

©Copyright 2021

Jessica Amaral

**Investigating Perceptual Subgroups in Speakers with Ataxic Dysarthria:
An Auditory Free- Classification Approach**

Jessica Amaral

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2021

Committee:

Kristie Spencer (Chair)

Kaitlin Lansford

Tanya Eadie

Edythe Strand

Program Authorized to Offer Degree:

Speech and Hearing Sciences

University of Washington

Abstract

Investigating Perceptual Subgroups in Speakers with Ataxic Dysarthria:

An Auditory Free- Classification Approach

Jessica Amaral

Chair of the Supervisory Committee:

Professor Kristie Spencer

Speech and Hearing Sciences

Ataxic dysarthria has presented with considerable heterogeneity in the presentation of speech characteristics. Converging evidence supports the existence of subgroups, specifically related to the instability and inflexibility of motor patterns (Hartelius et al., 2000) as a possible explanation of this variability. To further examine the alignment of the speech characteristics of ataxic dysarthria with the instability/inflexibility framework, 23 graduate student listeners participated in an auditory free classification task (Clopper, 2008) and a guided classification task. Listeners grouped 15 speakers with ataxic dysarthria based on their judgment of the most salient perceptual characteristics during two speaking tasks: alternating motion rates (AMR) and connected speech (1-2 sentences). Listener ratings were then compared with *a priori* expert

determinations of speakers who fit the instability subgroup profile and the inflexibility subgroup profile. Results of both the free classification and guided classification listening paradigms provided definitive evidence of subgroups, particularly for the AMR task, in the context of excellent interrater reliability. This study supports growing evidence of the existence of instability and inflexibility subgroups in ataxic dysarthria and serves as a proof of concept for use of the auditory free-classification paradigm in dysarthria subtype research.

Table of Contents

Abstract.....	
List of Figures.....	
List of Tables.....	
Introduction.....	
Methods.....	
Participants.....	
Speakers.....	
Speaker Samples.....	
Perceptual Characterization of Speaker Samples.....	
Listeners.....	
Perceptual Classification Tasks.....	
Data Preparation and Analysis.....	
Results.....	
Reliability.....	
Free Classification Task.....	
Cluster Analysis.....	
Guided Classification Results.....	
Discussion.....	
Confirmed Use of AMRs for Subgroup Differentiation.....	
Possible Use of Connected Speech for Subgroup Differentiation	
Theoretical Justification of Ataxic Dysarthria Subgroups.....	
The Auditory Free Classification Approach.....	
Continued Investigation of Subgroups within Ataxic Dysarthria is Warranted.....	
References.....	
Appendix A: Listener Task PowerPoint Slides.....	

List of Figures

Figure 1: Screenshot of the grid system used for the free-classification task.....

Figure 2: Example of the completed free-classification grid from one listener.....

Figure 3: Screenshot of Free Classification task layout.....

Figure 4: Screenshot of the Guided Classification task layout.....

Figure 5: Dendrogram with the two-cluster solution for the AMR task.....

Figure 6: Multidimensional scaling plot reflecting the two-cluster solution for the AMR task.....

Figure 7: Dendrogram with a three-cluster solution for the Connected speech task.....

Figure 8: Multidimensional scaling plot reflecting the three-cluster solution for the Connected Speech task.....

Figure 9: Percentage match between gold standard expert classification and listener ratings during the guided classification task.....

List of Tables

Table 1: Summary of salient speech characteristics aligning with instability and inflexibility patterns of ataxic dysarthria.....

Table 2: Summary of speaker features.....

Table 3: Demographics and dysarthria characteristics of the speakers.....

Table 4: Summary of listener characteristics.....

Table 5: Example Speaker Similarity Matrix.....

Table 6: Pooled similarity matrix for AMR speaker samples from the free classification task.....

Table 7: Pooled similarity matrix for connected speech samples from the free classification task..

Table 8: Percent of listener ratings across variable, robotic and indeterminant categories for the *a priori* “instability” speakers.....

Table 9: Percent of listener ratings across variable, robotic and indeterminant categories for the *a priori* “inflexibility” speakers.....

Acknowledgments

I cannot begin to express my appreciation for my mentor, Dr. Kristie Spencer, who has been an endless source of patience, wisdom, and support throughout this process. Executing this study in a fully remote format was an especially unique experience and I attribute its success to her endless enthusiasm and flexibility. I have learned so much throughout this process because she shares her boundless knowledge in ways that are accessible, applicable, and compelling. I feel incredibly lucky to have worked with a mentor that believed in my strengths as a student, supported my areas of growth, and guided me through the research process in such a thoughtful manner.

I would also like to extend my appreciation to my committee members. Dr. Kaitlin Lansford, Dr. Tanya Eadie, and Dr. Edythe Strand for their expert insight. Dr. Kaitlin Lansford provided methodological guidance that was invaluable to the success of this study. I would also like to thank John Sidtis for his efforts in gathering and providing speaker samples.

I would like to thank Michael Singleton, Ph.D., Biostatistics Consultation Lead at University of Washington and the Institute of Translational Health Sciences, and Ali Shojaie, PhD, Professor of Biostatistics and Adjunct Professor of Statistics at the University of Washington for their statistical support.

I would like to thank my friends and colleagues Erica Seely and Claire Nesbitt for their camaraderie, advice, and humor while we tackled this thesis process together. I would like to thank my partner, Matthew, for patiently teaching me how to use Microsoft Excel to my advantage for data manipulation. I would also like to thank the graduate students of the University of Washington classes of '21 and '22 MedSLP and CoreSLP cohorts who volunteered as listeners despite their hectic schedules.

Finally, I would like to thank my friends and family for their constant encouragement throughout this journey. Without my support system I would not be where I am today. Thank you all for supporting my endeavors, academic and otherwise.

Introduction

Ataxic dysarthria is a motor speech disorder (MSD) associated with a disruption to the cerebellar control circuit. Etiologies of ataxic dysarthria are numerous, and include cerebellar infarction, tumor, traumatic brain injury, spinocerebellar ataxia, and neurotoxicity. The neurologic basis underlying ataxic dysarthria is decreased coordination of movement, thus, all speech domains can be affected, but most prominently articulation and prosody. Articulation deficits are typically characterized by vowel distortions, phoneme prolongation, and irregular articulatory breakdowns; speakers typically complain of “drunk” or intoxicated speech because distortions result in slurred speech production. Prosodic deficits include excess or insufficient loudness variation, excess and equal stress, and trouble coordinating breathing with speaking. Though there are commonly reported features of ataxic dysarthria, there is also considerable heterogeneity in the clustering and presentation of speech characteristics. This creates substantial challenges in research, motor speech diagnosis, and the development of appropriate treatments for individuals with ataxic dysarthria.

The heterogeneity observed in this population is exemplified by the inherent contradiction found in speech features. “Excess prosodic variation” and “insufficient prosodic variation,” two commonly reported features of ataxic dysarthria, highlight this incongruence. Excess prosodic variation is characterized as excessive and equalized stress patterns as well as pitch and loudness variability beyond that of typical speech. In contrast, insufficient prosodic variation encompasses the opposite with monopitch and monoloudness being the main characteristics. The presence of contradictory characteristics suggests the possibility of perceptual subgroups within ataxic dysarthria. Multiple theories, described below, seek to explain the variation observed in speakers with ataxic dysarthria. The ability to identify possible

subtypes of ataxic dysarthria would allow for more specificity and control in speaker selection for future research, thus, further informing clinical and theoretical frameworks of the cerebellum's role in motor speech control. Additionally, future research could develop hypotheses that are specific to certain subtypes, leading to tailored evaluation and intervention approaches. The purpose of the proposed study is to investigate perceptual subgroups of ataxic dysarthria. First, possible explanatory theories will be summarized.

Instability/ Inflexibility Theory

Hartelius, Runmarker, Andersen, and Nord (2000) introduced the concept of subgroups within ataxic dysarthria based on overall patterns related to the temporal control of speech. Fourteen participants with multiple sclerosis (MS) and primarily ataxic dysarthria, as well as fifteen control participants, were recorded while repeating and reading a series of words and sentences. Syllable duration and interstress interval durations were measured and analyzed for group differences and the coefficient of variation was used as a representation of temporal control variability. Results suggested a wide range of speech characteristics that the authors attributed to either the instability or the inflexibility of the temporal control of speech movements. That is, inflexibility in temporal control presented as reduced variability in syllable length, speaking rate, intonation, etc. and was perceived primarily as a prosodic insufficiency, also referred to as “scanning speech”. Conversely, instability in temporal control manifested as increased variability in interstress intervals and across repetitions of the same utterance as well as excess pitch and loudness variation, which was perceived as excess prosodic variation. Thus, this study established subgroups of instability and inflexibility, which highlighted marked differences in timing, rate and intonation across speakers.

Support for this theory is apparent in the literature even before its conception. Prior to the identification of instability and inflexibility by Hartelius et al. (2000), Boutsen, Bakker, and Duffy (1997) investigated ataxic dysarthria subgroups based on diadochokinetic (DDK) task performance. Twenty-seven speakers with ataxic dysarthria performed an oral reading task and maximum syllable repetition diadochokinetic tasks. Oral reading rate, syllable repetition rate, and variation in syllable repetition rate and intensity were measured across syllables. Q-factor analysis revealed three subgroups in their speakers. Group 1 presented with equalized syllable repetition rate and intensity across all DDK utterance types (/pΛ/, /tΛ/, and /kΛ/). This group corresponds with the inflexibility subgroup from Hartelius et al. (2000). Equalized rate and intensity were perceptually perceived as reduced prosodic variability or prosodic insufficiency. Boutsen and colleagues (1997) also found that Group 2 and Group 3 presented with excess syllable repetition rate variability for specific DDK utterance types. Specifically, Group 2 showed excess variability for the /pΛ/ syllable while Group 3 demonstrated excess variability for the /kΛ/ syllable. Both of these groups correspond to the Hartelius et al. (2000) instability classification with highly inconsistent syllable repetition rates, perceptually perceived as prosodic excess. Though the authors did not identify their groups based on the instability/inflexibility paradigm, their results align with this framework.

Differential Subsystem Theory

Darley, Aronson, and Brown (1969a, 1969b) were among the first to report speech characteristics that deviated from typical in speakers with dysarthria secondary to a variety of diagnoses, one of which was cerebellar ataxia. They determined that perceptual speech characteristics of ataxic dysarthria clustered into three groups based on predominance of attributes: articulatory inaccuracy, prosodic excess, and phonatory-prosodic insufficiency. They

concluded that individuals with ataxic dysarthria can be divided into three groups based on their primary deviant speech characteristics and that these deficits mirror the neuroanatomical underpinnings. Thus, the authors provided early indications of heterogeneity, through the lens of speech subsystems, that would serve to inform future investigations of perceptual characteristics.

Joanette and Dudley (1980) also sought to differentiate subgroups of ataxic dysarthria based on the primary speech deficits of speakers with Friedreich's ataxia. The authors hypothesized that, using 16 of the 38 speech dimensions defined by Darley and colleagues (1969a, 1969b), they could identify clusters based on primary deficits in articulation, phonation, and prosody. Results suggested that salient speech dimensions clustered into two factors. First was the general dysarthric factor, which included the articulatory characteristics of prolonged phonemes and imprecise consonants. The second was phonatory-stenosis, which included harshness, abnormal pitch level, and pitch breaks. Thus, two subgroups emerged, differentiated by a predominance of articulatory versus phonatory characteristics.

In a related study, Ebert, Hefter, Dohle, and Freund (1995) analyzed the breathing movements of adults with cerebellar disease at rest, spontaneously, and while coordinating simultaneous arm movements. Results indicated that breathing patterns became highly irregular when coordinating forearm movement simultaneously, suggesting that the cerebellum is responsible for the coordination of multiple motor subsystems. Their results also demonstrated that deficits in motor coordination for individuals with ataxia can be observed in breathing patterns. The authors suggested that breathing disturbances, namely coordination impairments and moments of breathing apnea, are caused by disruption to different cerebellar subsystems that can be affected independently. Thus, we can speculate that various speech characteristics of

ataxic dysarthria may be caused by the involvement of different cerebellar subsystems and, by extension, subtypes may emerge based on the network that is impacted.

Disease Site and Progression

When considering subgroups of ataxic dysarthria, it is important to determine whether the perceived perceptual heterogeneity is primarily driven by lesion location or disease severity.

