



# Quantitative Speech Assessment in Ataxia—Consensus Recommendations by the Ataxia Global Initiative Working Group on Digital-Motor Markers

Adam P. Vogel<sup>1,2,3</sup> · Anna Sobanska<sup>4</sup> · Anoopum Gupta<sup>5</sup> · Gessica Vasco<sup>6</sup> · Marcus Grobe-Einsler<sup>7,8</sup> · Susanna Summa<sup>6</sup> · Stephanie Borel<sup>9</sup>

Accepted: 19 October 2023  
© The Author(s) 2023

## Abstract

Dysarthria is a common and debilitating symptom of many neurodegenerative diseases, including those resulting in ataxia. Changes to speech lead to significant reductions in quality of life, impacting the speaker in most daily activities. Recognition of its importance as an objective outcome measure in clinical trials for ataxia is growing. Its viability as an endpoint across the disease spectrum (i.e. pre-symptomatic onwards) means that trials can recruit ambulant individuals and later-stage individuals who are often excluded because of difficulty completing lower limb tasks. Here we discuss the key considerations for speech testing in clinical trials including hardware selection, suitability of tasks and their role in protocols for trials and propose a core set of tasks for speech testing in clinical trials. Test batteries could include forms suitable for remote short, sensitive and easy to use, with norms available in several languages. The use of artificial intelligence also could improve accuracy and automaticity of analytical pipelines in clinic and trials.

**Keywords** Speech · Dysarthria · Patient-reported outcome · Clinical trials · Acoustics

## Introduction

Ataxia leads to changes in speech [1–6]. These changes worsen as the disease progresses [7] and can improve with effective treatment [8–10]. Subtle changes can even occur prior to disease onset [3]. Broadly, the ataxia speech phenotype is characterised by a reduced rate of speech, imprecise production of consonants, distorted vowels [1, 11], dysphonia [12] and hypernasality [13]. The dysarthric profile also includes poor vocal control (incoordination of pitch and loudness) and diminished breath support [14]. These deficits, in part, are the result of mis-timed and inaccurately targeted articulatory movements resulting in slower and slurred speech [15, 16]. Combined, dysarthria resulting from ataxia impacts naturalness and intelligibility. Dysarthria can lead to daily disadvantage and prevent simple communication exchanges from occurring (e.g. signalling food preferences, need for toileting). It can also trigger altered self-identity [17] and impede or prevent both social and professional interactions [18], leading to daily disadvantage, producing social marginalisation [19] and underemployment [18]. Seventy percent of people with a communication disorder are unemployed or in the lowest income brackets [20].

✉ Adam P. Vogel  
vogela@unimelb.edu.au

<sup>1</sup> Centre for Neuroscience of Speech, The University of Melbourne, Melbourne, Australia

<sup>2</sup> Division of Translational Genomics of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Germany & Center for Neurology, University Hospital Tübingen, Tübingen, Germany

<sup>3</sup> Redenlab Inc., Melbourne, Australia

<sup>4</sup> Department of Clinical Neurophysiology, Institute of Psychiatry and Neurology, Warsaw, Poland

<sup>5</sup> Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

<sup>6</sup> Bambino Gesù Children's Hospital, IRCCS, 00050 Rome, Italy

<sup>7</sup> German Center for Neurodegenerative Diseases, Bonn, Germany

<sup>8</sup> Department of Neurology, University Hospital Bonn, Bonn, Germany

<sup>9</sup> Sorbonne Université, Paris Brain Institute (ICM Institut du Cerveau), AP-HP, INSERM, CNRS, University Hospital Pitié-Salpêtrière, F-75013 Paris, France

Despite the debilitating daily impact of dysarthria, objective measurement of speech is rarely addressed in ataxia clinical practice, clinical trials and research. This may be due to the relatively limited influence of speech on overall scores in commonly used disease severity scales (i.e. SARA, ICARS, mFARS) [21]. On the other hand, speech is considered key feature for measurement in patients (<https://www.ataxia.org/ataxiapfdd/>), becoming the most important quality of life when individuals become non-ambulant [22, 23]. When it is examined, published cohort studies are often small and restricted to specialised centres which are informative, but their generalizability is limited. There also is limited published longitudinal and natural history data and inadequate evidence-based interventions for speech [24, 25]. In clinical disciplines involved in managing ataxia, decisions about disease-related dysarthria are mainly based on subjective assessment of speech symptoms. Yet a strong body of evidence has consistently shown that more precise speech measurement can increase sensitivity of clinical decisions and provide greater information on the nature of neurological change as well as determining the potential benefits of pharmaceutical and behavioural therapies aimed at forestalling symptomatic progression [9, 10].

