



The Role of the Motor System in Speech and Language in Autism

Citation

O'Brien, Amanda M. 2025. The Role of the Motor System in Speech and Language in Autism. Doctoral Dissertation, Harvard University Graduate School of Arts and Sciences.

Link

<https://dash.harvard.edu/handle/1/42719737>

Terms of use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at

<https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFIOWI2M2EwYTg>

Accessibility

<https://accessibility.huit.harvard.edu/digital-accessibility-policy>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#)

HARVARD
Kenneth C. Griffin



GRADUATE SCHOOL
OF ARTS AND SCIENCES

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Division of Medical Sciences
Speech and Hearing Bioscience and Technology
have examined a dissertation entitled

The Role of the Motor System in Speech and Language in Autism

presented by Amanda M. O'Brien

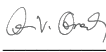
candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.

Signature: 
Nadine Gaab (Apr 30, 2025 14:39 EDT)

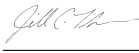
Typed Name: Dr. Nadine Gaab

Signature: 
Maria Mody (Apr 30, 2025 16:41 EDT)

Typed Name: Dr. Maria Mody

Signature: 

Typed Name: Dr. Karen Chenausky

Signature: 

Typed Name: Dr. Jill Thorson

Date: April 29, 2025

The Role of the Motor System in Speech and Language in Autism

A Dissertation Presented by

Amanda M. O'Brien

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Speech and Hearing Bioscience and Technology

Harvard University

Cambridge, Massachusetts

April 2025

© 2025 Amanda M. O'Brien

All rights reserved.

The Role of the Motor System in Speech and Language in Autism

Abstract

Communication challenges in autism are often attributed to social and cognitive factors, but emerging research suggests that motor processes also play a critical role. Despite growing interest in this area, the brain and behavioral underpinnings of speech-motor differences in autism, and their relationship to language, remain poorly understood. The present dissertation examined how motor system differences in the brain contribute to speech and language challenges in autistic individuals who use speech to communicate. This work integrated studies that bridge brain and behavior, while also incorporating the perspectives of autistic people.

The first study examined how three key language-relevant brain networks – speech perception, working memory, and speech production networks – related to language challenges in autistic children. FMRI results revealed that autistic children showed atypical and reduced engagement of the *speech production* network, but not more traditional language processing regions, during nonword repetition. This atypical motor engagement correlated with poorer task performance, higher autism traits, and lower social communication skills. The second study built on this finding to explore *how* motor system differences impact speech and language in autism beyond nonword repetition. Research in nonautistic individuals, including lesion-based work, has suggested the importance of key motor regions in both the production and perception of *prosody*. Thus, we turned to prosody as an area of language that may be impacted by brain-based motor differences. Prosody – the use of pitch, duration, and loudness in speech – plays a key role in communication, and atypical prosody is commonly noted in autism. Study two revealed that autistic individuals reported greater challenges with prosody, and that these challenges were

strongly correlated with autism traits in both autistic and nonautistic adults. Acoustic analyses showed that the autistic participants produced speech with atypical timing, with longer pauses and a slower speech rate, during a spontaneous picture description task. Notably, diadochokinetic and articulation rate did not differ between groups, suggesting that these prosody differences were unlikely due to an underlying speech-motor disorder (e.g., apraxia of speech), but may stem from broader motor system differences.

The third study further examined the motor-based brain differences in autism, by first precisely localizing each participants' speech production regions, and then examining how these functionally-defined regions were engaged during prosody perception. Results from this fMRI study revealed that nonautistic individuals engaged several of the same functional regions for speech production as prosody perception (i.e., listening to sentences with dynamic versus flat pitch contours). In contrast, autistic individuals showed a reduced activation in these regions. Significant group differences emerged across several motor regions, including the supplementary motor area. Moreover, activation in motor regions during prosody perception was positively correlated with participants' accuracy on an out-of-scanner prosody perception task, highlighting the behavioral relevance of this motor activation.

Taken together, these studies provide evidence that atypical engagement of the motor system contributes to language differences in autism. By reframing language differences in autism through the lens of speech motor system function, this dissertation offers new insights into the neurobiological mechanisms underlying language challenges. These findings lay a foundation for future research aimed at supporting autistic communication.