Ackerman, Vogel, Petersen, and Poremba (1992) studied speech characteristics of participants with cerebellar infarcts to determine if lesion location and extent affected the presentation of ataxic dysarthria. Their findings suggested that the overall severity of dysarthria may be related to the size of the cerebellar lesion, while the primary speech deficit may be related to the location of the lesion. That is, bilateral cerebellar lesions were associated with primary articulatory deficits while lesions extending into the dentate nucleus presented as articulatory and phonatory, thereby suggesting some differentiation based on site of lesion.

However, converging evidence suggests that lesion location (Spencer and Slocomb, 2007; Vandana and Manjula, 2015) and disease severity (Ackermann and Hertrich, 1994; Schalling et al., 2008; Brendel et al., 2015) do not dictate the speech profile of individuals with ataxic dysarthria. For instance, Ackermann and Hertrich (1994) explored the effects of cerebellar disease severity on ataxic dysarthria. They hypothesized that speakers with slowly progressive cerebellar disease and speakers with an acute ataxic syndrome would present with different speech characteristics, thereby demonstrating subgroups based on disease progression. When comparing syllable durations, both groups presented with reduced *intrautterance* variability and increased *interutterance* variability with the latter pattern being more prominent. No significant differences were found between speakers with acute and chronic cerebellar disease, suggesting that disease progression alone does not account for the heterogeneity of speech patterns found in

ataxic dysarthria. Ackermann and colleagues (2007) explored this further, hypothesizing that two stages of cerebellar degeneration existed; first, temporal instability caused high variability in speech parameters, then progression led to “scanning speech” in which reduced speech parameter variability reflects inflexibility. Consonant imprecision and excess and equal stress were later found to be indicative of severity in speakers with spinocerebellar ataxia (Sidtis et al., 2011). However, multiple studies have demonstrated an inconsistent relationship between disease severity and speech profiles (Schalling et al., 2008; Brendel et al., 2015).

A literature review from Spencer and Slocomb (2007) echoed the same notion. The authors reviewed studies investigating the neural bases of ataxic dysarthria through neuroimaging reports of lesions. Bilateral and generalized cerebellar disease was most often associated with ataxic dysarthria; however, generalized disease does not allow for specific attributions regarding speech motor control in the cerebellum. For studies of speakers with focal lesions, there were a wide variety of lesion locations associated with ataxic dysarthria. While some lesion location patterns were revealed, none were prominent enough to suggest that ataxic dysarthria subgroups could be informed by lesion location alone. The authors explained that these results support the notion that different speech subsystems may be represented in distributed areas of the cerebellum and that there may be individual differences in motor speech subsystem representation across the cerebellum. However, the alignment of specific perceptual speech characteristics, or *patterns* of speech characteristics, with lesion location has yet to be established.

Feedforward/ Feedback contributions to motor control

There is evidence to suggest that patterns of impaired speech may be mediated by the functioning of feedback and feedforward motor control systems. Feedforward systems allow

speakers to predictively adapt their output based on learning from past productions. That is, if consistent auditory perturbation is provided over an extended period, speakers change their output to oppose the perturbation. Feedback systems enable speakers to adapt their output reactively, so unpredictable auditory perturbation would lead to online speech adaptations to oppose the auditory feedback disturbance. This notion has been examined in several clinical populations, including speakers with dysfluencies, apraxia of speech and ataxia. For example, speakers with apraxia of speech have been characterized as having difficulty implementing feedback control and/or an impaired feedforward mechanism for speech motor control (Ballard, Tourville, and Robin, 2014). However, not all speakers with apraxia of speech demonstrate patterns consistent with feedforward control deficits (Maas, Mailend and Guenther, 2015), suggesting that differentiated speech patterns may emerge based on which aspect of motor control, feedforward or feedback, is more impaired (Civier, Tasko, and Guenther, 2010).

The role of the cerebellum in speech motor control was assessed by Parrell, Agnew, Nagarajan, Houde, and Ivry (2017) in terms of anticipatory and reactive control, also known as feedforward and feedback control, respectively. Both control mechanisms were assessed in participants with cerebellar degeneration and matched controls to determine if the cerebellum is more heavily involved in feedforward control, feedback control, or both. The authors did not report how many, if any, participants with cerebellar degeneration also had a diagnosis of ataxic dysarthria. Using perturbations of auditory feedback during speech tasks, Parrell and colleagues (2017) demonstrated that adults with cerebellar degeneration relied more on feedback/reactive control of their speech motor system as evidenced by more reactive corrections and fewer anticipatory adjustments to speech in response to the perturbations. As such, the authors

suggested that the cerebellum plays a crucial role in feedforward control of speech movements and a less impactful role in feedback motor control.

In light of these findings, it is reasonable to anticipate varied speech patterns depending on disruption to feedforward versus feedback control mechanisms. Though the explicit link to ataxic dysarthria was not established by Parrell and colleagues (2017), a connection between feedforward/feedback control and subgroups of ataxic dysarthria can be surmised. That is, speakers primarily affected by impaired feedforward control mechanisms may lack the motor programming refinement required for well-coordinated speech subsystems, thus presenting with irregular articulatory breakdowns, excess pitch and loudness variation, variable rate, etc. These speakers could be considered a subgroup based on impairments in feedforward control.

Alternatively, another subgroup may be established based on impaired feedforward control mechanisms with subsequent hyperactivity of feedback control systems. An overreliance on feedback, or reactive, control may lead to reduced flexibility in on-line speech adaptation, resulting in rigid patterns of equalized stress, overall reduced variation, and reduced pitch and loudness variation.

Current Research

As demonstrated above, there are numerous theories that can inform the heterogeneity of speech characteristics in ataxic dysarthria. The instability/ inflexibility framework (Hartelius et al. 2000), in particular, emerged with theoretical and empirical support for understanding heterogeneity in ataxic dysarthria. Thus, this theory became the focus of numerous studies by Spencer and colleagues, which are summarized below.

Spencer and France (2016) investigated pre-recorded samples of ten speakers with ataxic dysarthria to examine patterns of speech characteristics. Based on the Hartelius et al. (2000)

framework, the investigators considered speech patterns through the lens of instability and inflexibility, operationally defining specific perceptual characteristics related to respiration, articulation, and phonation (see Table 1). Ten experienced SLPs listened to audio samples of sustained phonation and alternating/sequential motion rates (AMRs/SMRs) and rated perceptual speech features using a Visual Analog Scale. Results indicated that speakers clustered into three patterns: instability, inflexibility, and mixed. The most common pattern in this participant group was instability, followed by a mixed profile, then inflexibility. The authors echoed suggestions from others (Ackermann and Hertrich, 1994; Hartelius et al. 2000) that speakers more commonly align with the instability profile, which primarily reflects impaired coordination and timing. The inflexibility pattern was less prominent with only one speaker fitting this profile. Spencer and France (2016) thus provided evidence of heterogeneity within ataxic dysarthria and demonstrated that speakers could be perceptually grouped by experienced listeners into one of three speech patterns.

Table 1. Summary of salient speech characteristics aligning with instability and inflexibility patterns of ataxic dysarthria from Spencer and France, 2016.

	Instability	Inflexibility
Respiration	Incoordination of breathing with speaking Paradoxical breath patterning	Consistently reduced breath support
Articulation	Variable; possible undershoot and overshoot	Relatively consistent articulatory imprecision
Prosody	Inconsistent rate Uncontrolled loudness variation Variable pitch Variable stress patterns Variable rhythm (especially in AMRs/SMRs)	Decreased rate Monoloudness Monopitch Isochronous, metered stress patterns Steady rhythm (especially in AMRs/SMRs)

Note: AMRs, alternating motion rates; SMRs, sequential motion rates (Duffy 2013).

As a follow-up investigation, Spencer and Dawson (2019) contrasted the instability/inflexibility theory (Hartelius et al. 2000) with the differential subsystem theory (e.g.,

Joanette and Dudley, 1980). Diadochokinetic, sustained phonation, and monologue samples of eight adults with hereditary ataxia were perceptually rated by dysarthria experts using Visual Analog Scales. Results indicated that the participants' speech patterns were consistent with the instability/inflexibility profiles, but not the differential subsystems framework. Additionally, the authors reported that the speakers in the instability subgroup demonstrated the highest intelligibility scores, while speakers in the inflexibility subgroup demonstrated the lowest intelligibility scores. This pattern was not simply a function of disease duration and ataxia severity; instead, it suggested that the instability profile may be less detrimental to intelligibility than the inflexibility profile.

Most recently, Gore and Spencer (2020) expanded the investigations of instability and inflexibility patterns among speakers with ataxic dysarthria by replicating the AMR/SMR task and including a novel polysyllabic word repetition task. The AMR/SMR task was anticipated to best capture unusual variability of speech production patterns (i.e., corresponding to the instability theory) while the polysyllabic repetition task was expected to capture equalized stress patterns and monopitch (i.e., corresponding to the inflexibility theory). Speech samples were perceptually rated by ten listeners using an online visual analog scale system using a consistency/variability continuum that mapped onto the inflexibility and instability subgroups. Results revealed that nine of the eleven participants aligned with the instability pattern, congruous with previous studies (Ackermann and Hertrich, 1994; Hartelius et al., 2000; Spencer and France, 2016) in which instability was the most common profile. The authors also noted that Alternating Motion Rates was the most sensitive task for subgroup distinction. This study further supported the notion of instability/inflexibility subgroups in ataxic dysarthria.

Limitations of Previous Research

As with much of the research available regarding dysarthria, and specifically ataxic dysarthria, many investigations were limited by small sample sizes. Most studies employed samples of less than twenty speakers, which limits the examination of subgroups and the extent to which results can be used to infer broader explanations of speech heterogeneity. Additionally, some studies only included participants with a single medical diagnosis (e.g., Friedrich's Ataxia, Multiple Sclerosis, cerebellar infarction, etc.) which prevents findings from being generalized to the broader population of speakers with ataxic dysarthria.

To date, many of the studies of ataxic dysarthria have relied on visual analog scales or equal interval scales for judges to perceptually rate dimensions. While informative, these approaches are inherently limited because the dimensions rated by listeners are defined by the researchers, therefore constraining listener judgements. Constrained classification systems, in which researchers impose categories that constrict listener groupings, limit subgroup categorization to pre-described groups. This system prevents any additional groups from being revealed, thereby potentially limiting the representation of converging speech characteristics. Furthermore, the studies of Spencer and colleagues have used an untraditional version of the Visual Analog Scale by representing gradations of two phenomena, rather than one. This approach added complexity to the perceptual rating process and restricted the characteristics that could be rated.

Other studies have relied on acoustic measurements alone. Of the studies mentioned, Ackermann and Hertrich (1994), Parell et al. (2017), Boutsen et al. (1997), and Hartelius et al. (2000) reported acoustic measures without their perceptual correlates. Though acoustic measurements allow for objective differentiation between speakers with dysarthria, they are limited without the connection to perceptual groupings. Presently, the gold standard for

dysarthria diagnosis and subtype distinction is perceptual analysis by trained listeners, perhaps in combination with acoustic measures (Boutsen, Bakker, and Duffy, 1997).

Auditory Free-Classification

In light of the limitations of the extant literature, particularly the methodology used to examine subgroup differentiation, an alternative approach is warranted. Auditory free-classification tasks allow listeners to apply classifications based on the most salient perceptual characteristics, rather than researcher-imposed category labels. For example, listeners presented with spoken sentences by speakers with different accents could create perceptually similar groups by clustering icons on a visual grid. That is, the icons of speakers that sound like they may be from a similar region of the US would be placed in proximity on the grid. Clustering or scaling analyses would then reveal the most salient perceptual characteristics that are similar for each “group” generated by the listener.

Formative work by Clopper (2008) facilitated application of the free-classification paradigm, originally developed by Imai and colleagues (1966, Imai and Garner 1965), to the perceptual classification of speech stimuli. According to Clopper (2008), there are numerous advantages of auditory free-classification tasks in speech perception research. First, this method avoids bias towards experimenter’s pre-determined categories and researchers can allow for a fixed number or an infinite number of groupings based on their research question(s). It is also more efficient than methods that yield similar perceptual results. These methods include paired comparisons and oddball detection tasks, which are more time intensive and confusing to listeners than auditory free-classification tasks. Additionally, auditory free-classification tasks do not require specialized programming skills or software, thus, the method is more accessible and feasible for researchers.

