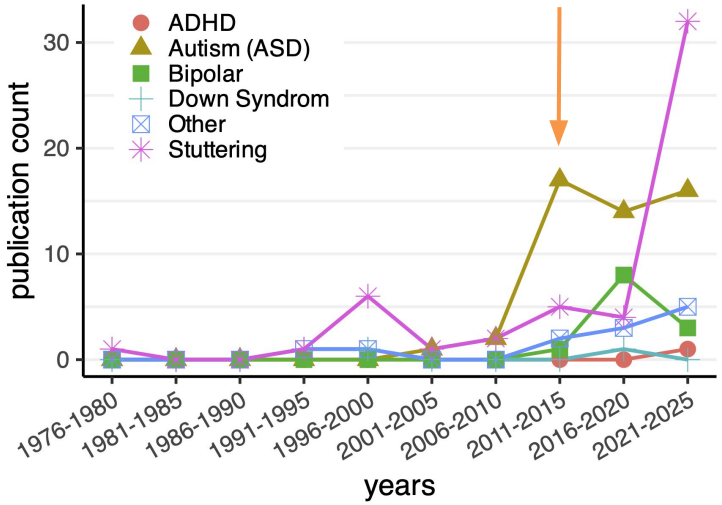
Stuttering and ASD became most popular conditions to study in recent years

Number of publications covered in scoping review

Key Findings #1

Increased Research Interests over Years

2013 ASD speech classification challenge

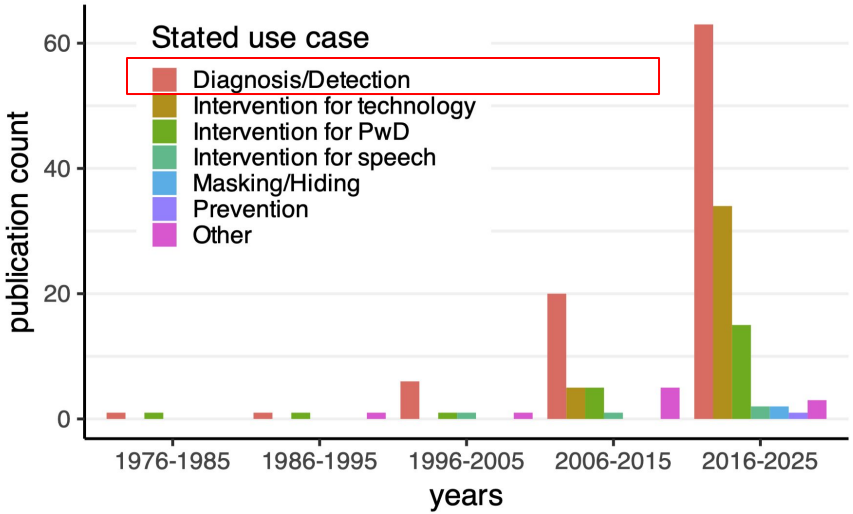


Sustained interests after the challenge

Number of publications covered in scoping review

Key Findings #2

Dominance of the medical model

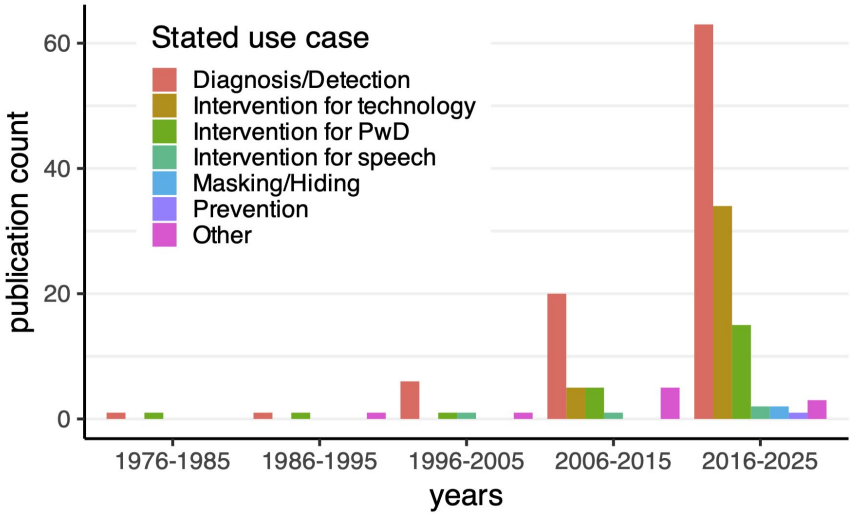


78% of all publications reviewed were motivated by the “diagnosis and detection” use case.