For over a century, speech disorders have been described by what the listener can subjectively hear, despite early attempts at quantification [26]. Advances in signal processing, cloud computing and hardware and remote data capture provide an opportunity to exploit the intrinsic utility of speech as a marker of disease progression and treatment response (see Fig. 1). Digital technologies have the potential to surpass clinical judgement for accuracy and accessibility as they can yield objective outcomes and can be administered in the clinic or home. Here we outline considerations for use of speech as a marker of performance and quality of life in clinical trials. We also provide recommendations for protocol design, hardware and software selection, features of importance for describing change and disease state, links to patient reported outcomes, existing datasets and ongoing natural history studies.

## Hardware and Software Selection

Audio files are typically recorded and stored for post testing analysis<sup>1</sup>. A microphone is used to capture speech and it is an important determinant of signal quality. Microphone quality and suitability for recording speech is determined by its frequency response and range, directionality, polar

pattern and power supply (see [27, 28] for details). Signal fidelity is also influenced by file format, sampling rate, physical elements of the device and noise. Captured audio can be stored in a lossless format (e.g. .wav), preserving all aspects of the signal within the predetermined sampling rate. Elements of speech important for communication fall within the first 5KHz, with a minimum suggested sampling rate twice the maximum frequency of interest [29]. Thus, to ensure adequate fidelity, it is recommended that files be sampled at a minimum of 16KHz, 8-bit quantization with post recording down-sampling completed if necessary. Noise can enter the signal through several sources including the environment (e.g. other speakers, air conditioning), low-quality or poorly insulated wiring, or inappropriate positioning of the microphone (e.g. too close/far, or variable distance to source). Portability, ease of use and budget also guide decisions about utility of recording set ups (see [30–32] for example comparisons between devices).

Specific hardware recommendations are not included in this review as technology is constantly evolving; each study or testing environment will likely benefit from a configuration tailored to the specific use case. For example, for large community samples, a bring-your-own device may be selected by investigators, or similarly, provisioned portable devices may be selected for use in a clinical trial that only interact with specific microphone solutions. Readers are encouraged to read published comparative studies [30–32] or tutorials [27, 28] for more information to assist with hardware selection.