Lansford, Liss, and Norton (2014) used an auditory free-classification task to determine if listeners could group speakers with dysarthria based on similarity of perceptual characteristics and whether the groups were congruent with established types of dysarthria (e.g., flaccid, spastic, ataxic, etc.). Twenty-three communication science graduate students listened to speech samples of 33 speakers with dysarthria and grouped speakers based on perceptual similarity. Grouping was conducted on a visual grid (see Figure 1), upon which listeners placed audio files in proximity of each other based on their similarity. Perceptually similar files were placed in proximity on the grid while divergent files were kept away from each other (see Figure 2).

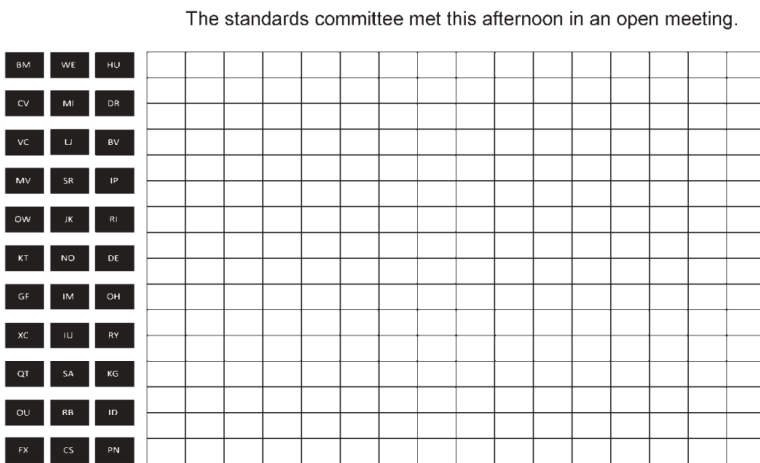


Figure 1. Screenshot of the grid system used for the free-classification task. Each black icon is paired with a specific speaker's recorded sentence production (Lansford, Liss, and Norton, 2014).

The standards committee met this afternoon in an open meeting.

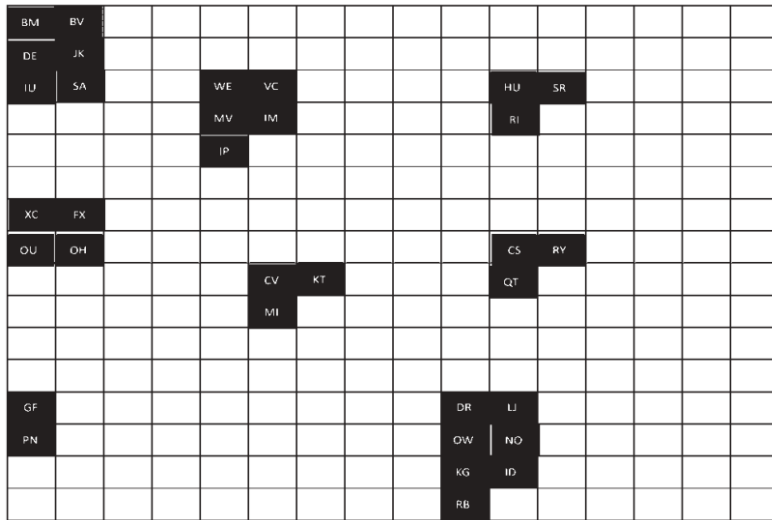


Figure 2. Example of the completed free-classification grid from one listener. Icons that touch one another were considered to be in the same group because they sounded similar to the listener (Lansford, Liss, and Norton, 2014).

After cluster analysis and multidimensional scaling, results revealed six clusters of perceptually similar speakers, which did not align with the typical seven dysarthria types. Most relevant to the proposed study, results also demonstrated that auditory free-classification tasks are feasible for studying perceptual comparisons in dysarthria. The authors explained that the auditory free-classification paradigm may reveal clusters more closely related to the true nature of a communication disorder than previously constrained classification systems.

In light of emerging evidence, this study examined perceptual subgroupings of ataxic dysarthria using an auditory free-classification paradigm. In addition, a second perceptual classification task was added, where listeners completed “guided classification.” For this task, grouping categories and definitions were provided, based on previous research (Spencer et al., 2016, 2019), to further examine listeners’ ability to detect perceptual differences among speakers with ataxic dysarthria. The following research question was addressed: Will speakers with ataxic

dysarthria be reliably classified into groupings that align with instability/inflexibility (or indeterminant/mixed) subgroups based on an unconstrained auditory free-classification listening task and a guided classification task? It was hypothesized that subgroups aligning with instability and inflexibility would be evident during the perceptual grouping tasks, and would be more apparent with the AMR task than the connected speech task (Spencer and France, 2016; Spencer and Dawson, 2019).

Methods

Due to the COVID-19 pandemic, no in-person data collection was feasible; thus, the study was conducted virtually and with existing speech files. All procedures were approved by the Institutional Review Board of the University of Washington.

Speakers

Sound files from 9 speakers with ataxic dysarthria from the Spencer lab and 20 speakers with ataxic dysarthria from the lab of John Sidtis, New York University, were considered for the study. Inclusion criteria for all speakers was a diagnosis of ataxia by a neurologist, age 18 or older, and native American English speakers. Exclusion criteria was presence of a mixed dysarthria. For Spencer lab files, the additional following exclusion criteria were met: a) disease duration less than one year, b) presence of another neurological disorder and/or history of a head trauma, c) severe or uncontrolled psychiatric disorder, d) presence of dementia determined by a Mini-Mental State Examination score of $< 26/30$, e) moderate-severe depression based on the Beck Depression Inventory-II score of > 19 , and f) alcohol or drug dependency.

From the 29 original speakers, five were excluded for mixed dysarthria ($n = 2$) or negligible dysarthria evident in the speech sample ($n = 3$) based on judgment of the researchers,

and two were excluded for poor recording quality. Of the remaining 22 speakers, 15 were chosen based on the PI's expert judgment of best fit with the instability/inflexibility framework (details below), and to ensure the listeners' perceptual task was of reasonable length/difficulty. Six speaker files were from the Spencer lab and nine speaker files were from the Sidtis lab.

Speaker Samples

All speaker samples included diadochokinetics and a short connected speech sample. From the Spencer lab, sound files used for the present experiment included AMRs and 15-word *Sentence Intelligibility Test* (Yorkston, Beukelman, & Hakel, 1996) sentences for the connected speech sample. From the Sidtis lab, recordings of AMRs were available for all speakers. However, connected speech sample sources varied and included repetitions of phonetically balanced sentences ($n = 4$), 1 sentence from the Cookie Theft picture description ($n = 4$), or 1 sentence from a spontaneous language sample ($n = 1$). The selection of Sidtis files for the connected speech sample were based on availability and the closest possible approximation to the 15-word sentences used for speakers from the Spencer lab, without violating grammatical or prosodic markers. See Table 2 for source, word count, syllable count, and duration of each speaker's extracted sample of connected speech.

Sound files were edited in Audacity for Windows. Sound quality correction, including balance and static/noise reduction, was completed as needed to correct for recording issues. Speaker files were cropped to ensure that AMRs and connected speech samples were of comparable length between Sidtis and Spencer lab speakers. AMR productions longer than 7 seconds were cropped at a syllable break to remain as close to 7 seconds as possible. One second of silence was inserted between each AMR syllable sequence (i.e., between the /pʌ/ sequence, /tʌ/ sequence, and /kʌ/ sequence).

Table 2. Summary of speaker features.

Speaker	AMR duration (sec)			Connected Speech			
	p Δ	t Δ	k Δ	Source	duration (sec)	# word*	# syllables*
S1	5.445	6.780	6.769	SIT	6.145	16 ^a	27
S2	4.231	5.823	5.651	SIT	5.241	15	21
S3	5.529	6.106	6.905	SIT	4.602	15	20
S4	4.088	3.737	3.533	SIT	6.093	15	19
S5	3.353	3.670	4.921	SIT	7.332	15	23
S6	7.098	7.129	7.197	Cookie Theft	3.321	17	24
S7	7.077	7.027	7.105	Sentence-question repetition	4.054	13	17
S8	7.261	7.080	7.103	Spontaneous speech	6.472	16	25
S9	6.910	6.544	5.650	Sentence-question repetition	6.225	14	18
S10	7.063	7.156	7.526	Sentence-question repetition	3.250	12	16
S11	6.996	6.423	8.347	SIT	4.616	15	18
S12	7.182	7.142	7.142	Cookie Theft	5.539	14	22
S13	7.038	7.136	7.257	Cookie Theft	7.986	16	23
S14	7.330	7.186	7.017	Sentence-question repetition	6.451	14	16
S15	7.073	7.108	7.084	Cookie Theft	6.725	10	16
Mean	6.245	6.403	6.614		5.603	14.47	20.33
SD	1.3095	1.1343	1.1563		1.3559	1.7158	3.4383
Min	3.353	3.67	3.533		3.25	10	16
Max	7.33	7.186	8.347		7.986	17	27

* Filler words including “uh” and “um” were excluded from the word and syllable count for connected speech samples.

^a Represents a 15-word *Sentence Intelligibility Test* sentence + one repeated word.

Perceptual Characterization of Speaker Samples

Dysarthria severity and speech characteristics were first determined through consensus agreement by two researchers who independently characterized speech based on diadochokinetic and extended connected speech samples of all speakers. *A priori* determination of alignment with inflexibility, instability, or mixed perceptual subgroups was completed for Spencer files prior to this study (Spencer & Dawson, 2019; Gore & Spencer, 2020). Determination of inflexibility/instability subgroup alignment for the Sidtis lab files, and replication of perceptual grouping of the Spencer lab files, was completed by the independent perceptual evaluation of two motor speech experts. The experts referred to previously established perceptual criteria and grouped speakers into subgroups based on the AMR task, previously shown to be most sensitive to subgroup discrimination (Spencer and France, 2016; Spencer & Dawson, 2019). Independent ratings revealed 93.3% agreement (14/15), with full consensus following discussion. Of the 15 participants, 10 were characterized as aligning with the instability group (Speakers S1 – S10), 5 were determined to align with the inflexibility group (Speakers S11 – S15), and 0 none were characterized as mixed instability-inflexibility by expert rating. Speaker characteristics are summarized in Table 3. Ten male and five female speakers comprised the sample, with a mean age of 46.7 years. Dysarthria severity ranged from very mild to moderate.

Table 3. Demographics and dysarthria characteristics of the speakers with ataxic dysarthria.

Speaker	Gender	Age (years)	Etiology	Dysarthria Severity	Speech Characteristics	<i>A Priori</i> Determination
S1	M	52	SCA3	Mild-Moderate	Irregular articulatory errors, telescoping, irregular rate, difficulty coordinating breathing and speaking	Instability
S2	F	60	SCA3	Mild	Irregular articulatory errors, telescoping	Instability
S3	M	19	SCA2	Moderate	Irregular articulatory errors, prolonged phonemes, telescoping, distorted vowels, excess loudness variation	Instability
S4	F	55	SCA3	Mild	Irregular articulatory errors	Instability
S5	M	35	SCA3	Mild-Moderate	Irregular articulatory errors, telescoping	Instability
S6	F	21	SCA1	Very Mild	Irregular articulatory errors, telescoping	Instability
S7	F	23	SCA1	Very Mild	Irregular articulatory errors	Instability
S8	M	54	SCA1	Mild	Irregular articulatory errors, telescoping	Instability
S9	F	66	SCA5	Mild-Moderate	Irregular articulatory errors, telescoping, excess pitch variation	Instability
S10	M	51	SCA6	Moderate	Irregular articulatory errors, telescoping, distorted vowels, difficulty coordinating breathing and speaking	Instability
S11	M	38	SCA1	Moderate	Irregular articulatory errors, distorted vowels, telescoping, mildly equalized stress, prolonged phonemes	Inflexibility
S12	M	29	SCA1	Mild-Moderate	Irregular articulatory errors, monopitch, monoloudness, equalized stress, slower rate	Inflexibility
S13	M	66	SCA5	Moderate	Irregular articulatory errors, equalized stress, slow rate	Inflexibility
S14	M	68	SCA5	Mild-Moderate	Irregular articulatory errors, equalized stress, slow rate	Inflexibility
S15	M	63	SCA6	Moderate	Irregular articulatory errors, equalized stress, slow rate, mild monopitch and monoloudness, distorted vowels	Inflexibility

Listeners

Listeners were 23 graduate students from the University of Washington Speech-Language Pathology program who completed a course in motor speech disorders and had foundational experience with dysarthria identification. No listeners reported additional experience with ataxic dysarthria beyond their graduate school training. All listeners had self-reported normal hearing. As shown in Table 4, the listener sample was comprised of 22 female and one male student with an average age of 25.87 years.

