

# Making Room for Speech Diversity: A 50 Year Retrospective of Speech Science and Technology through a Neurodivergent Lens

Rebecca Lietz<sup>1</sup>, Jingjin Li<sup>2</sup>, Peiyao Liu<sup>1</sup>, Norman Su<sup>1</sup>, Shaomei Wu<sup>2</sup>

<sup>1</sup>Computational Media, University of California Santa Cruz, USA

<sup>2</sup>AImpower.org, USA

rlietz@ucsc.edu, jingjin@aimpower.org, pliu62@ucsc.edu, normsu@ucsc.edu,  
shaomei@aimpower.org

## Abstract

Interspeech’s latest conference theme is “fair and inclusive speech science and technology.” With this increased interest in diversity in speech research and the goal to ensure such research will “benefit all individuals and communities,” we take this to be an opportune moment to critically and systematically examine papers on neurodivergent speech from 1976–2024 in Interspeech and ICASSP. Our content and discourse analyses of publications identify three key themes. Papers (1) take a medicalizing and interventionist approach to “correcting” neurodivergent traits, thereby reinforcing a deficit view of disability; (2) distance stakeholders by excluding them from the research process; and (3) contribute to the othering of PwD by using ableist language. With the benefit of hindsight, we propose pathways through which Interspeech can achieve its goal of celebrating speech diversity in its research and its papers.

**Index Terms:** neurodivergent speech, ableism, fairness, inclusivity, discourse analysis

## 1. Introduction

Human speech is highly variable. The human brain demonstrates a remarkable adaptability to a wide range of voices, accents, and speaking styles [1]. Speech technologies, in contrast, are catered to a narrow set of “typical” speech patterns, underperforming for individuals with speech diversities such as stuttering [2, 3], deaf speech [4], second language speech [5, 6], as well as regional vernaculars and ethnic dialects [7].

Recognizing this gap, the Interspeech community has shown an increasing interest in speech diversity research, reflected most recently in its 2025 conference theme of “*fair and inclusive speech science and technology*.” Given the growing number of publications contributing new models, datasets, and systems aiming to promote this theme, we believe it an opportune moment to critically examine – with the benefit of hindsight – the past and present of Interspeech to identify future opportunities to better promote inclusivity.

As a step toward this goal, our study systematically reviews existing papers on fair and inclusive speech technology through the lens of neurodiversity. As an umbrella term encompassing a wide range of conditions – such as autism, dyslexia, ADHD, and more recently, stuttering – neurodiversity is an apt lens to evaluate inclusivity in Interspeech for three reasons: (1) it affects a significant portion of the population – estimated at 15-20% [8]– and has frequently been the target for inclusion-oriented research efforts; (2) it often impacts one’s speech and communication functions in a non-degenerative and non-reversible way; and (3) as a social and political movement, neurodiversity has fostered rich discourse and scholarship on the design and evaluation of assistive and everyday technolo-

gies [9, 10, 11]. Leveraging neurodiversity as a case study, we conduct a critical examination of the existing body of work on speech diversities in the past 48 years. Our systematic review covers Interspeech and ICASSP – two of the largest speech science and technology conferences – publications from 1976 to 2024 that involve the speech of neurodiverse communities, aiming to uncover key trends, common practices, and opportunities for future, better inclusivity.

Through our content and discourse analysis, we observe that research in the context of neurodivergent speech often takes an interventionist and medicalizing approach, aiming to “cure” or “correct” neurodivergent traits in order to fit into the neurotypical norm. Other stakeholders, like clinicians or people with disabilities (PwD), are rarely involved in the research process, creating a distance between researchers and the populations they aim to support. The use of ableist language and misconceptions about neurodiverse conditions further contribute to the othering of PwD. We hope our scope review inspires reparative reflections within the Interspeech community, encouraging the field to explore new problems grounded in the experiences of people with speech diversities, foster partnership between researchers and end users, and cultivate anti-ableist mindsets around the research process and its impact.

## 2. Background & Related Work

### 2.1. Models of disability

Different models of disability have been developed to conceptualize disability and people with disabilities.