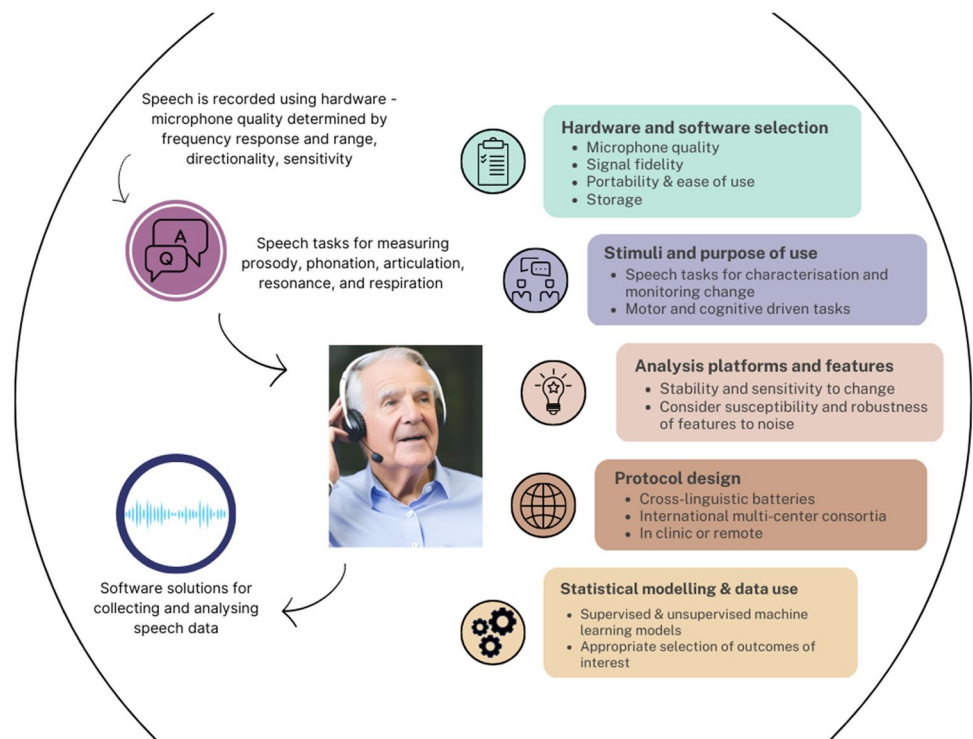
Software, apps and digital interfaces used to capture speech should allow for modification or set minimum standards when collecting audio. A significant proportion of speech capture software now provides cloud storage of audio rather than on-device storage. Remote storage can assist in data dissemination, access to analytic platforms and secure storage. It is recommended that systems used to capture and store audio encrypt data at rest and in transit are not directly identifiable beyond the audio file itself and meet multi-region privacy regulations outlined in the Health Insurance Portability and Accountability Act (HIPAA) and Europe's General Data Protection Regulations (GDPR).

## Stimuli and Use

Speech protocols should be theory driven and influenced by strong empirical evidence supporting their use. The motivation for testing also shapes protocol design. For example, assessment for characterisation is not necessarily well suited to detecting change from treatment [33, 34]. Characterisation of speech deficits requires an in-depth investigation describing specific impairments (e.g. voice quality), their impact on function (e.g. intelligibility) and their influence on participation (e.g. quality of

<sup>1</sup> There are circumstances where a microphone is not used (e.g. electroglottography) and where data are not stored on the device (e.g. contemporaneous analysis of filtered components of the waveform or spectrogram, on device).

**Fig. 1** Considerations for using speech as a clinical outcome measure



life). Characterisation protocols may include speaking in a variety of contexts, across multiple tasks, and include listener ratings alongside patient reported outcomes. Batteries that support investigation of key speech domains of prosody, voice, articulation, resonance and respiration are appropriate for phenotyping studies. Tasks could focus on connected speech (e.g. conversation) to assess articulation, prosody and resonance, and challenge activities such as diadochokinesis (DDK) (i.e. PATA) for timing, coordination and articulation, and maximum phonation time (MPT) (for breath support and voice quality). Maximal challenge tasks such as MPT or DDK and oral motor mobility tasks such as those in cranial nerve exams may be appropriate to measure performance across severity levels. They can test a speaker's maximal abilities, provide data on severity and are independent of language. Beyond singular measures of severity, global features like intelligibility and naturalness of spontaneous speech bring together information on all speech subsystems and are a strong reflection of daily life difficulties. Measures of intelligibility can be derived via standardised clinician perceptual scores or via composite measures of multiple acoustic features. Speech to text tools can also provide an estimate of intelligibility; however, these estimates are dynamic as they are built on models that are constantly evolving. It is possible to rely on these tools in circumstances where measures are based on a specific version of the model; that model can be "frozen" for persistent use [35]. Accuracy of speech to text

models also vary based on sex, accent, age and language of the speaker [36].

Testing for the purpose of detecting change in performance can be achieved through a brief, easy-to-administer and complete battery that is motivating and provides capacity for comparison over multiple time points (see [34] for discussion on tracking change in speech studies). Performance should remain stable in the absence of true change, and change when central nervous system function is compromised, through disease or physiology (e.g. fatigue) [37].