Table 4 Summary of listener characteristics.

Listener #	Age (years)	Sex	Education (years)	Native Language	Other Fluent Languages
L1	23	F	18	English	None
L2	22	F	16	English	None
L3	24	F	18	English	None
L4	23	F	18	English	None
L5	23	F	18	English	None
L6	33	F	18	English	None
L7	48	F	18	French	English
L8	24	F	18	English	None
L9	30	F	19	English	Spanish
L10	25	F	19	English	None
L11	24	F	21	Hindi	English
L12	23	F	17	English	Korean
L13	24	F	18	English	None
L14	23	F	17	English	Japanese
L15	22	F	18	English	None
L16	27	F	18	English	None
L17	27	F	19	English	None
L18	24	M	19	English	None
L19	25	F	18	English	None
L20	25	F	17	English	None
L21	27	F	18	English	None
L22	25	F	17	English	None
L23	24	F	18	English	None
Mean	25.87	1 M, 22 F	18.04		
SD	5.46		0.95		

Perceptual Classification Tasks

All listeners reviewed the IRB approved consent form, sent a statement of consent via email, and completed an intake questionnaire prior to starting the classification tasks.

Compensation for participation was \$20.

Following the procedure of Lansford et al. (2014), a single Microsoft PowerPoint slide with a white background was used for each auditory free-classification and guided-classification task: one for connected speech and another for AMRs. Speaker recordings were embedded onto black rectangles with a random two-letter identifier in white font to create an icon for each speaker sample. These icons and audio files were merged into the PowerPoint slide, allowing the listener to double-click each rectangle to hear the audio file. Icons were randomly placed in 3 columns to the left of an 8x8 cell grid (see Figure 3) and each icon was sized to fit exactly into each cell. The PowerPoint “notes” section was available for optional listener notations.

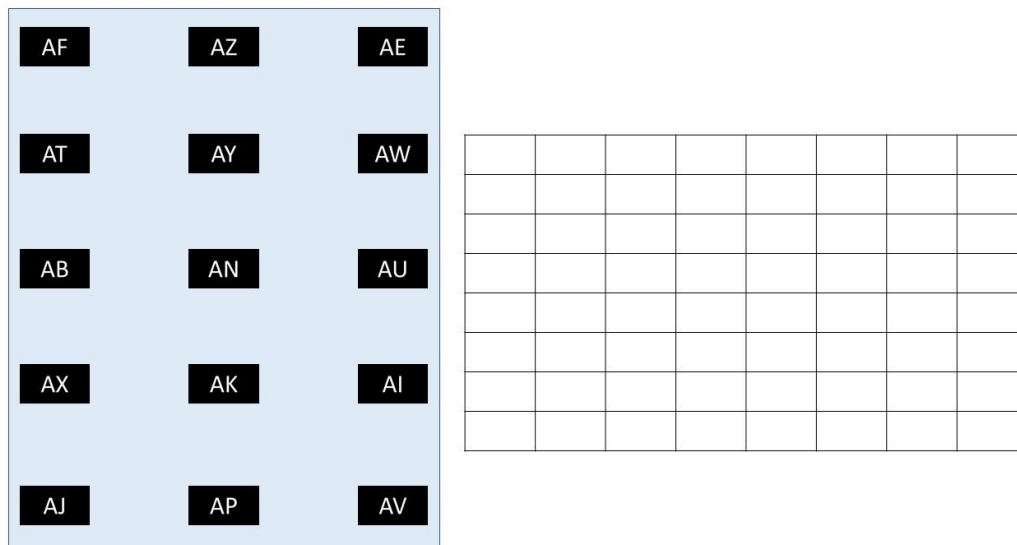


Figure 3. Screenshot of Free Classification task layout with speaker files and 8x8 grid on a blank Microsoft PowerPoint Slide.

Listeners used their personal computers to complete the task. They were instructed to find a quiet workspace and were encouraged to use high-quality over-ear headphones when possible. Listeners were informed that all speakers were diagnosed with ataxic dysarthria, but they were given no other information about the speakers. Listeners were permitted, but not required, to use the notes section at the bottom of the PowerPoint slide to take notes about perceptual similarity throughout the tasks, to reduce working memory load. Notes were saved for descriptive purposes.

Listeners completed two tasks: a free classification task and a guided classification task. The free classification task was always presented first, to avoid any influence of the subgroup characterizations provided in the guided classification task. To accomplish this, downloadable PowerPoint files and instructions were provided to listeners in a Google Form. The Google Form's settings required listeners to download, complete, and upload the free classification task before advancing to the page containing the downloadable file for the guided classification task. Once listeners reached the guided classification page of the Google Form, they could not go back to the free classification page. For both the free classification and guided classification task, speaking tasks (AMRs and connected speech) were counterbalanced.

In the free classification task, participants were instructed to listen to the audio files and place them on the blank grid by clicking, dragging, and dropping them into groups with speakers that sound similar— placed so the icons were touching each other on the grid. They were not given further instructions about how to make their similarity judgments and were not informed of the purpose of the study until their debrief. Listeners were permitted to make any number of groups with any number of speakers in each group. They were also permitted to listen to each file and rearrange their groupings as many times as necessary.

For the guided classification task, listeners were presented with characterizations of two subtypes, listed as “variable” and “robotic”, corresponding to instability and inflexibility subtypes, respectively (see Figure 4). The “variable” group was described in the following manner: *These speakers have more unstable motor control. They are typically more variable with sound errors, stress, rhythm, and pitch. They tend to be faster or have inconsistent rate. They may have incoordination of breathing and speaking.* The “robotic” group was described as: *more inflexible motor control; more equalized stress and rhythm, and more consistent errors; tend to be slower; may have monopitch/monoloudness.* Listeners were also provided with a third category, indeterminant, for speakers that the listener believed did not fit either subtype profile. Participants were instructed to listen to the audio files and place them on the grid by clicking, dragging, and dropping them into the category they felt matched the speaker’s dysarthria characteristics. They were permitted to put any number of speakers in each category and were permitted to listen to and rearrange speakers as many times as necessary. See Appendix A for all PowerPoint slides with instructions and the free classification and guided classification tasks.

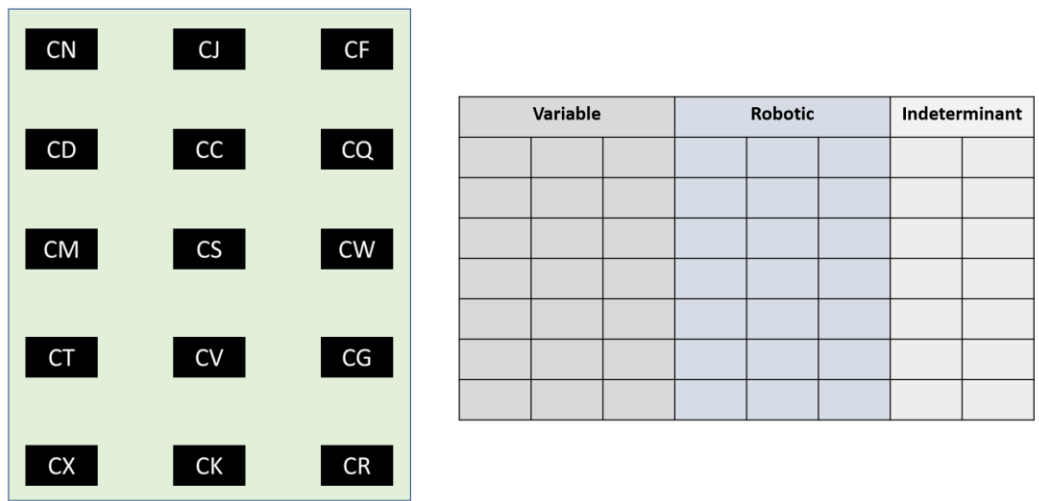


Figure 4. Screenshot of the Guided Classification task layout with speaker files and divided grid on a blank Microsoft PowerPoint Slide.

Data Preparation and Analysis

Following the procedures of Lansford et al. (2014), the PowerPoint slides with each listener’s speaker groupings were coded alphanumerically and transferred to a Microsoft Excel spreadsheet. The similarity data obtained from each listener were arranged into a 15 x 15 speaker-similarity matrix in Excel (see Table 5 for an example). A “1” was entered into cells corresponding to two speakers who were grouped together by a listener, while a “0” was entered if the speakers were not grouped together. Individual listener matrices were then combined into a pooled speaker similarity matrix for each of the free classification tasks (AMR and connected speech). The pooled matrix was then used to conduct the cluster analysis (*hclust()* function) and multidimensional scaling (*cmdscale()* function) which were conducted using R Statistical Software (version 4.0.2). Descriptive and Chi-Square analyses were used for the guided classification task and conducted using SPSS (v.26).

Table 5. Example Speaker Similarity Matrix. Each cell represents the frequency at which each speaker pair was judged to be perceptually similar by the listener.

L1	S1	S2	S3	S4	S5	S6	S7	S8	S8	S10	S11	S12	S13	S14	S15
S1	0	1	1	1	1	0	0	1	1	0	0	0	0	0	0
S2	1	0	1	1	1	0	0	1	1	0	0	0	0	0	0
S3	1	1	0	1	1	0	0	1	1	0	0	0	0	0	0
S4	1	1	1	0	1	0	0	1	1	0	0	0	0	0	0
S5	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0
S6	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0
S7	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0
S8	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0
S9	1	1	1	1	1	0	0	1	0	0	0	0	0	0	0
S10	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
S11	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
S12	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0
S13	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1
S14	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1
S15	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0

Results

Reliability

Reliability of data entry was completed by two independent researchers on 17.4% (4/23) of listener similarity matrices with 99.8% agreement between coders.

Interrater reliability among the 23 listeners was calculated in SPSS (v.26) for the guided classification task to determine the agreement between listeners when provided with categorical labels and definitions. A two-way mixed effects intraclass correlation coefficient (ICC) using absolute agreement revealed excellent reliability with an intraclass correlation coefficient of 0.988 (95% CI 0.978 - 0.995; $p < .001$). Interrater reliability was not conducted for the free classification task as perceptual groupings were inherently idiosyncratic.

Free Classification Task

Data from the individual similarity matrices (see Table 5) were summed to create a pooled similarity matrix for both the AMR task and the Connected speech task (see Tables 6, 7).

Table 6. Pooled similarity matrix for AMR speaker samples from the free classification task.

AMR task	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	0	12	11	8	12	10	10	13	11	9	2	2	1	2	1
S2	12	0	14	14	12	10	11	10	9	5	0	2	0	0	0
S3	11	14	0	11	13	8	9	8	7	6	0	2	0	0	0
S4	8	14	11	0	10	14	12	12	8	7	0	6	0	1	1
S5	12	12	13	10	0	7	11	10	9	8	0	4	0	0	0
S6	10	10	8	14	7	0	13	14	6	11	2	3	1	1	1
S7	10	11	9	12	11	13	0	10	9	11	1	6	0	1	0
S8	13	10	8	12	10	14	10	0	9	9	2	2	1	1	1
S9	11	9	7	8	9	6	9	9	0	10	2	3	1	3	2
S10	9	5	6	7	8	11	11	9	10	0	3	5	2	3	1
S11	2	0	0	0	0	2	1	2	2	3	0	10	17	11	13
S12	2	2	2	6	4	3	6	2	3	5	10	0	9	6	6
S13	1	0	0	0	0	1	0	1	1	2	17	9	0	14	14
S14	2	0	0	1	0	1	1	1	3	3	11	6	14	0	17
S15	1	0	0	1	0	1	0	1	2	1	13	6	14	17	0

Table 7. Pooled similarity matrix for connected speech samples from the free classification task.