As a dominant framework in research and society, the **medical model of disability**, views disability as a biological deficit in need of interventions to prevent, cure, or mitigate [12]. While the medical model of disability has provided utilities in designing assistive technologies that address real-world issues encountered by PwD, it is increasingly pushed back by disability scholars and activists for framing disability as deterministically and biologically inferior and reinforcing social stigma around disability-related traits [11, 13].

The **social model of disability**, in contrast, argues disability arises from the way society is structured: an individual is disabled not by their physical condition but by environmental, attitudinal, and systemic factors [14]. Developed by disabled people, the social model of disability has been embraced by the disability community to recognize the value of disability experiences and advocate for societal changes that ensure full inclusion of PwD [11]. The stuttering community, for example, has leveraged the social model to emphasize the role of listeners and the communication environment in “constructing a disabled speaker” [15]. By celebrating stuttering gain and stuttering pride [16], people who stutter are finding value and strength

in their stuttering experiences while advocating for recognition and the right to stutter without stigma [12].

## 2.2. Neurodiversity movement

The neurodiversity movement originated in the late 1990s as an alternative to medicalized views of cognitive and communicative differences [17]. Initially an effort to reframe conditions such as autism, ADHD, and dyslexia as natural variations in human cognition rather than pathologies, the idea of neurodiversity has become a social and political movement for the rights and liberation of those living with developmental neurological conditions. Pushing back on medical interventions that seek to eliminate these natural variations, neurodiversity advocates emphasize the unique skills and advantages of different neurological traits and demand the acceptance, recognition, and the right to live authentically with these traits [18].

In recent years, the neurodiversity framework has expanded to include speech diversity such as stuttering [13]. This perspective empowers people who stutter to find meaning and joy in their stuttering experience while acknowledging their structural vulnerability that necessitates the accommodations from listeners – including speech technologies.

## 2.3. Technoableism and disfluent speech

While the social model of disability and the neurodivergent movement challenge the notion that disability is an unfortunate individual condition that must be “fixed,” ableism – the assumption that able-bodied and neurotypical experiences are superior – remains deeply embedded in society. Reflected in technological development, technoableism prevails in technologies designed to make PwD *appear* or *act* “normal” rather than remove structural barriers [10]. While these technologies may benefit some users, they often function as disability dongles – tools developed without meaningful input from disabled communities that fail to address their real-world needs [19].

Speech technologies often exemplify technoableism by treating fluent speech as the default and the ideal. Even products specially designed for people with speech disfluency often focus on eliminating disfluencies, either by altering the behavior of the speaker or by modifying their speech. Technologies that “fix” speakers, such as SpeechEasy<sup>1</sup>, aim to train their users to speak more fluently – often with significant effort. Technologies that modify the speech, such as Google’s Project Relate<sup>2</sup>, directly operate on “atypical” speech to make it more fluent.

Disfluent communities have increasingly resisted the ableist desire for fluent speech, challenging the medical necessity of speech restructuring for merely auditory aesthetics [20] and establishing an affirmative narrative around speech disfluency [16, 21]. Following the “**nothing about us without us**” principle from the disability justice movement [22], the stuttering community recently organized a public campaign against an AI-assisted fluency shaping app marketed by Samsung, demanding that the voices of people who stutter be heard rather than suppressed by speech technology<sup>3</sup>.

## 3. Methods & Data

We first identified papers about disfluent or atypical speech published at two large speech technology conferences, Interspeech

<sup>1</sup><https://speecheasy.com>

<sup>2</sup><https://sites.research.google/relate/>

<sup>3</sup><https://www.spacetostutter.org/impulse>

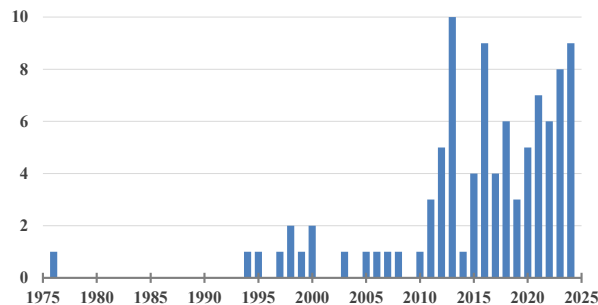


Figure 1: Histogram of publication year

and ICASSP, using the word stems ‘stutter,’ ‘disfluen,’ ‘dysfluen,’ ‘disab,’ ‘disord,’ and ‘atyp’ as search terms. Our IEEE Xplore search yielded 684 ICASSP papers. A Python script to extract metadata and search the ICSA archives yielded 823 papers published at Interspeech and its predecessors, EuroSpeech and ICSLP.