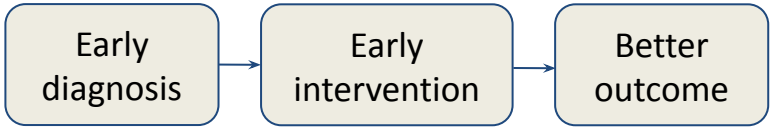
What’s the use case that motivated the study?

Key Findings #2

Dominance of the medical model



What's the use case that motivated the study?

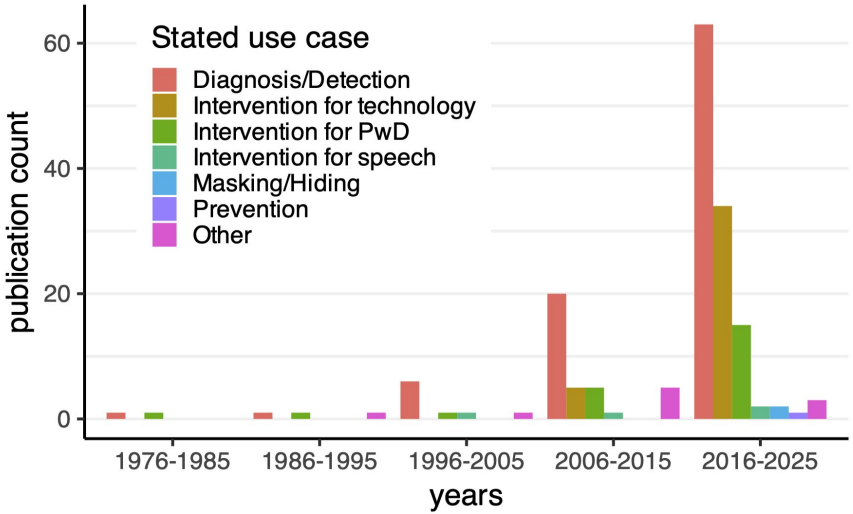


"it is critical to address speech disorder problems in early childhood, since delays in medical interventions can result in wide-ranging social and mental difficulties"

– A lightly supervised approach to detect stuttering in children's speech, InterSpeech 2018

Key Findings #2

Dominance of the medical model

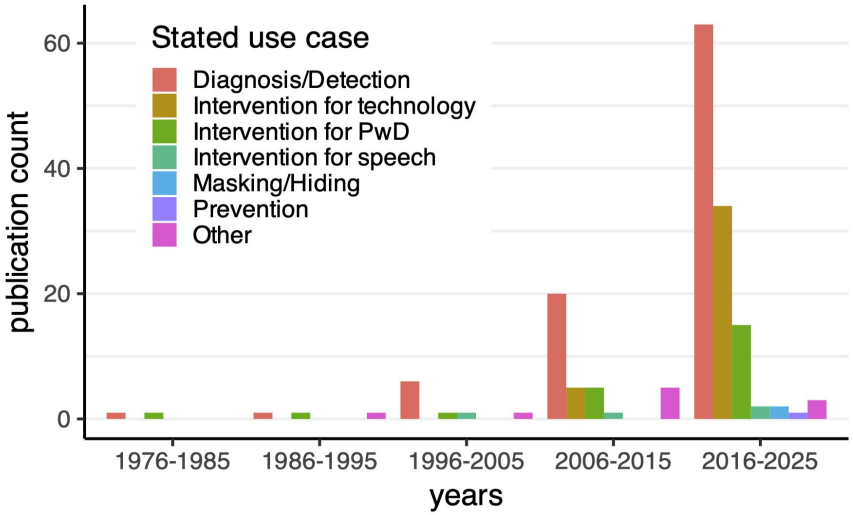


What's the use case that motivated the study?