Connected Speech task	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15
S1	0	6	4	8	6	9	7	10	2	10	6	2	1	1	0
S2	6	0	3	18	4	6	7	8	1	4	8	4	1	0	0
S3	4	3	0	1	1	2	2	6	13	2	3	3	6	5	3
S4	8	18	1	0	4	5	5	6	2	6	7	5	0	0	0
S5	6	4	1	4	0	1	2	6	2	6	6	5	5	5	5
S6	9	6	2	5	1	0	20	5	1	7	1	1	0	0	0
S7	7	7	2	5	2	20	0	4	2	6	3	0	0	0	0
S8	10	8	6	6	6	5	4	0	3	9	8	5	1	0	2
S9	2	1	13	2	2	1	2	3	0	2	1	2	8	8	7
S10	10	4	2	6	6	7	6	9	2	0	8	6	0	0	1
S11	6	8	3	7	6	1	3	8	1	8	0	10	1	1	3
S12	2	4	3	5	5	1	0	5	2	6	10	0	6	3	4
S13	1	1	6	0	5	0	0	1	8	0	1	6	0	17	11
S14	1	0	5	0	5	0	0	0	8	0	1	3	17	0	12
S15	0	0	3	0	5	0	0	2	7	1	3	4	11	12	0

From the pooled similarity matrix, hierarchical clustering analyses were completed using an agglomerative (bottom-up) approach with average linkage. The results of the cluster analysis are best visualized by dendrograms (see Figures 5, 7) which illustrate how speakers with ataxic dysarthria were linked together. Agglomerative clustering begins by assuming each speaker is in a cluster alone (as indicated by the individual lines for each speaker at the bottom of Figures 5, 7), then iteratively merges speakers together based on free classification similarity until there is a single cluster left (as indicated by the single line at the top of Figures 5, 7). The number of clusters resulting from the analysis were defined by the researchers based on large gaps in the dendrogram (James, Witten, Hastie, & Tibshirani, 2017) together with findings from multidimensional scaling (see below) rather than by a pre-determined set number of clusters.

For the AMR task (Figure 5), two clusters emerged. Cluster 1 contained all ten speakers (S1-S10) determined by expert judgment, pre-experiment, to reflect the instability subgroup, whereas cluster 2 contained all five speakers (S11-S15) determined by expert judgment to reflect inflexibility subgroup.

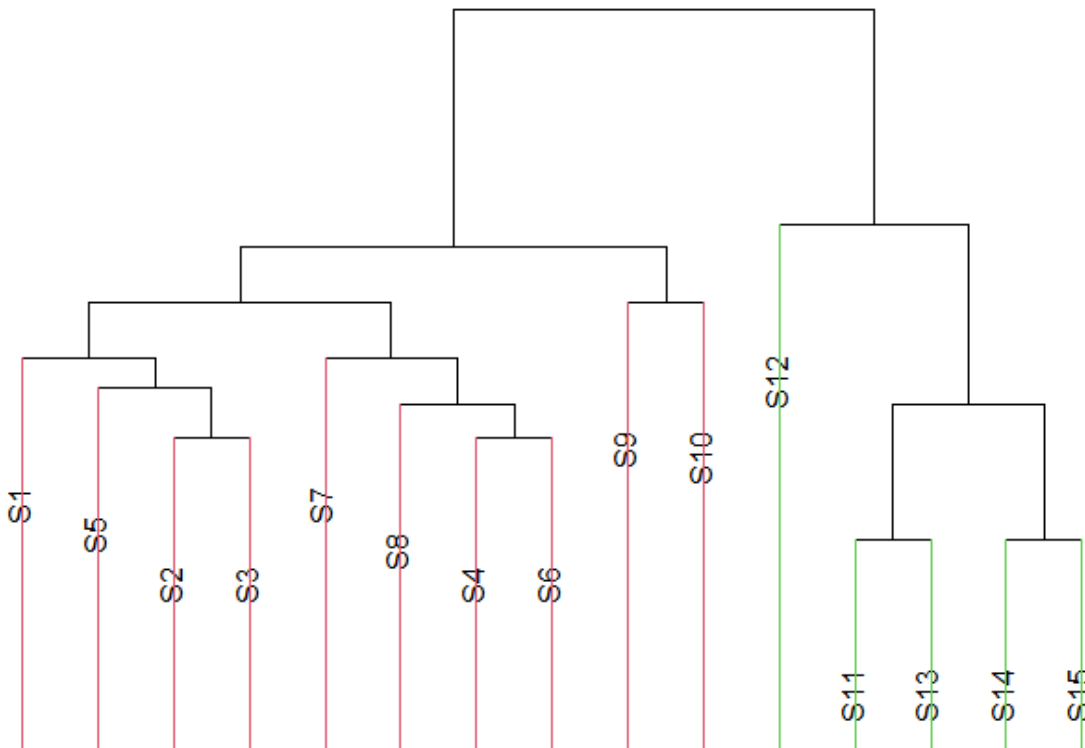


Figure 5. Dendrogram with the two-cluster solution for the AMR task that shows the progressive linking of speakers based on listener groupings during auditory free classification. The red and green lines are used to demarcate the two clusters.

Multidimensional scaling (MDS) is useful for visualizing the similarity (or dissimilarity) of the data; simply put, speakers closer together were rated more similarly by listeners than speakers farther apart. The two clusters identified based on hierarchical clustering are corroborated with findings from the MDS, and are shown in the lower-dimensional projection by blue circles in Figure 6.

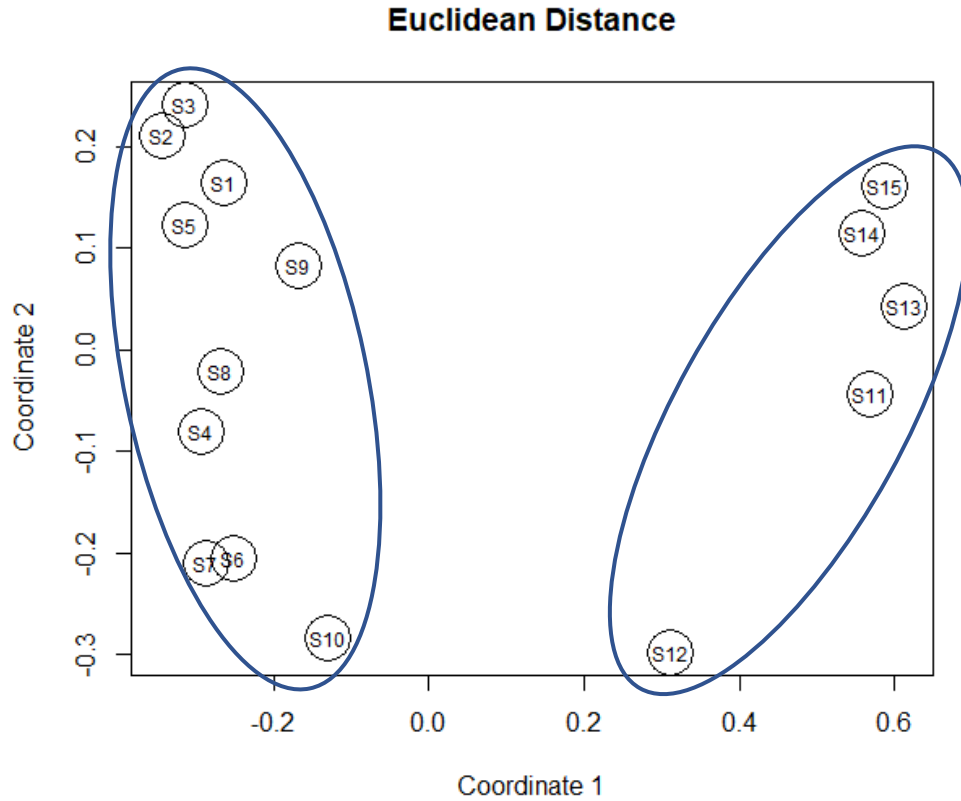


Figure 6. Multidimensional scaling plot reflecting the two-cluster solution for the AMR task.

For the Connected speech task, three clusters emerged (Figure 7). Cluster 1 contained 7 speakers, all determined by expert judgment, pre-experiment, to reflect the instability subgroup. Cluster 2 contained a mix of speakers determined by expert judgment to reflect the instability subgroup (n=2) and the inflexibility subgroup (n=3). Cluster 3 contained 3 speakers, two with a predetermined inflexible pattern and one with an unstable pattern. The three clusters identified based on hierarchical clustering are corroborated with findings from the MDS and are shown in the lower-dimensional projection by blue circles in Figure 8.

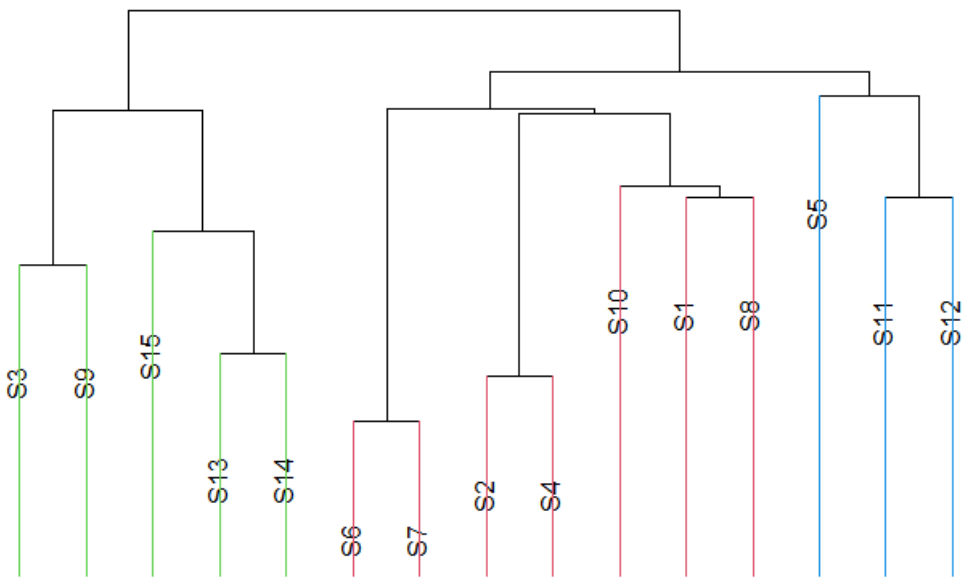


Figure 7. Dendrogram with a three-cluster solution for the Connected speech task that shows the progressive linking of speakers based on listener groupings during auditory free classification. The green, red, and blue lines are used to demarcate the three cluster.

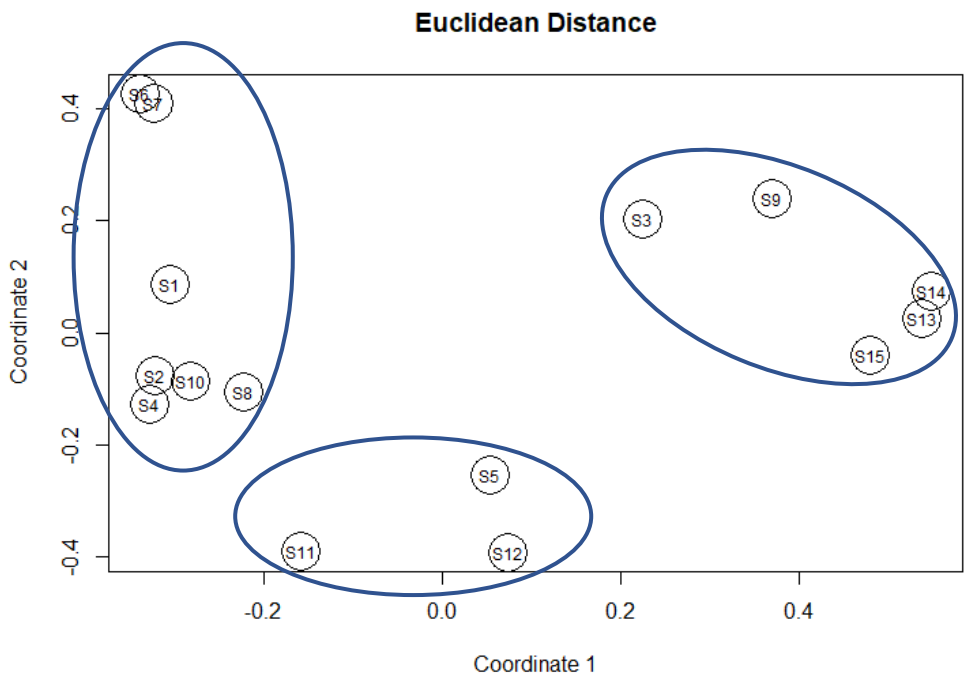


Figure 8. Multidimensional scaling plot reflecting the three-cluster solution for the Connected Speech task.