Table of Contents

Title Page	i
Copyright	iv
Abstract	iii
Table of Contents	v
List of Figures	vii
List of Tables	viii
Acknowledgements	ix
Dedication	xi
Chapter 1: Introduction	1
Overview of Autism	2
Key Terminology	4
Major Themes in Speech and Language Research in Autism.....	10
The Language and Motor Systems in Autism.....	14
The Current Dissertation.....	17
References.....	22
Chapter 2: Altered Engagement of the Speech Motor Network is Associated with Reduced Phonological Working Memory in Autism	28
Abstract	30
Introduction	31
Materials and Methods	39
Results.....	53
Discussion	60
References.....	68
Chapter 3: Acoustic and Self-Reported Prosodic Differences in Autistic Adults	82
Abstract	83
Introduction	84
Materials and Methods	92
Results.....	103
Discussion	111
References.....	122
Chapter 4: Precision fMRI Reveals Engagement of Speech Motor Regions During Prosody Perception and Differences in Autism	129

Abstract	130
Introduction	131
Materials and Methods	139
Results.....	152
Discussion	159
References.....	169
<i>Chapter 5: Discussion & Conclusions</i>	<i>176</i>
The Motor System, Language System, and Autism	179
Heterogeneity of Speech and Language Profiles in Autism.....	180
Motor Prediction and Autism	181
Considerations for Future Work	182
References.....	184
<i>Appendix A: Supplementary Information for Chapter 2.....</i>	<i>186</i>
<i>Appendix B: Supplementary Information for Chapter 4.....</i>	<i>197</i>

List of Figures

Figure 1.1. Visualization of the Intersections Between Communication, Language, Speech, and Prosody, as Used in the Present Dissertation

Figure 2.1. Flowchart Depicting Exclusionary Criteria and Numbers of Included Participants in Each Group

Figure 2.2. Task Design and Accompanying Visual Prompt

Figure 2.3. Networks of Interest (NOI) Derived from Neurosynth.org Used for the Univariate and Multivariate Analyses

Figure 2.4. A Schematic Representation of How Patterns of Activity May be Decoded Between Two Conditions (Two Syllable NWR and Five Syllable NWR) Within the Speech Production NOI

Figure 2.5. In-scanner Performance on Nonword Repetition Task

Figure 2.6. Effects of Syllable Length in fMRI

Figure 3.1. Self-Reported Prosody Challenges and Speech Features, Based on the VOX-Q

Figure 3.2. Primary Results From Acoustic Features Extracted From The Picture Description Task

Figure 3.3. Visualization of the Three Clusters Identified by the K-Means Clustering Approach

Figure 4.1. Task Design and Timing For Speech Production Localizer Task and Passive Prosody Perception Task

Figure 4.2. Activation in precisely localized speech production parcels during passive prosody perception.

Figure 4.3. Whole Brain Analyses for the Critical Prosody Contrast

Figure 4.4. Correlations Between Brain Activation and Out-of-Scanner Prosody Task

List of Tables

Table 1.1. Research Questions, Aims and Hypotheses for Each Study Comprising the Dissertation

Table 2.1. Phenotypic Cognitive, Behavioral, Language, and Reading Characteristics in the Primary Analysis Group (NT and ASD), as well as for the RD Groups

Table 2.2. Peak Coordinates and Anatomical Locations of Significant Clusters for Within and Between Group Analyses for Five-Syllable > Two-Syllable Contrast

Table 2.3. Decoding Accuracy for NT and ASD Groups Within the NOIs, Intersections of Interest, Complement Regions, and Combination of All Networks

Table 3.1. Participant Characteristics by Group

Table 3.2. Summary of Acoustic Features Extracted as Measures of Global Prosody From Speech Samples

Table 4.1. Group Averages and Standard Deviations of Participant Demographic and Assessment Characteristics

Table 4.2. Activation Difference for the Critical Prosody Contrast (Dynamic > Flat) in Motor Parcels

Table 4.3. Regions of Significant Activation for the Critical Prosody Perception Contrast in Whole Brain Analysis

Table 5.1. Research Questions, Aims, Hypotheses, and Key Findings for Each Study Comprising the Dissertation

Acknowledgements

At the heart of my PhD experience are the people – my mentors, collaborators, colleagues, family, and friends – who have contributed to this work, guided my thinking, and enriched my life in countless ways.

First and foremost, thank you to my advisor, **John D. E. Gabrieli**. John, your unwavering encouragement, insight, and brilliance over the past six years have meant so much. I am deeply grateful for the trust that you placed in me, which provided me the freedom to explore a wide range of research directions and build an incredible network of mentors and collaborators in the process. Thank you for cultivating a lab culture of kindness in which creativity can thrive.