We reduced this initial corpus of 1199 papers to 130 papers by only including papers that focus on developmental neurodivergent conditions affecting speech: stuttering, Attention Deficit Hyperactivity Disorder (ADHD/ADD), Autism Spectrum Disorder (ASD), Bipolar Disorder (BPD), and Down Syndrome (DS). 35 papers were then excluded because they did not meet the inclusion criteria of being about speech and focusing on the conditions mentioned above. For example, some papers only mention stuttering as an example of disfluency in their abstract or use electromagnetic (but no speech) signals.

The finalized corpus of papers our team analyzed in-depth includes 95 papers published from 1976 to 2024 at Interspeech (61 papers), ICASSP (25 papers), ICSLP (5 papers), and EuroSpeech (4 papers). Within this corpus, there are an increasing number of publications addressing neurodivergence in speech technology starting in the 2010s (Figure 1), indicating a rising interest in addressing topics of neurodiversity and accessibility. Autism and stuttering were the most represented neurodivergent conditions, with 43 and 38 papers, respectively. Bipolar disorder (12 papers), Down syndrome (3 papers), and other conditions (8 papers) were less frequent. No papers about ADHD/ADD met all the inclusion criteria.

Our analysis critically examined each paper’s type and topic, main contribution, stated use case, intended audience and end users, data source, annotation method, evaluation metrics, and use of ableist language. Our analysis, drawing from both content and discourse methods, foregrounds not only the *content* of an individual paper but also how it *delivers* its content through words, tables, and imagery [23, 24]. We met weekly to discuss observations and questions. The corpus and evaluation rubric are provided in the supplemental material.

## 4. Results

The research in this corpus is largely technical and primarily aims to assist clinicians. This approach emphasizes the medicalization of neurodivergent conditions over the individuals themselves. The following sections detail how papers viewed neurodivergence through a medicalizing, distancing, and othering lens by aiming to “correct” neurodivergent traits, rarely involving key stakeholders like clinicians or PwD, and using ableist language. The Discussion proposes opportunities for more fair and inclusive speech science and technology.

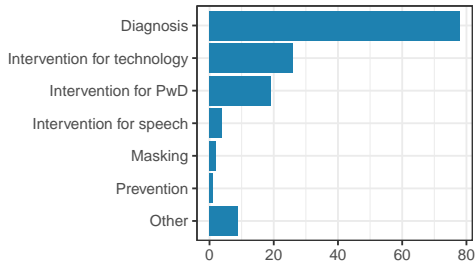


Figure 2: Use cases stated in papers

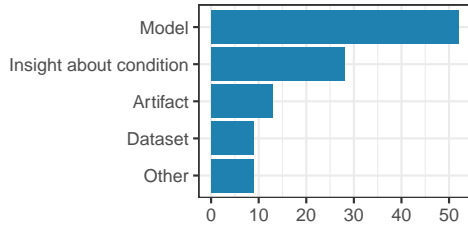


Figure 3: Main contributions claimed by papers

#### 4.1. Medicalizing conditions

The reviewed speech technology research often viewed neurodivergence through a medicalized lens. Figures 2 and 3 show the intended use cases and main contributions for the corpus.

As evident from Figure 2, diagnosis and detection is by far the most common use case (82%). The overall goal of these contributions is to make diagnosis and detection less labor-intensive for clinicians with little focus on PwD. Fewer papers focused on developing interventions for existing technologies (e.g., improving ASR for stuttered speech [25]), PwD (e.g., developing a symptom monitoring system to be used by PwD [26]), or speech (e.g., modifying speech to increase prosody [27]). Intervention use cases were also often tied to detection and diagnosis as the overarching goal.