Across the three clusters, there did not appear to be an influence of the source/type of the connected speech samples across the Spencer and Sidtis labs. As reflected in Table 8, there are no clear patterns with respect to type of SCA. Cluster 3 was comprised of relatively younger, male speakers, though group membership is small ($n=3$). Average age differed across all clusters (Cluster 1 = 45.14, Cluster 2 = 56.4, Cluster 3 = 34.0). There is not a strong pattern related to severity as all three clusters had a mix of mild and moderate speakers; however, the overall severity slightly increased with each cluster. This weak trend is also reflected in average intelligibility scores for Cluster 1 (96.51), Cluster 2 (94.16), and Cluster 3 (95.23). The speech characteristics of Cluster 1 reflected irregularity in terms of articulatory errors, incoordination of breathing and speaking, etc. Three of the five members of Cluster 2 manifested equalized stress and slow rate. Finally, two of the three speakers in Cluster 3 had equalized stress patterns. As a reminder, the speech characteristics reported in Table 8 were derived from complete speech samples across multiple tasks, whereas listeners were making connected speech judgments based on ~ 15 words. Optional listener notes made during the experiment are included in Appendix B.

Table 8. Speaker characteristics grouped by listener-derived clusters for the connected speech task.

	Speaker	Gender	Age (years)	Etiology	Dysarthria Severity	% Intelligibility	Speech Characteristics	<i>A Priori</i> Determination
Cluster 1	S1	M	52	SCA3	Mild-Moderate	96.0	Irregular articulatory errors, telescoping, irregular rate, difficulty coordinating breathing and speaking	Instability
	S2	F	60	SCA3	Mild	96.0	Irregular articulatory errors, telescoping	Instability
	S4	F	55	SCA3	Mild	95.7	Irregular articulatory errors	Instability
	S6	F	21	SCA1	Very Mild	100	Irregular articulatory errors, telescoping	Instability
	S7	F	23	SCA1	Very Mild	100	Irregular articulatory errors	Instability
	S8	M	54	SCA1	Mild	93.3	Irregular articulatory errors, telescoping	Instability
	S10	M	51	SCA6	Moderate	94.6	Irregular articulatory errors, telescoping, distorted vowels, difficulty coordinating breathing and speaking	Instability
Cluster 2	S3	M	19	SCA2	Moderate	91.0	Irregular articulatory errors, prolonged phonemes, telescoping, distorted vowels, excess loudness variation	Instability
	S9	F	66	SCA5	Mild-Moderate	92.9	Irregular articulatory errors, telescoping, excess pitch variation	Instability
	S13	M	66	SCA5	Moderate	97.3	Irregular articulatory errors, equalized stress, slow rate	Inflexibility
	S14	M	68	SCA5	Mild-Moderate	97.1	Irregular articulatory errors, equalized stress, slow rate	Inflexibility
	S15	M	63	SCA6	Moderate	92.5	Irregular articulatory errors, equalized stress, slow rate, mild monopitch and monoloudness, distorted vowels	Inflexibility
Cluster 3	S5	M	35	SCA3	Mild-Moderate	94.3	Irregular articulatory errors, telescoping	Instability
	S11	M	38	SCA1	Moderate	95.0	Irregular articulatory errors, distorted vowels, telescoping, mildly equalized stress, prolonged phonemes	Inflexibility
	S12	M	29	SCA1	Mild-Moderate	96.4	Irregular articulatory errors, monopitch, monoloudness, equalized stress, slower rate	Inflexibility

Note. For Speakers S1 -S5 and S11 (Spencer lab), intelligibility scores were derived from the Sentence Intelligibility Test scores. For Speakers S6-S10 and S12-S15 (Sidtis lab), intelligibility was determined from connected speech recordings which were, on average, 51.3 words (SD 18.4), and ranged from 28 to 74 word samples.

Guided Classification Results

The agreement between listeners' categorizations from the guided classification task and researchers' *a priori* classifications are reflected in Tables 9 and 10. Speakers judged pre-experimentally to reflect the instability subgroup were most frequently matched with the parallel listener rating of "variable". Similarly, speakers judged pre-experimentally to reflect the inflexibility subgroup were most frequently matched with the parallel listener rating of "robotic". These patterns are also depicted in Figure 9.

Table 9. Percent of listener ratings across variable, robotic and indeterminant categories for the *a priori* "instability" speakers.

	AMRs			Connected Speech		
	Variable	Robotic	Indeterminant	Variable	Robotic	Indeterminant
Total/230*	223	6	1	153	48	29
% of ratings	97.0	2.6	0.4	66.5	20.9	12.6

*Reflects 10 speakers x 23 listeners

Chi-Square analyses were conducted to test the null hypothesis that the variable, robotic, and indeterminant categories are independent and would lead to expected frequency counts that are relatively equal across categories. All comparisons were significant, suggesting a strong association among the variables and departure from the expected frequencies. Specifically, for the *a priori* "instability" speakers, significant results were found for the AMR task: $\chi^2 (2, N = 230) = 419.12, p < 0.001$, and the connected speech task: $\chi^2 (2, N = 230) = 116.36, p < 0.001$.

Table 10. Percent of listener ratings across variable, robotic and indeterminant categories for the a priori “inflexibility” speakers.

	AMRs			Connected Speech		
	Variable	Robotic	Indeterminant	Variable	Robotic	Indeterminant
Total/115*	6	105	4	33	80	2
% of ratings	5.2	91.3	3.5	28.7	69.6	1.7

*Reflects 5 speakers x 23 listeners

Similarly, for the a priori “inflexibility” speakers, significant results were found for the AMR task: $\chi^2(2, N = 115) = 173.96, p < 0.001$, and the connected speech task: $\chi^2(2, N = 115) = 80.47, p < 0.001$.

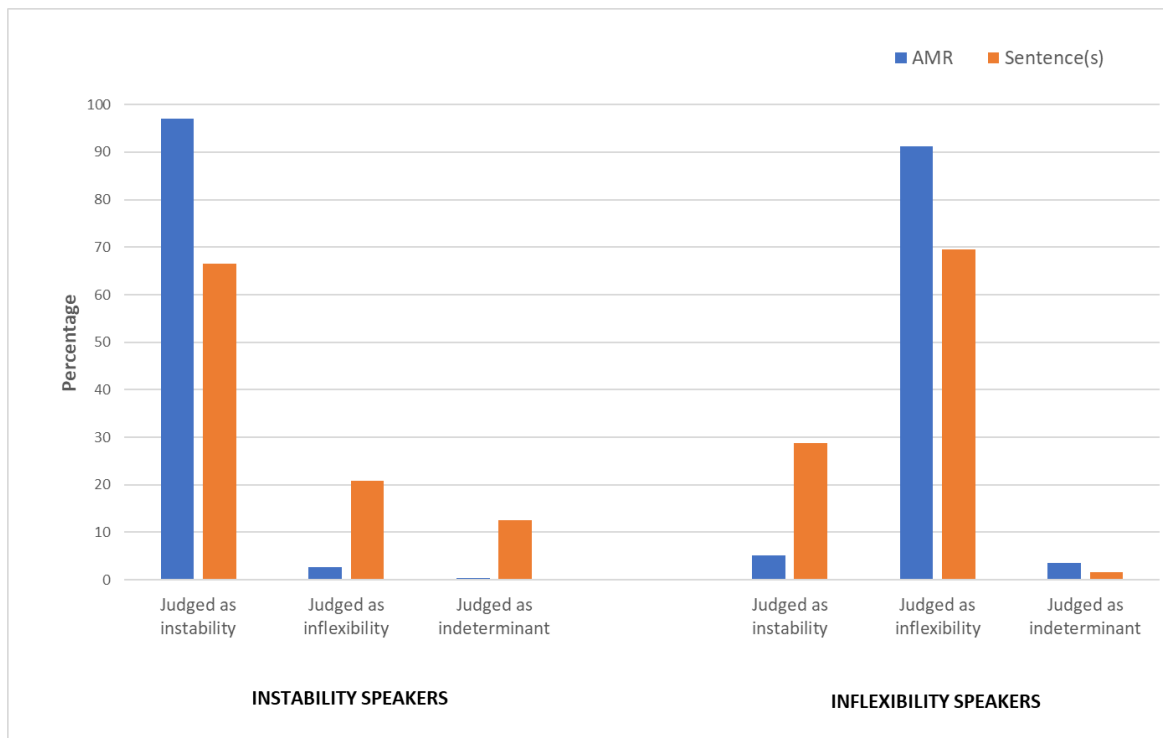


Figure 9. Percentage match between a priori expert classification and listener ratings during the guided classification task.

Discussion

The goal of this study was to examine the speech patterns of speakers with ataxic dysarthria in the context of instability and inflexibility subtypes using a free classification approach. Results add to the growing empirical evidence, including Spencer and France (2016), Spencer and Dawson (2019), and Gore and Spencer (2020), that found speech profiles of speakers with ataxic dysarthria to align with the instability/inflexibility theory of Hartelius et al. (2000). Overall, the present study builds on the results of previous studies with the addition of a free classification approach, which allowed for unconstrained and uninfluenced classification by listeners.

Confirmed Use of AMRs for Subgroup Differentiation

Based on findings from both the free classification and guided classification tasks, the instability/inflexibility subgroups appear to be detectable by relatively inexperienced listeners, particularly during the AMR maximum performance task. This pattern was also found in previous research (Spencer and France, 2016; Spencer and Dawson, 2019; and Gore and Spencer, 2020) and suggests that AMR tasks are more sensitive to irregularities in cerebellar motor control. AMRs require accurate repetition of motor sequences compared to connected speech, which has inherent variability of motor patterns; thus, abnormalities of timing and steadiness become more apparent (Duffy, 2020). At a broader level, limb diadochokinetics are also considered sensitive and informative indicators of cerebellar dysfunction. In a 2016 consensus paper, Bodranghien and colleagues discussed how dysdiadochokinesia results from a dissociation in the timing and force of muscle contractions which should act synergistically—a phenomenon known as asynergia. Specifically, asynergia reflects poor timing of the activation of agonist and antagonist muscles, best captured in multijoint and fast movement sequences.

Asynergia is a hallmark of cerebellar disease and has been linked to difficulties with feedforward processing, that is, an inability to predict the consequences of movements (Pisotta & Molinari, 2014). Fast movement sequences, such as speech and limb diadochokinetics, where there is the additional diminished ability to use afferent feedback, are particularly sensitive to asynergia (Bodranghien et al., 2016).

During the auditory free classification task, which asks listeners to group similar sounding speakers without any guidance, the subgroup differentiation for AMRs was definitive. That is, the cluster analysis revealed two clusters that perfectly matched the pre-experiment classifications by experts. One cluster contained all ten speakers that were determined by experts to reflect the instability motor pattern (S1-S10) and the other contained all five speakers determined pre-experiment to reflect the inflexibility motor pattern (S11 – S15). Moreover, during the guided classification task, where listeners were provided with three category labels (“variable,” “robotic”, and “indeterminant”) and brief descriptions to facilitate their perceptual decisions, listeners were highly successful at matching the *a priori* subgroup determinations for AMRs. That is, for the *a priori* “instability” speakers, listeners chose the parallel “variable” category 97% of time. Similarly, for the *a priori* “inflexibility” speakers, listeners chose the parallel “robotic” category 91.3% of the time. Despite their limited experience with dysarthria, the listeners were highly consistent overall, reflected by the excellent interrater reliability (ICC of .988).

Converging evidence thus suggests that the AMR task is sensitive to motor control differences and may be particularly useful for subgroup differentiation for speakers with ataxic dysarthria. Diadochokinetics are quick and easy to administer clinically and are relatively impervious to language or cognitive impairment. While the relationship between

diadochokinetics and intelligibility or dysarthria severity is mixed (Wang, Kent, Duffy, & Thomas, 2009), when implemented properly, diadochokinetics can provide a wealth of information about the rate, regularity, and precision of motor control (Duffy, 2020), and may ultimately be used in a clinical setting for identifying ataxic dysarthria subgroups.

Possible Use of Connected Speech for Subgroup Differentiation

As anticipated, listeners were more successful at matching the *a priori* subgroup judgments during the AMR task than the connected speech task. However, the listeners judgements still aligned quite well with the *a priori* groups, despite their inexperience with dysarthria and the limited sample with which they were making perceptual judgments. That is, connected speech judgments were limited to 1-2 sentence or approximately 15 words, similar to previous perceptual studies (Lansford et al., 2014). This task would be challenging for the most seasoned listener. However, during the guided listening task, listeners were able to match the *a priori* “Instability” subgroups 67% of the time and the “Inflexibility” subgroups 70% of the time. Moreover, during the free classification task, the listeners were still able to identify a large cluster of seven out of fifteen speakers, all of whom matched the “Instability” motor pattern according to expert listeners. The other two smaller clusters were a mix of “Instability” and “Inflexibility” speakers. Possible, but weak, trends arose with respect to age, gender, and severity.