We know that speech is sensitive to disease in ataxia (see Table S1 for exemplar studies); however, it is rare for other influencing factors to be considered in study design. Speech changes with fatigue [37], repeated application [33], depression [38], altered feedback [34], the role of the assessor [39], the duration of the sample [40], phonetic context [41] and emotional states like boredom [42]. The influence these factors exert on speech production highlights the need for informed protocol design when the aim is monitoring change. Further, recognition that cerebellar disorders can lead to concomitant cognitive deficits [43] alongside motor dysfunction dictates the need for speech protocols to include simple, brief tasks that fit along a continuum of motor/cognitive complexity [44]. A similar model of assessment has been applied to other neurodegenerative diseases with motor and cognitive decline (e.g. Huntington's disease [45] and Fronto-temporal dementia [46]). Protocol establishment should be developed alongside intrinsic properties of

methods for analysing data and the features they yield. These include listener-based judgement, standardised assessments for measuring aspects of speech (e.g. [47]), instrumental assessments (e.g. electromagnetic articulography) or acoustic analysis.

### Analysis Platforms and Features

To establish the suitability of tasks (and analysis algorithms) for tracking change, they should be subjected to both stability and sensitivity challenges [33, 41]. Stability can be evaluated by eliciting speech repeatedly over brief and extended inter-recording intervals. This is designed to examine susceptibility and robustness of tasks and features to change. It is important to interrogate error or noise arising from technological issues relating to equipment or biological change like diurnal variability, altered motivation or fatigue. Following establishment of task and feature stability (the absence of change), sensitivity needs to be considered. Tasks and features may be stable because they are truly robust to noise, or they may simply be insensitive to change and therefore unsuitable for tracking change. Sensitivity can be estimated through challenges like sustained wakefulness [37], noise [44] or disease itself compared to a norm [5, 48].

Unlike a decade ago, there are now a plethora of software solutions for collecting and analysing speech data. When selecting appropriate digital resources for speech, there are data security, quality and usability features to consider. Ensure data are secure, encrypted at rest and in transit, are not stored alongside any personally identifiable information and are not altered (e.g. compressed) before storage. If using normative data provided by a software provider, check its veracity and suitability for comparison with your own dataset. There are reputable software options available from academic and commercial entities as well as normative datasets.

### Protocol Design

Batteries for the assessment of dysarthric speech have been developed in some languages, such as the MonPaGe battery (French) [48] and the Bogenhausen Dysarthria Scales (BoDyS) (German) [49]. But the challenge for improving research in ataxia is now to develop trans-linguistic batteries that can be used as biomarkers in international multicentric studies. Such protocols include language-independent tasks like prolonged vowel production and syllable repetition. Although there is considerable overlap between sites, investigators and batteries, the ad hoc approach to study design for each study or each language do not allow for multi-centre or inter-pathology comparison. The scientific and clinical community need to develop all together a core of the protocol that would be short, sensitive and easy to use, with norms available in several languages. There are some

exemplar initiatives bringing protocols together including the SpeechATAXIA project established within the Ataxia Global Initiative (<https://ataxia-global-initiative.net/projects/speech-ataxia-a-multinational-multilanguage-consortia-for-speech-in-hereditary-ataxias/>), the Friedreich ataxia Clinical Outcomes Study (FA-COMS) run by the Friedreich's Ataxia Research Alliance (<https://www.curefa.org/clinical-trials-active-enrolling/clinical-outcome-measures-in-friedreich-s-ataxia-a-natural-history-study>) and the new FA Global Clinical Consortium (FA-GCC) which combines FA-COMS and EFACTS (the European Friedreich's Ataxia Consortium for Translational Studies).

Speech studies can be run face to face in the clinic or remotely at home. Data can be collected on specialised audio equipment or consumer grade devices. Users can bring their own device (BYOD) to studies or use provisioned setups where hardware are provided by investigators. BYOD and remote testing can be advantageous in some settings and may provide freedom of users to complete tests when and where they choose. There is also an ability for investigators to collect data in what is perceived to be more ecologically valid testing conditions, such as in the home during daily activities. The latter leads to legitimate concerns around privacy and data use. Out-of-clinic recordings can also be hindered by reduced sound quality through non-provisioned devices or background noise for example.