- Neurodiversity is highly complex, speech is only a small part it.

Key Findings #2

Dominance of the medical model



What's the use case that motivated the study?

- Neurodiversity is highly complex, speech is only a small part it.
- Is intervention always necessary?

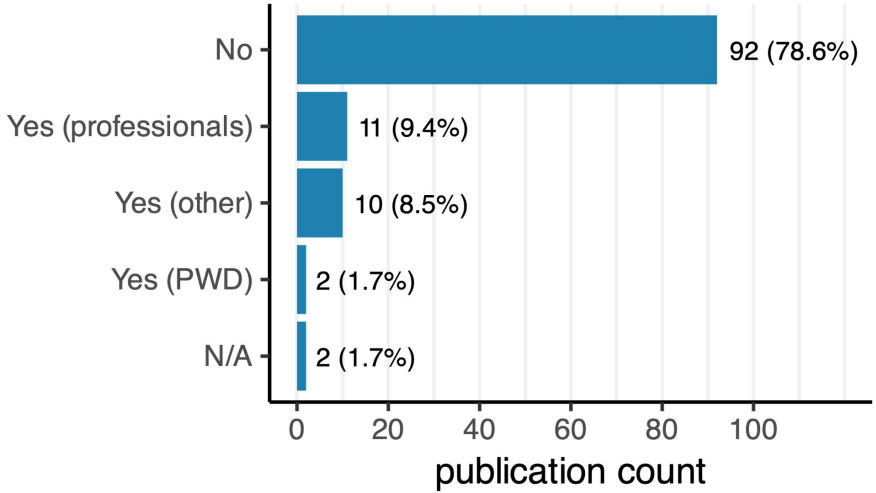
➔ 80% of children who stutter naturally stop stuttering as their brains develop (Smith and Weber 2025)

➔ Top predictors for persistent stuttering are **family history** and **environmental stress** (Yairi et al 1996)

➔ Receiving fluency shaping therapy has **no predictive power** (Yairi et al 1996)

Key Findings #3

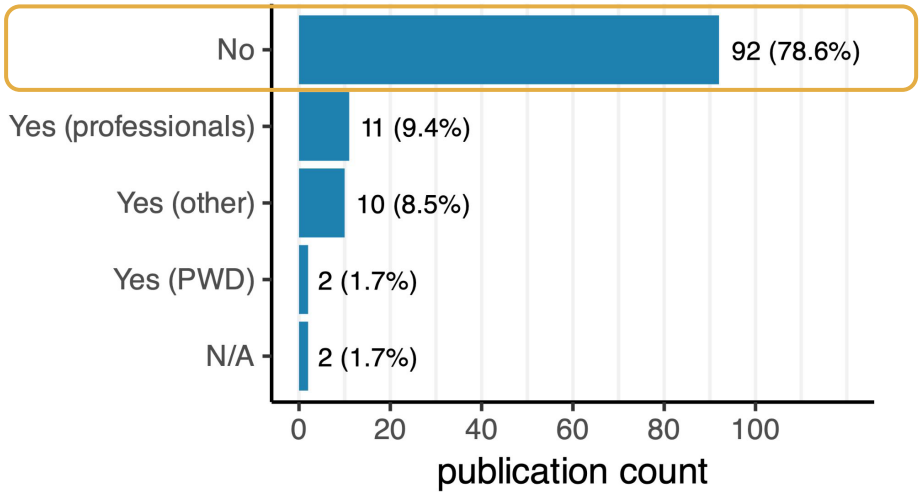
Minimum inputs from impacted communities



Who were involved in system evaluations?