The results from the connected speech task were possibly influenced by the inconsistent type of speech samples available, given the necessity to combine samples across labs. The samples thus represented a mix of sentence reading, sentence-question combination reading, a Cookie Theft picture description, and spontaneous speech. While samples were equated in length and

type as much as possible, there may have been an influence of speech elicitation method on perceptual judgments.

Theoretical Justification of Ataxic Dysarthria Subgroups

Our understanding of the basis of inflexible and unstable cerebellar motor patterns, as originally postulated by Hartelius and colleagues (2000), may be furthered by consideration of feedforward and feedback motor control processes. As discussed, coordinated movement relies on the combination of feedback control and anticipatory, or feedforward, systems (Pisotta and Molinari, 2014). Feedforward systems allow speakers to predictively adapt their output based on learning from past productions. For example, if consistent perturbation is provided over an extended period, speakers change their output to oppose the perturbation. Feedback systems, on the other hand, enable speakers to adapt their output reactively; unpredictable perturbations would therefore lead to online, compensatory speech adaptations to oppose the perturbation.

With respect to speech, feedforward control primes the cerebellum for the syllabic content of the intended utterance so the cerebellum can refine the temporal and prosodic characteristics of the motor plan (Duffy, 2020). Feedback control is used in speech production (both auditory and somatosensory) but does not seem to be critical for maintaining relatively accurate speech function (Parrell and Houde, 2019). Because speech is rapid, it is not as easily informed by the feedback control loop, using the error of one movement to inform the next (Tourville & Guenther, 2011). Instead, speakers rely more heavily on the feedforward control loop (Parrell et al, 2017) and the transformation of a desired articulatory movement to the motor command needed to complete the action.

Unstable, variable motor control is associated with a compromised cerebellar feedforward control loop (Parrell et al., 2017) and would align with the instability subgroup of ataxic dysarthria. Conversely, an overreliance on feedback loops, which can emerge if the feedforward system is disrupted (Parrell et al., 2017), may sacrifice speed and prosodic variation, and lead to slower and more “scanning” speech—a pattern more characteristic of the inflexibility subgroup. Recent work from Parrell's lab (2021, unpublished), which examined speaking and reaching movements in individuals with cerebellar degeneration, failed to replicate the overreliance on feedback loops, as demonstrated in their 2017 study. Feedforward motor control deficits were convincingly demonstrated in both studies. However, the researchers reported considerable individual variability in performance suggesting that some individuals with cerebellar disease may still exhibit abnormal feedback control systems.

The Auditory Free Classification Approach

Unlike prior studies that relied on visual analog scales and/or equal interval scales to measure speech characteristics and subsequent alignment with the instability/inflexibility subtypes, the present study implemented a free-classification approach. Listeners were asked to group speakers based on their perception of the most salient speech patterns rather than investigator-determined metrics of speech characteristics. The consistent clustering of speakers, particularly for the AMR task, suggests successful use of this perceptual paradigm for subgroup differentiation. This outcome is even more compelling given the level of experience of the listeners. That is, listeners in the present study were first- and second-year graduate student clinicians with limited experience with ataxic dysarthria, or dysarthria in general, and minimal clinical exposure. Thus, this study serves as a proof of concept for the free-classification approach for use in dysarthria subtype research, and builds on the work from Lansford, Liss, and

Norton (2014), who used the same approach for classification of general dysarthria types. Their results showed a relationship between free-classification outcomes and acoustic and perceptual metrics, supporting the ability of this task to corroborate gold standard measures of speech characteristics. Currently, research is also underway using the auditory free classification approach to examine perceptual subgroups within Huntington's disease (Kim, Diehl, de Riesthal, Mefferd, 2020).

Continued Investigation of Subgroups within Ataxic Dysarthria is Warranted

The heterogeneity observed in speakers with ataxic dysarthria creates obstacles in research and in clinical practice. Thus, the identification of subgroups of ataxic dysarthria would have several important outcomes. First, it would facilitate differential speech diagnosis. Second, it would lead to more refined selection of participants in research on ataxic dysarthria as opposed to inclusion of speakers with disparate speech patterns. Third, treatment for ataxic dysarthria could be better tailored. For example, the treatment approach for someone fitting the inflexibility profile may preferentially focus on activities that bring variability to speech, such as contrastive stress or flexible breath groups. Conversely, the treatment approach for someone fitting the instability profile may target activities to bring stability to speech, such as attenuated use of pitch/loudness for emphasis or regulated breath groups. Finally, identification of subgroups could inform theoretical and clinical models of cerebellar functioning and disease, which remain limited.

Limitations and Future Directions

This study has several limitations. Namely, generalization of these results are limited due to a relatively small speaker sample size that only included speakers with hereditary ataxia. It is

uncertain if and how the results would translate to speakers with acquired forms of ataxia, though earlier work (Spencer & France, 2016) suggested parallels between the acquired and hereditary ataxia speakers. Additionally, restrictions set by the COVID-19 pandemic did not allow for in-person data collection of speaker samples or listener tasks. The use of pre-recorded speaker samples reduced researcher control over the available stimuli, leading to inconsistencies in stimuli between speakers. Uniformity across speaker sample recordings would strengthen the methodological rigor. Similarly, reduced control over conditions of listener tasks, namely headphone quality and possible environmental noise, may have negatively impacted the quality of listener's participation. However, asynchronous and virtual platforms have been shown to be effective and valid in other perceptual investigations of dysarthria (Lansford, Borrie, and Bystricky, 2016).

Future studies should seek to have a larger group of speakers to more accurately represent the heterogeneity of the ataxic dysarthria population, such as speakers with a variety of etiologies, including tumor, stroke, TBI and degenerative processes. Perceptual judgments from the free-classification approach should be compared to a comprehensive set of speech metrics to facilitate interpretation of clusters. Additionally, it would be interesting to examine the free- and guided-classification paradigms with listeners who are experienced clinicians and/or dysarthria experts.

References

- Ackermann, H., & Hertrich, I. (1994). Speech Rate and Rhythm in Cerebellar Dysarthria: An Acoustic Analysis of Syllabic Timing. *Folia Phoniatrica Et Logopaedica*, 46(2), 70-78.
- Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. *The Cerebellum*, 6, 202-213. doi:10.1080/14734220701266742
- Ackerman, H., Vogel, M., Petersen, D., & Poremba, M. (1992). Speech deficits in ischaemic cerebellar lesions. *Journal of Neurology*, 239(4), 223-227.
- Ballard, K. J., Tourville, J. A., & Robin, D. A. (2014). Behavioral, computational, and neuroimaging studies of acquired apraxia of speech. *Frontiers in Human Neuroscience*, 8, 892.
- Bodranghien, F., Bastian, A., Casali, C., Hallett, M., Louis, E. D., Manto, M., ... van Dun, K. (2016, June). Consensus Paper: Revisiting the Symptoms and Signs of Cerebellar Syndrome. *Cerebellum*, 15(3), 369-391.
- Boutsen, F. R., Bakker, K., & Duffy, J. R. (1997). Subgroups in Ataxic Dysarthria. *Journal of Medical Speech-language Pathology*, 5(1), 27-36.
- Brendel, B., Synofzik, M., Ackermann, H., Lindig, T., Schölderle, T., Schöls, L., & Zeigler, W. (2015). Comparing speech characteristics in spinocerebellar ataxia type 3 and type 6 with Friedreich ataxia. *Journal of Neurology*, 262, 21-26. doi:10.1007/s00415-014-7511-8
- Civier, O., Tasko, S. M., & Guenther, F. H. (2010). Overreliance on auditory feedback may lead to sound/syllable repetitions: Simulations of stuttering and fluency-inducing conditions with a neural model of speech production. *Journal of Fluency Disorders*, 35(3), 246-279.
- Clopper, C. G. (2008). Auditory free classification: Methods and analysis. *Behavior Research Methods*, 40(2), 575-581.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Differential Diagnostic Patterns of Dysarthria. *Journal of Speech and Hearing Research*, 12(2), 246-269.

- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Clusters of Deviant Speech Dimensions in the Dysarthrias. *Journal of Speech and Hearing Research*, 12(3), 462-496.
- Duffy, J.R. (2020). *Motor speech disorders: Substrates, differential diagnosis, and management*, 4th edition. St. Louis: Elsevier.
- Ebert, D., Hefter, H., Dohle, C., & Freund, H. (1995). Ataxic breathing during alternating forearm movements of various frequencies in cerebellar patients. *Neuroscience Letters*, 193(3), 145-148.
- Gore, M. and Spencer, K. A. (November, 2020). Perceptual subgroups in speakers with ataxic dysarthria. Poster presentation accepted for the annual meeting of the American Speech-Language-Hearing Association. Conference cancelled.
- Hartelius, L., Runmarker, B., Andersen, O., & Nord, L. (2000). Temporal Speech Characteristics of Individuals with Multiple Sclerosis and Ataxic Dysarthria: 'Scanning Speech' Revisited. *Folia Phoniatica Et Logopaedica*, 52(5), 228-238.
- Imai, S., & Garner, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology*, 69(6), 596-608.
- Imai, S. (1966). Classification of sets of stimuli with different stimulus characteristics and numerical properties. *Perception & Psychophysics*. 1(1), 48-54.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2017). An introduction to statistical learning: with applications in R. New York: Springer.
- Joanette, Y., & Dudley, J. G. (1980). Dysarthric symptomatology of Friedreich's ataxia. *Brain and Language*, 10(1), 39-50.
- Kim, D., Diehl, S., de Riesthal, M., and Mefferd, A. (2020). Dysarthria subgroups in talkers with Huntington's disease: Free classification versus feature-constrained classification. Proceedings of the *Journal of the Acoustical Society of America*, 148(4).
- Lansford, K. L., Borrie, S. A., & Bystricky, L. (2, May). Use of Crowdsourcing to Assess the Ecological Validity of Perceptual-Training Paradigms in Dysarthria. *American Journal of Speech-Language Pathology*, 25, 233-239.

- Lansford, K. L., Liss, J. M., & Norton, R. E. (2014). Free-classification of perceptually similar speakers with dysarthria. *Journal of Speech, Language, and Hearing Research, 57*(6), 2051-2064.
- Maas, E., Mailend, M., & Guenther, F. H. (2015). Feedforward and feedback control in apraxia of speech: Effects of noise masking on vowel production. *Journal of Speech, Language, and Hearing Research, 58*(2), 185-200.
- Parrell, B. and Houde, J. (2019). Modeling the role of sensory feedback in speech motor control and learning. *Journal of Speech, Language, and Hearing Research, 62*, 2963-2985.
- Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired Feedforward Control and Enhanced Feedback Control of Speech in Patients with Cerebellar Degeneration. *The Journal of Neuroscience, 37*(38), 9249-9258.
- Pisotta, I. & Molinari, M. (2014). Cerebellar contribution to feedforward control of locomotion. *Frontiers in human neuroscience, 8*, 475.
<https://doi.org/10.3389/fnhum.2014.00475>
- Schalling, E., Hammarberg, B., & Hartelius, L. (2008). A longitudinal study of dysarthria in spinocerebellar ataxia (SCA): Aspects of articulation, prosody, and voice. *Journal of Medical Speech-Language Pathology, 16*(2), 103-117.
- Sidtis, J. J., Ahn, J. S., Gomez, C., & Sidtis, D. (2011). Speech characteristics associated with three genotypes of ataxia. *Journal of Communication Disorders, 44*, 478-492.
 doi:10.1016/j.jcomdis.2011.03.002
- Spencer, K. A., & Dawson, M. (2019). Dysarthria Profiles in Adults With Hereditary Ataxia. *American Journal of Speech-language Pathology, 28*(2S), 915-924.
- Spencer, K. A., & France, A. A. (2016). Perceptual ratings of subgroups of ataxic dysarthria. *International Journal of Language & Communication Disorders, 51*(4), 430-441.
- Spencer, K., & Slocumb, D. (2007). The neural basis of ataxic dysarthria. *Cerebellum (London, England), 6*(1), 58-65.
- Tourville, J., & Guenther, F. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes, 26*(7), 952-981.