Figure 3 illustrates the breakdown of main contributions claimed by papers. Over 50% of papers contributed a novel model or training method (“Model”) and ~30% generated new insights into the condition under study (“Insight about condition”), which includes speech patterns and characteristics of people with the condition. 13 papers developed or tested an artifact (“Artifact”) and 9 papers presented a new dataset (“Dataset”). With few exceptions, the people benefiting from these models, insights, and systems were clinicians working with PwD rather than PwD themselves. For example, [28, 29, 30, 31] propose systems to help clinicians and speech therapists count and annotate disfluencies to aid diagnosis of stuttering or ASD.

A small number of papers aimed at “treating” the symptoms of neurodivergent conditions by training PwD to better hide or mask them. For example, [32] proposes a method for the automatic discrimination and evaluation of soft/hard consonant articulation. They propose using this method to develop self-guided speech therapy interventions for people who stutter with the goal of helping them practice speech patterns that effectively mask their stutter.

Altogether, most papers tied their contributions to medical use cases such as diagnosis, symptom monitoring, or treat-

ment of neurodivergent conditions, which reinforces the medical model of disability.

#### 4.2. Distancing stakeholders

Similar to findings in other fields [33, 9], many papers in our corpus took a “distanced position”, where PwD were rarely involved, and evaluation metrics were usually limited to technical benchmarks. Contrary to the “nothing about us without us” principle [22], this “distanced position” has been criticized for undermining the agency of PwD and leading to interventions that are not suited to the needs of the community [34, 33].

Distancing is partially achieved by approaching the diagnosis of neurodivergent conditions from a purely technical perspective. In other words, papers are framed as contributing to (improved) algorithmic diagnosis and detection of “symptoms.” Such contributions are evaluated through quantitative benchmarks, such as accuracy, recall, or F1 scores. Without discounting the merits of this approach, we argue that the sole reliance on technical perspectives distances the reader and researcher from PwD, reducing them to a few algorithmic features. Indeed, only a few papers discuss the limitations of this computation-centered approach [35, 36, 37, 38, 39]. [40], for example, states that their approach to ASD detection using acoustic and text data “is not yet valuable as a method for diagnosing ASD, which still needs to be diagnosed by expert clinicians.” Limitations beyond technical capabilities, model architecture, and data availability were rarely mentioned.

Distancing is also evident in the lack of participation from PwD or practitioners in research problem formation, design, and evaluation. No paper mentioned anyone outside of the research team being involved in the problem formation, and very few recruited clinicians or PwD for evaluation [41, 29]. In most cases, researchers engaged PwD solely as data contributors, either directly through participant recruitment for data collection [31, 40, 32] or indirectly by utilizing existing datasets (e.g., SEP-28k, CPSD, PRIORI, USC CARE). In a few cases, when PwD’s data was sourced from third parties, it was unclear whether participants’ explicit consent was obtained. For example, [42] sourced videos of children with autism experiencing a meltdown from YouTube by searching terms like “autism child shouting” but does not mention seeking consent from the uploader or people in the videos. Similarly, [43], one of the most highly cited papers in our corpus, presents a dataset of stuttered speech from podcasts with people who stutter without mention of seeking consent from the podcast hosts or guests.

Though human evaluation is not common in tech-focused venues such as Interspeech, a few notable examples in our corpus show it is possible for such papers to be more inclusive—to involve other stakeholders, acknowledge participants’ needs and contributions, and discuss limitations inherent in a technology-centered approach. [29] and [41] both aim to identify instances of stuttered speech. [29] presents a computer tool that assists speech-language therapists (SLP) in counting and annotating clients’ disfluencies. To assess whether the tool aligns with the needs of speech therapists, the researchers conduct a user study with therapists who evaluate and correct the automatically generated transcriptions and annotations. Based on insights from a follow-up questionnaire, they were able to implement improvements detailed in [30]. [41], on the other hand, remove stuttered speech from test samples using a proposed frame-level technique for stutter detection. They recruit three raters from Amazon mTurk to assess the presence of stuttering and the quality of the resultant transcripts from each sam-

ple. Both of these publications make largely technical contributions while incorporating human evaluation techniques.