Key Findings #3

Minimum inputs from impacted communities

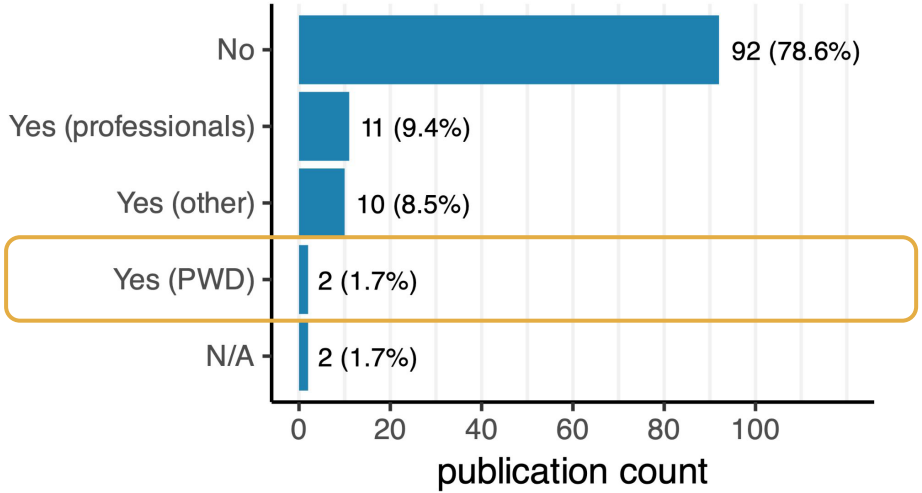


No human evaluation, evaluated with benchmarking datasets and static metrics

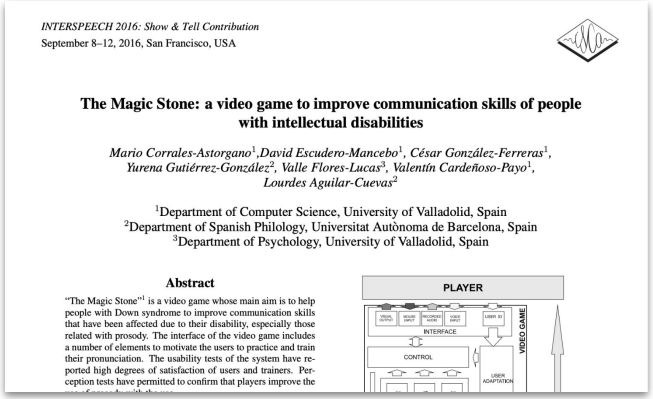
Who were involved in system evaluations?

Key Findings #3

Minimum inputs from impacted communities



Who were involved in system evaluations?



INTERSPEECH 2016: Show & Tell Contribution
September 8–12, 2016, San Francisco, USA

The Magic Stone: a video game to improve communication skills of people with intellectual disabilities

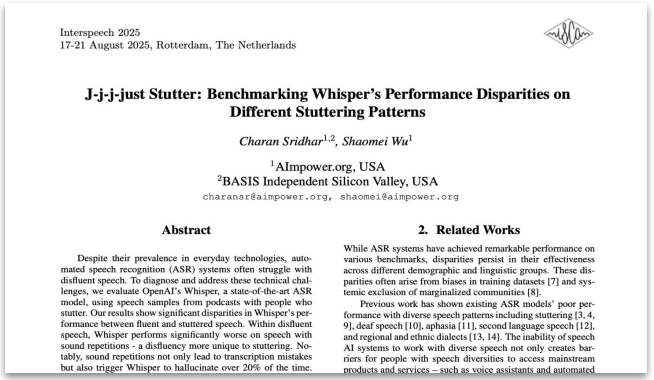
Mario Corrales-Astorgano[1], David Escudero-Mancebo[1], César González-Ferreras[1], Yurena Gutiérrez-González[2], Valle Flores-Lucas[3], Valentín Cardeñoso-Payo[1], Lourdes Aguilar-Cuevas[2]

[1]Department of Computer Science, University of Valladolid, Spain
[2]Department of Spanish Philology, Universitat Autònoma de Barcelona, Spain
[3]Department of Psychology, University of Valladolid, Spain

Abstract

"The Magic Stone"[1] is a video game whose main aim is to help people with Down syndrome to improve communication skills that have been affected due to their disability, especially those related with prosody. The interface of the video game includes a number of elements to motivate the users to practice and train their pronunciation. The usability tests of the system have reported high degrees of satisfaction of users and trainers. Perception tests have permitted to confirm that players improve the