### Potential Application of Machine Learning and Data-Driven Statistical Models

Artificial intelligence (AI) and big data analysis are methods that may enhance our ability to identify symptom onset or monitor disease progression in ataxia. Attempts to expand its use in diagnosis are underway [50]. The purpose is not to consider each single digital parameter as a biomarker but to define all the relevant information contained in the speech signal and use them as a parsimonious subset as determined by machine learning (ML) and deep learning (DL) algorithms. Learning feature representations is a central tenet of deep learning—the model can learn patterns directly from the audio time series that are informative for downstream tasks such as disease classification or severity estimation. Machine learning models can be trained in a supervised or unsupervised manner. In supervised learning, the sample data is already labelled, and it is used to train a classification or regression model. Then unlabelled data is given to this trained system for labelling (e.g. classify) based on the features. In unsupervised learning, the training set is not labelled. The system itself learns the structure of the data, for example to identify clusters or latent factors. In both the methods, selection of features plays an important role, as does the sample data dimension used to train the model. There are some ML studies seeking to separate



ataxic speakers from healthy speakers [51–54]; however, the value of this exercise is diminished by the knowledge that ataxia is a multi-faceted disease group requiring multi-modal assessments for diagnosis. An alternative or additional and potentially more valuable use of ML for speech draws on communication outcomes that are meaningful for patients and clinicians, such as intelligibility and naturalness [55]. This approach treats speech as an outcome in its own right, in addition to its role as a subcomponent of a diagnostic workup. In addition to those papers cited, we can gain insight into AI utility from other neurological disorders with similar symptoms [36, 54, 56]. As mentioned, binary or ternary classifiers are commonly used to distinguish between the healthy and pathological conditions [56, 57]. Often these discriminative models apply very simple feed-forward artificial neural networks (ANN) and support vector machines (SVM) [56, 57]. There are also studies that use binary or ternary classifiers to discriminate different levels of dysarthria severity using the Mahalanobis distance and reaching 95% or higher accuracy in splitting groups [51]. Other binary algorithms examples are linear discriminant analysis (LDA) and k-nearest neighbour but with lower accuracy [54]. As is the case with other behavioural markers, adding sensitivity beyond binary outcomes (e.g. adding levels of intelligibility) can lead to decreases in accuracy [58]. Some recent examples of hierarchical machine-learning model (combination of machine and deep learning algorithms) revealed promising results in ataxic groups [55, 59–61]. It is reasonable to assume AI will have role in future clinical practice, but it is important to understand its current limitations; for example, AI requires suitable and sufficient data [57].

## Conclusion

Speech disorder caused by hereditary ataxia triggers altered self-identity, and impedes social and professional interactions, leading to daily disadvantage, producing social marginalisation and underemployment. These changes typically worsen as the disease progresses but may improve with treatment. Subtle changes can even occur prior to diagnosis. The centrality of speech in daily life highlights its importance in clinical care and as a marker of brain health. We have provided clear information on the practical and theoretical factors driving protocol design, data collection, features of interest and links to meaningfulness for stakeholders.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12311-023-01623-4>.

**Author Contribution** A.V. and S.B. wrote the main manuscript text. A.S., A.G., G.V., M.G.-E. and S.S revised the draft.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions A.P.V. was supported by National Health and Medical Research Council, Australia (1135683), Australian Research Council (220100253).