Overall, the reviewed papers showed little end-user involvement. This, in addition to the tendency to reduce the complex experiences of neurodivergent individuals to a small number of metrics, distances neurodivergent communities from research that is precisely meant to benefit them.

### 4.3. Othering neurodivergence

Neurodivergence was often singled out as a hard edge case for technology. Under this lens, data from neurodivergent individuals served to showcase the robustness of technical contributions beyond “normal” cases. [44], for example, presents a method for speaker diarization, specifically for child-adult conversations. While the introduction is framed around autism assessment in children and the researchers use recordings from ASD assessment sessions as data, the main focus is on the opportunities and limitations of adversarial learning. ASD is merely the problem domain. This approach to neurodivergence is especially reflected in the language of papers.

Ableist language was very common in the reviewed papers. Neurodivergent conditions were often described as “abnormal” and “pathological.” The behavioral characteristics of people with the condition were termed “symptoms,” “impairments,” and “errors,” illustrating a deficit approach to neurodivergence. In contrast, data collected from control groups consisting of people without neurodivergent conditions were often labeled as “normal” and “correct”.

In some cases, introduction sections contain misleading information about the condition of interest to underpin the value of their contribution. For example, a number of papers about stuttering [45, 46, 47] perpetuate the notion that “if [stuttering] remains untreated or undiagnosed [it] develops into a lifetime disorder” [47]. While this idea ostensibly serves to justify stuttering detection research, it is misleading – 80% of children who stutter will naturally grow out of it as their brains develop, regardless of interventions [48]. People who stutter into adulthood often experienced negative reactions toward stuttering during their younger years, driven by social stigma and the pressure to “correct” it as soon as possible. In fact, current SLP research and clinical practice has advocated for the long-term mental-health benefits of disfluency affirmative approaches [49, 21]. While there is real value in developing effective stuttering detection models (e.g., making speech AI systems more accessible for people who stutter), the incorporation of misleading medical information only serves to further reinforce the social stigma associated with neurodivergence.

## 5. Discussion

There has been an increased interest in the speech science and technology research community over the past years in working with diverse groups and promoting inclusive, fair practices and contributions in this domain. However, as our review shows, the literature on speech and neurodivergent conditions does not yet reflect this framing. Researchers commonly take a medicalizing, distancing, and othering approach to PwD. In order to move towards more fair and inclusive practices in the field, we make two constructive suggestions to the research community.

*Increased direct engagement with PwD during the research process:* PwD have so far primarily served as data contributors, and their first-hand perspectives have rarely been consulted when forming research questions and evaluating the benefits

and limitations of technical contributions. When those outside of the research team are involved, they are mostly clinicians, thereby reinforcing a medical over social perspective on neurodivergence. We propose more direct engagement with PwD throughout the research process. PwD are often best positioned to lead the research related to their own experiences. For example, in [50], the authors – several of whom are people who stutter – leveraged their connections within the stuttering community to collect and precisely annotate 50 hours of Mandarin stuttered speech from 70 speakers. Their work in creating the first and largest Mandarin stuttered speech dataset contributed an invaluable resource for the research and development of stuttering-friendly speech technology [3]. Researchers can also engage with PwD by utilizing methods from user research and participatory design. For example, during the formative phase, they can conduct focus groups or surveys with PwD to generate relevant, meaningful, and valid research questions. In the evaluation phase, researchers could supplement technical benchmarks with human evaluators, similar to [41]. Incorporating these methods not only yields richer data, enables technology to accurately reflect users’ needs and make more impactful contributions, but also builds trust and empathy between researchers and PwD communities.