Vandana, V. P., & Manjula, R. (2015). Speech intelligibility in ataxic dysarthria due to lesions in different cerebellar loci. *Language in India*, 15(5), 381.

Wang, Y., Kent, R. D., Duffy, J. R., & Thomas, J. E. (2009, April). Analysis of Diadochokinesis in Ataxic Dysarthria Using the Motor Speech Profile Program. *Folia Phoniatrica et Logopaedica*, 61(1), 1-11.

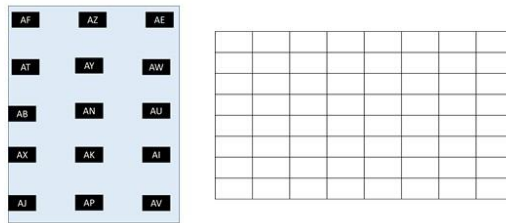
Yorkston, K.M., Beukelman, D., Hakel, M. (1996). Speech Intelligibility Test. Lincoln, NE: Tice Technology Services.

Appendix A: Listener Task PowerPoint Slides

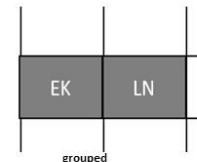
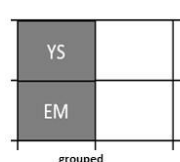
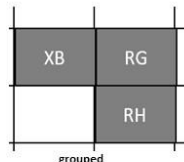
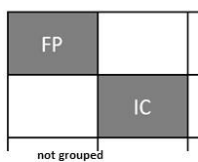
Free Classification Task Slides 1-6, counterbalance group A

Thank you for participating in our experiment!

- To get started, please make sure you are in a quiet environment and using the highest quality headphones possible.
- We will be asking you to listen to speech samples and group them according to how similar they sound. This is a challenging perceptual task with no right answers! Just do the best you can!
- On the following PowerPoint slides, you will find speaker files on the left and an empty grid on the right. Each speaker is given random initials. Here's an example:



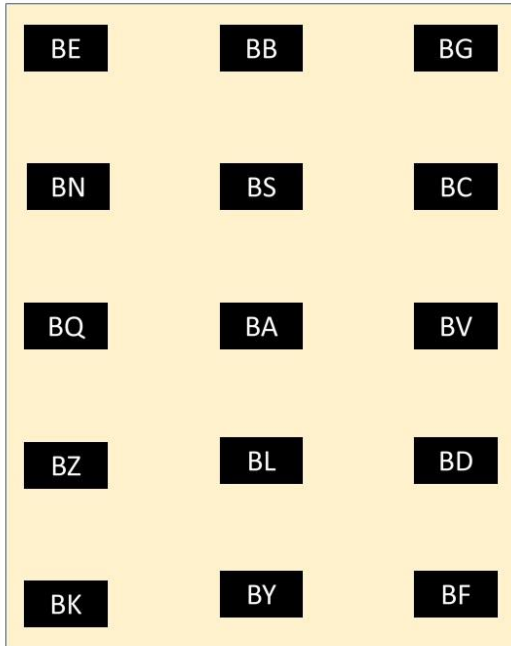
- To listen to a file, click on the one you would like to hear and press "play" when the control bar appears.
- Each file represents an individual speaker with ataxic dysarthria.
- On one slide, the speakers are performing diadochokinetics (repetitions of /p/, /t/, and /k/). On the other slide, they are saying 1-2 sentences.
- Group the files by clicking and dragging them into the grid depending on how similar they sound to you. The PowerPoint must remain in "Edit" mode, rather than Slideshow, for you to make your groupings.
- You can listen to each file as many times as necessary.
- When grouping the speaker files, be sure that speakers in the same group are touching in the grid. The grid does not have specific axes or dimensions, so files can be placed anywhere. See examples below:



- Keep in mind that speakers are a mix of males and females, and perceived sex should not be used as a factor in your grouping decisions.
- You are permitted to make as many groups as needed with as many speakers in each group as needed.
- There is no time limit. You can listen to and re-arrange the files as many times as necessary.
- The notes section at the bottom of the PowerPoint slide is available for you to use however you see fit, but notes are not required.
- **The final product should be groups of speaker files that sounded similar to you.**
- When you are finished, save the file and add an underscore followed by your assigned listener number to the end of the file name (e.g., FREE_A_ListenerTask_L6) then upload it to the google form. If you have any questions before beginning, please contact Jessica Amaral at (510)714-8409 or jramaral@uw.edu.

Thank you in advance for your help!

AF	AZ	AE
AT	AY	AW
AB	AN	AU
AX	AK	AI
AJ	AP	AV



You are halfway done!

- Please save this PPT file and add your listener number to the end of the file name as indicated on slide 3, then upload it to the Google Form.
- Once this file is uploaded, you can advance in the Google Form to receive the second task.
- Thank you so much and keep up the great work!

Welcome to Task #2!

- You will see a setup similar to the first task.
- One slide is for diadochokinetics and the other is for 1-2 sentences of connected speech.
- The PowerPoint must remain in “Edit” mode, rather than Slideshow, for you to make your groupings.
- When you are finished, save the file, add your listener number to the end of the file name (e.g., GUIDED_A_ListenerTask_L6), then upload it to the google form. If you have any questions before beginning, please contact Jessica Amaral at (510)714-8409 or jramaral@uw.edu.

This time, we will ask you to group the speakers according to whether they sound more **variable** or more **robotic**....

VARIABLE group: These speakers have more unstable motor control. They are typically more variable with sound errors, stress, rhythm, and pitch. They tend to be faster or have an inconsistent rate. They may have incoordination of breathing and speaking.

ROBOTIC group: These speakers have more inflexible motor control. They typically have more equalized stress and rhythm, and more consistent errors. They also tend to be slower. They may have monopitch and monoloudness.

Try to group all speakers as VARIABLE or ROBOTIC. However, if you feel a speaker doesn't fit either profile, you can put them in the **INDETERMINANT group.**

VARIABLE group: more variable sound errors, stress, rhythm, pitch; tend to be faster. Incoordination breathing & speaking.
ROBOTIC group: more equalized stress and rhythm, and more consistent errors; tend to be slower. Monopitch/monoloudness.

CN	CJ	CF
CD	CC	CQ
CM	CS	CW
CT	CV	CG
CX	CK	CR

Variable			Robotic			Indeterminant	

VARIABLE group: more variable sound errors, stress and rhythm; tend to be faster. Incoordination breathing & speaking.
ROBOTIC group: more equalized stress and rhythm, and more consistent errors; tend to be slower. Monopitch/monoloudness

DH	DC	DY
DN	DU	DV
DG	DJ	DS
DF	DL	DI
DB	DQ	DK

Variable			Robotic			Indeterminant	

Tasks complete!

- Please save this PPT file, add your listener number to the file name, upload it to the Google Form, and click submit.
- Thank you again for your participation. Your compensation will be sent through your method of choosing (Venmo or PayPal). If you do not receive your compensation, please contact Jessica Amaral at (510)714-8409 or jamaral@uw.edu.

Appendix B: Optional listener notes organized by speaker and cluster.

Optional listener notes from the free-classification task with AMR recordings.

Cluster 1	S1	<ul style="list-style-type: none"> • Very arrhythmic/fast pace • quick, ok voice with some breathy at end, popcorn-esque • arrhythmic, fast, imprecise
	S2	<ul style="list-style-type: none"> • quick, popcorn sounds • Starts very fast with arrhythmic dds, Mostly steady/rhythmic • fast, imprecise, better rhythm
	S3	<ul style="list-style-type: none"> • slurred, frequent artic/coord errors, resp/phon • More incoordinated, More slurred, Sounds strained/stops midway runs out of breath? Waves of volume up and down/in and out. • quick, slurry sounds
	S4	<ul style="list-style-type: none"> • fast, imprecise, better rhythm • often, less steady rate, rushed/pressed speech, Voicing on and off, Discoordinated articulation?. breathy • fairly steady, quicker rate, weak/soft voice, consistent across sounds
	S5	<ul style="list-style-type: none"> • weak, slurry • breathy • oral low pressure on consonants, Inconsistent voicing, Slurred ts, Low breath support • slurred, frequent artic/coord errors, resp/phon
	S6	<ul style="list-style-type: none"> • quick ,breathy, fairly steady • Quiet, Steady, Consistent voicing • arrhythmic, fast, imprecise
	S7	<ul style="list-style-type: none"> • breathy, quick slurry sounds • breathy and quiet • K-g-k-g-g-k (inconsistent voicing) • arrhythmic, fast, imprecise
	S8	<ul style="list-style-type: none"> • breathy voice, quicker rate, slurry t's • Arrhythmic, pressed speech, slurred T • arrhythmic, fast, imprecise
	S9	<ul style="list-style-type: none"> • quick, prosody like baby babble, ok voice • slurred, frequent artic/coord errors, resp/phon
	S10	<ul style="list-style-type: none"> • breathy voice, quicker rate, challenge with k • Loses momentum • arrhythmic, fast, imprecise
Cluster 2	S11	<ul style="list-style-type: none"> • arrhythmic, voicing errors, vowel errors • Alternating inconsistent between p/b, t/d, and k/g • slow and irregular • steady, slow rate, loud voice
	S12	<ul style="list-style-type: none"> • Fairly steady, slower rate, breathy voice • Steady, Devoiced first p then voicing ps, G sound for some ks, Breathy • groggy • slower, Breathy p, then b, inconsistent voicing • fast, mostly accurate
	S13	<ul style="list-style-type: none"> • arrhythmic, voicing errors, vowel errors • Slow movement of articulators • Steady, Loud, More coordinated articulation, Not breathy • slow and irregular • not so steady, slower rate, loud voice, worse from p to t to k

	S14	<ul style="list-style-type: none"> • steady, ok voice • Voiced all, steadier, slower, Gs for K • better rate but voicing errors
	S15	<ul style="list-style-type: none"> • arrhythmic, voicing errors, vowel errors • Breathy/devoiced on one or two, Voicing (b's for p's), Mostly steady but seems to lose this steadiness once or twice, Slower, Breathy p, then b, inconsistent voicing • slow and irregular • fairly steady, slower rate, loud voice, consistent across sounds

Optional listener notes from the free-classification task with connected speech recordings.

Cluster 1	S1	<ul style="list-style-type: none"> • fluent, some pitch variation, some segmentation • rate of speech and volume variation • break, nasality • distortions stand out most
	S2	<ul style="list-style-type: none"> • slow, slightly monotone • slow, hypernasality, weakness • hypernasality, ok prosody • distortions stand out most
Cluster 2	S3	<ul style="list-style-type: none"> • altered resonance, slow • slurring, strangled, hypernasality • articulation stands out most
Cluster 1	S4	<ul style="list-style-type: none"> • slow, slightly monotone • hypernasality • distortions stand out most
Cluster 3	S5	<ul style="list-style-type: none"> • fluent, some pitch variation, some segmentation • Halting speech, slight pauses, minimally slurred speech • slow rate, breaks, hyponasality • distortions stand out most
Cluster 1	S6	<ul style="list-style-type: none"> • fluent, can achieve pitch variation • Rapid speech • ok prosody, ok everything • distortions stand out most
	S7	<ul style="list-style-type: none"> • slow, slightly monotone • prosody ok, sounds ok • distortions stand out most
	S8	<ul style="list-style-type: none"> • altered resonance, slow • Vocal fry • “pressed” quality—strained, some distortions • nasality, flat prosody, ok rate • articulation stands out most
Cluster 2	S9	<ul style="list-style-type: none"> • altered resonance, slow • strangle, break, nasality • articulation stands out most

Cluster 1	S10	<ul style="list-style-type: none"> • fluent, can achieve pitch variation • “pressed” quality—strained, some distortions • slight surring, ok prosody • distortions stand out most
Cluster 3	S11	<ul style="list-style-type: none"> • slow, slightly monotone • prosody minimal, slow rate • prosody stands out more
	S12	<ul style="list-style-type: none"> • segmented, monotone, effortful • more slurry quality, greater distortions • prosody stands out more
Cluster 2	S13	<ul style="list-style-type: none"> • segmented, monotone, effortful • Rhythmic with some inconsistent rate, Excess and equal stress, Abnormal prosody • slow, breaks, some nasality • articulation stands out most
	S14	<ul style="list-style-type: none"> • segmented, monotone, effortful • slurring, effort • prosody stands out more
	S15	<ul style="list-style-type: none"> • altered resonance, slow • greater impact on prosody • slow rate, hyponasality (nose plug) • articulation stands out most