**Data Availability** Not applicable.

## Declarations

**Ethical Approval** Not applicable

**Competing Interests** A.P.V. is Chief Science Officer of Redenlab Inc., a speech biomarker company.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Folker JE, Murdoch BE, Cahill LM, Delatycki MB, Corben LA, Vogel AP. Dysarthria in Friedreich's ataxia: a perceptual analysis. *Folia Phoniatr Logop.* 2010;62(3):97–103.
2. Noffs G, Perera T, Kolbe SC, Shanahan CJ, Boonstra FMC, Evans A, et al. What speech can tell us: a systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmun Rev.* 2018;17(12):1202–9.
3. Vogel AP, Magee M, Torres-Vega R, Medrano-Montero J, Cyngler MP, Kruse M, et al. Features of speech and swallowing dysfunction in pre-ataxic spinocerebellar ataxia type 2. *Neurology.* 2020;95(2):e194–205.
4. Vogel AP, Rommel N, Oettinger A, Horger M, Krumm P, Kraus E-M, et al. Speech and swallowing abnormalities in adults with POLG associated ataxia (POLG-A). *Mitochondrion.* 2017;37:1–7.
5. Noffs G, Boonstra FMC, Perera T, Butzkueven H, Kolbe SC, Maldonado F, et al. Speech metrics, general disability, brain imaging and quality of life in multiple sclerosis. *Eur J Neurol.* 2021;28(1):259–68.
6. Noffs G, Boonstra FMC, Perera T, Kolbe SC, Stankovich J, Butzkueven H, et al. Acoustic speech analytics are predictive of cerebellar dysfunction in multiple sclerosis. *Cerebellum.* 2020;
7. Rosen KM, Folker JE, Vogel AP, Corben LA, Murdoch BE, Delatycki MB. Longitudinal change in dysarthria associated with Friedreich ataxia: a potential clinical endpoint. *J Neurol.* 2012;259(11):2471–7.
8. Vogel AP, Skarrat J, Castles J, Synofzik M. Video game-based speech rehabilitation for reducing dysarthria severity in adults with degenerative ataxia. *Eur J Neurol.* 2016;23:227.
9. Vogel AP, Stoll LH, Oettinger A, Rommel N, Kraus E-M, Timmann D, et al. Speech treatment improves dysarthria in multisystemic ataxia: a rater-blinded, controlled pilot-study in ARSACS. *J Neurol.* 2019;266(5):1260–6.
10. Yiu EM, Tai G, Peverill RE, Lee KJ, Croft KD, Mori TA, et al. An open-label trial in Friedreich ataxia suggests clinical benefit with high-dose resveratrol, without effect on frataxin levels. *J Neurol.* 2015;1-10