*Reflect on framing and social impact:* Many of the reviewed papers use ableist language, treat PwD as a medical special case, and neglect discussions about the limitations of their chosen methods. Far from promoting inclusivity, such framings serve to reinforce a deficit model that views PwD as merely an interesting object of technical study. We therefore encourage researchers to critically examine how their work is framed and the social impact of their work. Authors could include a limitations section in their papers, a practice common among many academic disciplines, but usually missing from the reviewed papers. In addition, authors could more clearly state the expected impact of their contribution on PwD – not only on clinicians’ work. Instead of framing research in terms of “curing” or “fixing” neurodiversity, researchers should explore how their work can embrace values such as acceptance and autonomy.

Lastly, we acknowledge that our analysis is one with the benefit of hindsight. Each paper is inescapably a product of its time, as is our own. Our intention is not to call out individual papers but to take this moment to encourage the speech research community to reflect upon the past and present. Researchers have the opportunity to shape a more just future where speech technology supports PwD through academic research and industry partnerships. We hope that by embracing diverse perspectives, assessing the social impact of one’s research, and building strong relationships with PwD communities, the Inter-speech community can truly achieve its goals to “celebrate and incorporate the vast speech diversity both within and between individuals, as well as within and between languages.”

## 6. References

- [1] S. E. Brennan and M. F. Schober, “How listeners compensate for disfluencies in spontaneous speech,” *J. Mem. Lang.*, vol. 44, no. 2, pp. 274–296, 2001.
- [2] C. Lea, Z. Huang, J. Narain, L. Tooley, D. Yee, D. T. Tran, P. Georgiou, J. P. Bigham, and L. Findlater, “From user perceptions to technical improvement: Enabling people who stutter to better use speech recognition,” in *Proc. of CHI’23*, 2023.
- [3] Q. Li and S. Wu, “Towards fair and inclusive speech recognition for stuttering: Community-led chinese stuttered speech dataset creation and benchmarking,” in *Proc. of EA CHI’24*, 2024.