Corrales-Astorgano *et al,* 2016



Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

J-j-j-just Stutter: Benchmarking Whisper's Performance Disparities on Different Stuttering Patterns

Charan Sridhar[1,2], Shaomei Wu[1]

[1]AImpower.org, USA
[2]BASIS Independent Silicon Valley, USA
charansr@aimpower.org, shaomei@aimpower.org

Abstract

Despite their prevalence in everyday technologies, automated speech recognition (ASR) systems often struggle with disfluent speech. To diagnose and address these technical challenges, we evaluate OpenAI's Whisper, a state-of-the-art ASR model, using speech samples from podcasts with people who stutter. Our results show significant disparities in Whisper's performance between fluent and stuttered speech. Within disfluent speech, Whisper performs significantly worse on speech with sound repetitions - a disfluency more unique to stuttering. Notably, sound repetitions not only lead to transcription mistakes but also trigger Whisper to hallucinate over 20% of the time.

2. Related Works

While ASR systems have achieved remarkable performance on various benchmarks, disparities persist in their effectiveness across different demographic and linguistic groups. These disparities often arise from biases in training datasets [7] and systemic exclusion of marginalized communities [8].

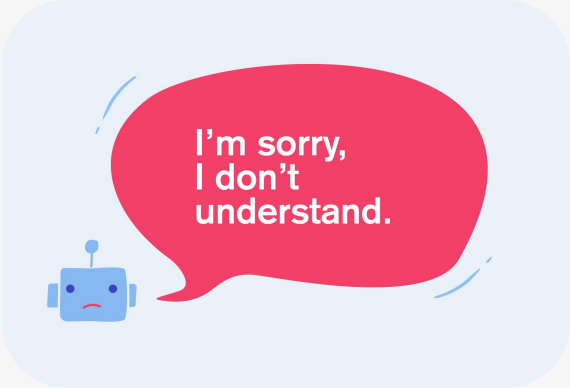
Previous work has shown existing ASR models' poor performance with diverse speech patterns including stuttering [3, 4, 9], deaf speech [10], aphasia [11], second language speech [12], and regional and ethnic dialects [13, 14]. The inability of speech AI systems to work with diverse speech not only creates barriers for people with speech diversities to access mainstream products and services – such as voice assistants and automated

Scridhar and Wu, 2025

Discussion

What can the InterSpeech community do

- Incorporate social model of disability
  - Investigating environmental and social barriers
  - Shifting attention from correction to accommodation



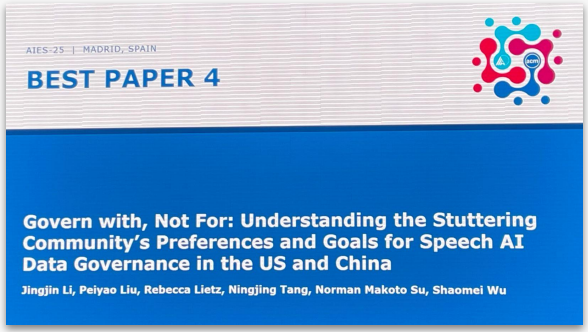
What can the InterSpeech community do

- Incorporate social model of disability
- Direct engagement with impacted people
  - Using HCI methods to evaluate how systems perform and benefit (or harm!) PWD
  - Having PWD as researchers and collaborators



What can the InterSpeech community do

- Incorporate social model of disability
- Direct engagement with impacted people
  - Using HCI methods to evaluate how systems perform and benefit (or harm!) PWD
  - Having PWD as researchers and collaborators



Li et al. AIES 2025



Tang et al. FAccT 2026



September 27th - October 1st, 2026

Speech AI for All: Challenging Deficit-based Approaches to Speech Science and Technology

at International Convention Centre (ICC) on Gadigal Country in Sydney, Australia

Questions?

Contact:
Shaomei@aimpower.org

Download full paper

bit.ly/neurodiversity-interspeech

Access review dataset

zenodo.org/records/20754795