11. Brendel B, Synofzik M, Ackermann H, Lindig T, Schölderle T, Schöls L, et al. Comparing speech characteristics in spinocerebellar ataxias type 3 and type 6 with Friedreich ataxia. *J Neurol*. 2015;262(1):21–6.
12. Vogel AP, Wardrop MI, Folker JE, Synofzik M, Corben LA, Delatycki MB, et al. Voice in Friedreich ataxia. *J Voice*. 2017;31(2):243.e9–e19.
13. Poole ML, Wee JS, Folker JE, Corben LA, Delatycki MB, Vogel AP. Nasality in Friedreich ataxia. *Clinical Linguistics & Phonetics*. 2015;29(1):46–58.
14. Schalling E, Hartelius L. Speech in spinocerebellar ataxia. *Brain Lang*. 2013;127(3):317–22.
15. Folker JE, Murdoch BE, Cahill LM, Delatycki MB, Corben LA, Vogel AP. Kinematic analysis of lingual movements during consonant productions in dysarthric speakers with Friedreich's ataxia: a case-by-case analysis. *Clinical Linguistics & Phonetics*. 2011;25(1):66–79.
16. Folker JE, Murdoch BE, Cahill LM, Rosen KM, Delatycki MB, Corben LA, et al. Differentiating impairment levels in temporal versus spatial aspects of linguopalatal contacts in Friedreich's ataxia. *Mot Control*. 2010;14(4):490–508.
17. Clarke P, Black SE. Quality of life following stroke: negotiating disability, identity, and resources. *J Appl Gerontol*. 2005;24(4):319–36.
18. Gibilisco P, Vogel AP. Friedreich ataxia. *BMJ*. 2013;347
19. Vogel AP. Speech disorder is an invisible form of disability. In: Gibilisco P, editor. *Design for All*. 11. New Delhi, India: Design for All Institute of India; 2016. p. 31–9.
20. ABS. *Australians Living with Communication Disability*. Australia Bureau of Statistics; 2017.
21. Rummey C, Harding IH, Delatycki MB, Tai G, Rezende T, Corben LA. Harmonizing results of ataxia rating scales: mFARS, SARA, and ICARS. *Ann Clin Transl Neurol*. 2022;9(12):2041–6.
22. Thomas-Black G, Dumitrascu A, Garcia-Moreno H, Vallortigara J, Greenfield J, Hunt B, et al. The attitude of patients with progressive ataxias towards clinical trials. *Orphanet J Rare Dis*. 2022;17(1):1.
23. Pandolfo M. Neurologic outcomes in Friedreich ataxia: study of a single-site cohort. *Neurol Genet*. 2020;6(3):e415.
24. Vogel AP, Folker JE, Poole ML. Treatment for speech disorder in Friedreich ataxia and other hereditary ataxia syndromes. *Cochrane Database Syst Rev*. 2014;10:CD008953.
25. Vogel AP, Graf LH, Magee M, Schöls L, Rommel N, Synofzik M. Home-based biofeedback speech treatment improves dysarthria in repeat-expansion SCAs. *Annals of Clinical and Translational Neurology*. 2022;n/a(n/a)
26. Hiller F. A study of speech disorders in friedreich's ataxia. *Arch Neurol Psychiatr*. 1929;22(1):75–90.
27. Vogel AP, Reece H. Recording speech: methods and formats. In: Ball M, editor. *Manual of Clinical Phonetics*. 1st ed. Routledge; 2021.
28. Vogel AP, Morgan AT. Factors affecting the quality of sound recording for speech and voice analysis. *International Journal of Speech-Language Pathology*. 2009;11(6):431–7.
29. Nyquist H. Certain topics in telegraph transmission theory. *Proc IEEE*. 2002;90(2):280–305.
30. Vogel AP, Maruff P. Comparison of voice acquisition methodologies in speech research. *Behav Res Methods*. 2008;40(4):982–7.
31. Vogel AP, Rosen KM, Morgan AT, Reilly S. Comparability of modern recording devices for speech analysis: smartphone, landline, laptop, and hard disc recorder. *Folia Phoniatri Logop*. 2014;66(6):244–50.
32. Noffs G, Cobler-Lichter M, Perera T, Kolbe S, Butzkueven H, Boonstra F, et al. Plug-and-play microphones for recording speech and voice with smart devices. *medRxiv*. 2023;
33. Vogel AP, Fletcher J, Snyder PJ, Fredrickson A, Maruff P. Reliability, stability, and sensitivity to change and impairment in acoustic measures of timing and frequency. *J Voice*. 2011;25(2):137–49.
34. Vogel AP, Maruff P. Monitoring change requires a rethink of assessment practices in voice and speech. *Logopedics Phoniatrics Vocology*. 2014;39(2):56–61.
35. Isaev DY, Vlasova RM, Martino MD, Stephen CD, Schmahmann JD, Sapiro G, et al. Uncertainty of vowel predictions as a digital biomarker for ataxic dysarthria. *Cerebellum*. in press
36. Schultz BG, Tarigoppula VSA, Noffs G, Rojas S, van der Walt A, Grayden DB, et al. Automatic speech recognition in neurodegenerative disease. *International Journal of Speech Technology*. 2021;
37. Vogel AP, Fletcher J, Maruff P. Acoustic analysis of the effects of sustained wakefulness on speech. *J Acoust Soc Am*. 2010;128(6):3747–56.
38. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry*. 2012;72(7):580–7.
39. Zraick RI, Gentry MA, Smith-Olinde L, Gregg BA. The effect of speaking context on elicitation of habitual pitch. *J Voice*. 2006;20(4):545–54.
40. Zraick RI, Birdwell KY, Smith-Olinde L. The effect of speaking sample duration on determination of habitual pitch. *J Voice*. 2005;19(2):197–201.
41. Caverlé MWJ, Vogel AP. Stability, reliability, and sensitivity of acoustic measures of vowel space: a comparison of vowel space area, formant centralization ratio, and vowel articulation index. *The Journal of the Acoustical Society of America*. 2020;148(3):1436–44.
42. Cowie R, McGuiggan A, McMahon E, Douglas-Cowie E. Speech in the process of becoming bored. *Barcelona: Proceedings of 15th International Congress of Phonetic Sciences*; 2003.
43. Schmahmann JD. Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *The Journal of Neuropsychiatry and Clinical Neurosciences*. 2004;16(3):367–78.
44. Vogel AP, Fletcher J, Maruff P. The impact of task automaticity on speech in noise. *Speech Comm*. 2014;65:1–8.
45. Chan JCS, Stout JC, Vogel AP. Speech in prodromal and symptomatic Huntington's disease as a model of measuring onset and progression in dominantly inherited neurodegenerative diseases. *Neurosci Biobehav Rev*. 2019;107:450–60.
46. Vogel AP, Poole ML, Pemberton H, Caverlé MW, Boonstra FM, Low E, et al. Motor speech signature of behavioral variant frontotemporal dementia: refining the phenotype. *Neurology*. 2017;89(8):837–44.
47. Yorkston KM, Beukelman DR. *Assessment of Intelligibility of dysarthric speech*. Austin, TX: Pro-Ed; 1984.
48. Laganaro M, Fougeron C, Pernon M, Leveque N, Borel S, Fournet M, et al. Sensitivity and specificity of an acoustic- and perceptual-based tool for assessing motor speech disorders in French: the MonPaGe-screening protocol. *Clin Linguist Phon*. 2021;35(11):1060–75.
49. Nicola F, Ziegler W, Vogel M. The Bogenhausener Dysarthria Scales (BODYDYS): an instrument for clinical diagnostic of dysarthria. *Forum Logopädie*. 2004;18:14–22.
50. Schultz BG, Joukhadar Z, del Mar QM, Nattala U, Noffs G, Rojas S, et al. The classification of neurodegenerative disease from acoustic speech data. *Research Square*. 2021; Preprint
51. Paja MS, Falk TH. Automated dysarthria severity classification for improved objective intelligibility assessment of spastic dysarthric speech. *Portland, OR: Interspeech*; 2012. p. 62–5.
52. Narendra NP, Alku P. Automatic assessment of intelligibility in speakers with dysarthria from coded telephone speech using glottal features. *Comput Speech Lang*. 2021;65:101117.