- [4] A. Glasser, "Automatic speech recognition services: Deaf and hard-of-hearing usability," in *Proc. of EA CHI'19*, 2019, p. 1–6.
- [5] S.-E. Kim, B. R. Chernyak, O. Seleznova, J. Keshet, M. Goldrick, and A. R. Bradlow, "Automatic recognition of second language speech-in-noise," *JASA Express Letters*, vol. 4, no. 2, 2024.
- [6] K. Wenzel and G. Kaufman, "Designing for harm reduction: Communication repair for multicultural users' voice interactions," in *Proc. of CHI'24*, 2024.
- [7] A. Koenecke, A. Nam, E. Lake, J. Nudell, M. Quartey, Z. Mengesha, C. Toups, J. R. Rickford, D. Jurafsky, and S. Goel, "Racial disparities in automated speech recognition," *PNAS*, vol. 117, no. 14, pp. 7684–7689, 2020.
- [8] N. Doyle, "Neurodiversity at work: a biopsychosocial model and the impact on working adults," *Br Med Bull*, vol. 135, no. 1, pp. 108–125, 2020.
- [9] R. M. Williams, L. Boyd, and J. E. Gilbert, "Counterinterventions: a reparative reflection on interventionist hci," in *Proc 2023 SIGCHI Conf Hum Factor Comput Syst*, 2023, pp. 1–11.
- [10] A. Shew, *Against Technoableism: Rethinking Who Needs Improvement*. W. W. Norton, 2023.
- [11] A. Hamraie and K. Fritsch, "Crip technoscience manifesto," *Catalyst: Feminist, Theory, Technoscience*, vol. 5, pp. 1–34, 04 2019.
- [12] C. Constantino, P. Campbell, and S. Simpson, "Stuttering and the social model," *J. Commun. Disord.*, vol. 96, p. 106200, 2022.
- [13] C. D. Constantino, "What can stutterers learn from the neurodiversity movement?" *Semin. Speech Lang.*, vol. 39, no. 4, pp. 382–396, Sep. 2018.
- [14] M. Oliver, "The social model of disability: thirty years on," *Disabil. Soc.*, vol. 28, no. 7, pp. 1024–1026, 2013.
- [15] J. S. Pierre, "The construction of the disabled speaker: Locating stuttering in disability studies," *Can. J. Disabil. Stud.*, vol. 1, pp. 1–21, 2012.
- [16] Patrick Campbell and Christopher Constantino and Sam Simpson, Eds., *Stammering Pride and Prejudice: Difference not Defect*. J&R Press, 2019.
- [17] J. Singer, "Why can't you be normal for once in your life? from a problem with no name to the emergence of a new category of difference," *Disability discourse*, pp. 59–67, 1999.
- [18] P. Jaarsma and S. Welin, "Autism as a natural human variation: reflections on the claims of the neurodiversity movement," *Health Care Anal.*, vol. 20, no. 1, pp. 20–30, Mar. 2012.
- [19] L. Jackson, A. Haagaard, , and R. Williams, "Disability dongle," <https://blog.castac.org/2022/04/disability-dongle/>, 2022, accessed: 2025-02-03.
- [20] Z. Richter, "On stuttering activism and resistance," in *Stammering Pride and Prejudice: Difference not Defect*, P. Campbell, C. Constantino, and S. Simpson, Eds. J & R Pres, 2020.
- [21] C. D. Constantino, "Fostering positive stuttering identities using stutter-affirming therapy," *Lang Speech Hear Serv Sch*, vol. 54, no. 1, pp. 42–62, 2023.
- [22] J. I. Charlton, *Nothing About Us Without Us: Disability Oppression and Empowerment*, 1st ed. UC Press, 1998.
- [23] N. Phillips and C. Hardy, *Discourse Analysis: Investigating Processes of Social Construction*, 2nd ed. Sage Publications, 2002.
- [24] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, 4th ed. SAGE Publications, 2019.
- [25] D. Mujtaba, N. R. Mahapatra, M. Arney, J. S. Yaruss, C. Herring, and J. Bin, "Inclusive asr for disfluent speech: Cascaded large-scale self-supervised learning with targeted fine-tuning and data augmentation," in *Interspeech 2024*, 2024, pp. 1275–1279.
- [26] J. Gideon, E. M. Provost, and M. McInnis, "Mood state prediction from speech of varying acoustic quality for individuals with bipolar disorder," in *Proc. of ICASSP'16*, 2016, pp. 2359–2363.
- [27] J. P. van Santen, L. Black, G. Cohen, A. B. Kain, E. Klappers, T. Mishra, J. de Villiers, and X. Niu, "Applications of computer generated expressive speech for communication disorders," in *EUROSPEECH 2003*, 2003, pp. 1657–1660.
- [28] S. Alharbi, M. Hasan, A. J. H. Simons, S. Brumfitt, and P. Green, "A lightly supervised approach to detect stuttering in children's speech," in *Interspeech 2018*, 2018, pp. 3433–3437.
- [29] P. A. Heeman, A. McMillin, and J. S. Yaruss, "Computer-assisted disfluency counts for stuttered speech," in *Interspeech 2011*, 2011, pp. 3013–3016.