Thank you to my Dissertation Advisory Committee: **Nadine Gaab, Jordan Green, Carol Wilkinson**. Nadine, your constant support and guidance throughout my time in the SHBT program have been pivotal to my success. Thank you for being a role model to me. Jordan, thank you for always offering fresh perspectives, and for encouraging me to shape this work into the cohesive dissertation has become. Carol, your clinical insight, feedback, and enthusiasm for this work has been so valuable and appreciated.

Thank you to my Dissertation Examination Committee: **Maria Mody, Jill Thorson, and Karen Chenausky**. I'm grateful for the time you have dedicated to evaluating and providing feedback on this work, much of which was inspired by yours.

Thank you to my F31 team: **Zhengan Qi, Frank Guenther, Helen Tager-Flusberg, and Rhiannon Luyster**. Zhengan, having the joy of learning from you has been a true highlight of my PhD years – thank you. Frank, it was an absolute privilege to receive your insight and feedback on my F31 project. Helen, learning from you has been a true gift. The rich body of work you have produced throughout your career continues to inspire me. Rhiannon, thank you for your mentorship related to the ADOS, learning from you has been such an enriching opportunity.

Thank you to the many other mentors who have supported me along the way: **Ev Fedorenko, Pawan Sinha, Jennifer Zuk, Tyler Perrachione, Stefanie Shattuck-Hufnagel, Sofia Vallila-Rohter, and Satra Ghosh**. When I started this PhD journey in 2019, I could never have imagined how many brilliant and thoughtful mentors I would have the pleasure and privilege of learning from.

Science is a team sport, and I have been incredibly fortunate to be surrounded by the best teammates. Working on the autism-projects team in the Gabrieli lab with **Cindy Li, Shruti Nishith, Luwei Quan, and Linday Bungert** has been both professionally and personally enriching. And an extra special thank you to **Annie Cardinaux**, who has been a true partner through these past six years. From juggling our ever-growing bag(s) of projects to navigating the many world changes – I could not picture this experience without you (nor would I want to try).

Thank you to the many lovely and brilliant lab members who have become friends, with special gratitude to **Sadie Zacharek, Halie Olson, and Steven Meisler**. Your friendship paired with writing dates, mug salads, excessive bowls of olives, and deliciously overpriced lattes got me through some/many challenging days.

A giant thank you to **Anila D'Mello**. You been an incredible mentor, as well a trusted and beloved friend. I am so grateful that our paths crossed and that we kept them intertwined with HP. I'm looking forward to the other collaborations in our future.

I am also grateful for the many post-docs I have had the joy of working with and learning from: **Alexandra Decker, Rebecca Marks, Jonathan Cannon**. And special thanks to **Ola**

Ozernov-Palchik: you have been a role model for me in so many ways and **Kristina Johnson**, for bringing me into Commalla and inspiring the best commuting conversations.

Thank you to the members of the McGovern Imaging Center, especially **Atsushi Takahashi** and **Steven Shannon**. I am so thankful for all the support you provided in setting up the project in Chapter 4, advising on unforeseen issues, and for sharing in engaging conversations during scans.

Thank you to the **SGBT program**, which has fostered such a special community of interdisciplinary scientists. I am grateful to my SGBT cohort, particularly **Ayelet Kershenbaum**. You are a true gem, and your friendship, comradery, thoughtfulness, and deep insight have been a constant source of support and inspiration during this journey. Thanks for picking up my strangely timed phone calls.

I would also like to thank my dear colleagues at Boston Children's Hospital, especially my mentor **Howard Shane**, whose continued support has meant so much. Thank you also to **Ralf Schlosser**, who instilled in me the value and importance of conducting rigorous science.

I would also like to thank **Meera Modi**, **Mustafa Sahin**, and **Charles Nelson**, who provided an invaluable opportunity for me to become involved in their research at Boston Children's Hospital, an experience that truly cemented my desire to pursue a PhD.

Thank you to my undergraduate mentor, **Jeremy Skipper**, who inspired me to question things deeply, and fostered my fascination with language in the brain.

I am deeply thankful for the **participants** who made this science possible. Thank you to the Simons Foundation Autism Research Initiative (SFARI) & Simons Foundation Powering Autism Research (**SPARK**) for facilitating recruitment for the study in Chapter 2. I am also incredibly thankful to the **autistic consultants** who participated in a focus group discussion. Your insights and perspectives were incredibly valuable in shaping this research.

Thank you to the **funding sources** who have enabled the research presented here, including The Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT, funding from the NIDCD (Institutional Training Grant T32DC000038 and Fellowship F31DC020864), and from The Amelia Peabody Award from Harvard University.

I am also grateful to the academy, though imperfect, it has been a safe haven for me amidst both joyful and challenging times. May it continue to be a safe place to land for those who follow.