53. Kadi KL, Selouani SA, Boudraa B, Boudraa M. Fully automated speaker identification and intelligibility assessment in dysarthria disease using auditory knowledge. *Biocybernetics and Biomedical Engineering*. 2016;36(1):233–47.
54. Kim J, Kumar N, Tsiartas A, Li M, Narayanan SS. Automatic intelligibility classification of sentence-level pathological speech. *Comput Speech Lang*. 2015;29(1):132–44.
55. Vogel AP, Maruff P, Reece H, Carter H, Tai G, Schultz BG, et al. Clinically meaningful metrics of speech in neurodegenerative disease: quantification of speech intelligibility and naturalness in ataxia. *medRxiv*; 2023. 2023.03.28.23287878
56. Rudzicz F. editor Phonological features in discriminative classification of dysarthric speech. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing. 2009;2009:19–24.
57. Schultz BG, Joukhadar Z, Nattala U, Quiroga MM, Bolk F, Vogel AP. Chapter 1 - Best practices for supervised machine learning when examining biomarkers in clinical populations. In: Moustafa AA, editor. *Big Data in Psychiatry & Neurology*. Academic Press; 2021. p. 1–34.
58. Selouani SA, Dahmani H, Amami R, Hamam H. Using speech rhythm knowledge to improve dysarthric speech recognition. *International Journal of Speech Technology*. 2012;15(1):57–64.
59. Tartarisco G, Bruschetta R, Summa S, Ruta L, Favetta M, Busà M, et al. Artificial intelligence for dysarthria assessment in children with ataxia: a hierarchical approach. *IEEE Access*. 2021;9:166720–35.
60. Vattis K, Luddy AC, Ouillon JS, Eklund NM, Stephen CD, Schmahmann JD, et al. Sensitive quantification of cerebellar speech abnormalities using deep learning models. *medRxiv*; 2023. 2023.04.03.23288094
61. Grobe-Einsler M, Faber J, Taheri A, Kybelka J, Raue J, Volkening J, et al. SARASpeech—feasibility of automated assessment of ataxic speech disturbance. *npj Digital Medicine*. 2023;6(1):43.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.