- [30] P. A. Heeman, R. Lunsford, A. McMillin, and J. S. Yaruss, "Using clinician annotations to improve automatic speech recognition of stuttered speech," in *Interspeech 2016*, 2016, pp. 2651–2655.
- [31] H. Kothare, V. Ramanarayanan, O. Roesler, M. Neumann, J. Liscombe, W. Burke, A. Cornish, D. Habberstad, A. Sakallah, S. Markuson, S. Kansara, A. Faerman, Y. Bensidi-Slimane, L. Fry, S. Portera, D. Suendermann-Oeft, D. Pautler, and C. Demopoulos, "Investigating the interplay between affective, phonatory and motoric subsystems in autism spectrum disorder using a multimodal dialogue agent," in *Interspeech 2021*, 2021, pp. 1967–1971.
- [32] K. Ochi, K. Mori, and N. Sakai, "Automatic evaluation of soft articulatory contact for stuttering treatment," in *Interspeech 2018*, 2018, pp. 1546–1550.
- [33] J. Mankoff, G. R. Hayes, and D. Kasnitz, "Disability studies as a source of critical inquiry for the field of assistive technology," in *Proc. of ASSETS'10*, 2010, p. 3–10.
- [34] K. Spiel, C. Frauenberger, O. Keyes, and G. Fitzpatrick, "Agency of Autistic Children in Technology Research—A Critical Literature Review," *ACM TOCHI*, vol. 26, no. 6, pp. 1–40, Dec. 2019.
- [35] M. Jouaiti and K. Dautenhahn, "Dysfluency classification in stuttered speech using deep learning for real-time applications," in *Proc. of ICASSP'22*, 2022, pp. 6482–6486.
- [36] S. P. Bayerl, D. Wagner, I. Baumann, F. Hönig, T. Bocklet, E. Nöth, and K. Riedhammer, "A stutter seldom comes alone – cross-corpus stuttering detection as a multi-label problem," in *Interspeech 2023*, 2023, pp. 1538–1542.
- [37] S. Khorram, J. Gideon, M. McInnis, and E. M. Provost, "Recognition of depression in bipolar disorder: Leveraging cohort and person-specific knowledge," in *Interspeech 2016*, 2016, pp. 1215–1219.
- [38] T. Chaspari, E. M. Provost, A. Katsamanis, and S. Narayanan, "An acoustic analysis of shared enjoyment in eca interactions of children with autism," in *Proc. of ICASSP'12*, 2012, pp. 4485–4488.
- [39] S. P. Whiteside, R. A. Varley, T. Phillips, and H. Garety, "The production of real and non-words in adult stutterers and non-stutterers: an acoustic study," in *Proc. of ICSLP'00*, vol. 2, 2000, pp. 515–518.
- [40] S. Cho, M. Liberman, N. Ryant, M. Cola, R. T. Schultz, and J. Parish-Morris, "Automatic detection of autism spectrum disorder in children using acoustic and text features from brief natural conversations," in *Interspeech 2019*, 2019, pp. 2513–2517.
- [41] J. Harvill, M. Hasegawa-Johnson, and C. D. Yoo, "Frame-level stutter detection," in *Interspeech 2022*, 2022, pp. 2843–2847.
- [42] J. Casebeer, H. Sarker, M. Dhuliawala, N. Fay, M. Pietrowicz, and A. Das, "Verbal protest recognition in children with autism," in *Proc. of ICASSP'18*, 2018, pp. 301–305.
- [43] C. Lea, V. Mitra, A. Joshi, S. Kajarekar, and J. P. Bigham, "Sep-28k: A dataset for stuttering event detection from podcasts with people who stutter," in *Proc. of ICASSP'21*, 2021, pp. 6798–6802.
- [44] R. Lahiri, M. Kumar, S. Bishop, and S. Narayanan, "Learning domain invariant representations for child-adult classification from speech," in *Proc. of ICASSP'20*, 2020, pp. 6749–6753.
- [45] K. S. S. Motepalli, V. Narasinga, H. Pathuri, H. Khan, S. Mahesh, A. K. Abraham, and A. K. Vuppala, "Stuttering detection application," in *Interspeech 2023*, 2023, pp. 680–681.
- [46] T. Kourkounakis, A. Hajavi, and A. Etemad, "Detecting Multiple Speech Disfluencies Using a Deep Residual Network with Bidirectional Long Short-Term Memory," in *Proc. of ICASSP'20*, 2020, pp. 6089–6093.
- [47] S. M. Fosnot, "Vowel development of /i/ and /u/ in 15-36 month old children at risk and not at risk to stutter," in *EUROSPEECH 1997*, 1997, pp. 1051–1054.
- [48] A. Smith and C. M. Weber, "How stuttering develops: The multifactorial dynamic pathways theory," *JSLHR*, vol. 60 9, pp. 2483–2505, 2017.
- [49] V. Sisskin, "Disfluency-affirming therapy for young people who stutter: Unpacking ableism in the therapy room," *Lang Speech Hear Serv Sch*, vol. 54, no. 1, pp. 114–119, 2023.
- [50] R. Gong, H. Xue, L. Wang, X. Xu, Q. Li, L. Xie, H. Bu, S. Wu, J. Zhou, Y. Qin *et al.*, "As-70: A mandarin stuttered speech dataset for automatic speech recognition and stuttering event detection," in *Interspeech 2024*, 2024, pp. 5098–5102.