Lastly, but most importantly, I thank my family. I am especially thankful to my **parents**, who taught me the value of hard work from a young age and who have always believed in me, and my brother, **Josh**, for paving the way for me in so many ways.

I am endlessly grateful to my husband and life partner, **Nick**, without whom this dream might have stayed *but a dream*. Thank you for keeping me anchored (metaphorically) while also encouraging me to set sail (metaphorically and literally), and for listening to all of my (many) presentations. And of course, I am deeply thankful for my daughter **Sophia**, who has brought so much joy to our lives, reminded me of the wonder of speech and language development, and who has given my life and work new purpose.

*This dissertation is dedicated to
my grandmothers, Eleanor Marie and Rose Marie, who sowed the seeds of my future,
and my daughter, Sophia Marie, for whom I now tend the garden.*

Chapter 1: Introduction

Introduction

Communication challenges in autism spectrum disorders (henceforth autism¹) are often attributed to social and cognitive factors. However, emerging research suggests that brain regions that are not typically considered in language processing – namely, speech motor regions – may also play a critical role in developmental differences in autism. Despite growing interest in this area, the brain and behavioral underpinnings of speech motor differences in autism remain poorly characterized. This dissertation investigates how *motor system* differences may contribute to speech and language challenges in autistic² individuals who use speech to communicate. Combining behavioral and brain-based research with perspectives from autistic individuals, this dissertation offers fresh insights and lays the groundwork for future research to support autistic communication.

Overview of Autism

Diagnostic Criteria for Autism

Autism is a heterogeneous neurodevelopmental condition that currently impacts 1 in 36 children (Centers for Disease Control [CDC], 2024). Autism is diagnosed based on behavioral differences in two key areas: (1) communication and social interaction and (2) restricted and repetitive behaviors and interests (American Psychiatric Association [APA], 2013). Social communication challenges may include difficulties or differences in social reciprocity, nonverbal communication behaviors, and the formation and maintenance of relationships. In the domain of restricted and repetitive behaviors, autistic individuals may engage in repetitive motor movements

¹ “Autism” will be used henceforth to refer to autism spectrum disorder for concision and in response to self advocates’ stated preferences (Kenny et al., 2016).

² This dissertation primarily uses identity-first language (“autistic person”) rather than person-first language (“autistic person”), consistent with the preferences of many – but not all – autistic people (Lei et al., 2021; Taboas et al., 2023).

or speech, struggle with changes in routines, develop strong personal interests, and exhibit hyper- or hypo-reactivity to sensory input. These characteristics must be present early in life – though they may not become fully evident until later when social demands increase – and must cause significant functional challenges (APA, 2013). Importantly, autism is a spectrum-based condition, meaning that individuals experience these symptoms differently and to varying degrees.

Beyond the core symptoms outlined in the diagnostic criteria, autistic individuals may have differences in language, cognition, and adaptive functioning. To account for these variations, clinicians can apply diagnostic modifiers that reflect the extent of these challenges (Lord et al., 2020). For example, both the Diagnostic Statistical Manual - Fifth Edition (DSM-5; APA, 2013) and the International Classification of Diseases–11th Revision (ICD-11; World Health Organization [WHO], 2022) allow clinicians to specify whether an individual has an accompanying language impairment and/or intellectual disability. Other conditions (e.g., anxiety, depression, attention deficit hyperactivity disorder, epilepsy) frequently co-occur with autism, adding further variability to each individual’s specific phenotype (Mosner et al., 2019).

Diagnostic Connections with Speech, Language, and Motor Differences

A central focus of the present dissertation is the *speech, language, and motor abilities* of autistic individuals, which are notably absent from the two core diagnostic domains described above. While speech and language impairments are no longer included as diagnostic criteria (APA, 2013; WHO, 2022), they were part of earlier criteria (e.g., APA, 1994) and continue to be among the first concerns that caregivers report (Richards, Mossey, & Robbins, 2016). Moreover, language is a primary predictor of long-term functional outcomes of autistic individuals (Miranda et al., 2023; Magiati et al., 2014), making it a critical domain for study. Additionally, although not comprising a core diagnostic domain, speech and language differences are clearly represented in

both key diagnostic areas (communication and social interactions; restricted and repetitive behaviors). For example, “difficulties with back and forth conversation” are included as criteria for social communication challenges, and “stereotyped speech” is included as a criterion for restricted and repetitive behaviors (APA, 2013).

Like language differences, motor differences are not included as core diagnostic criteria for autism, but are frequently observed in autistic individuals. Some researchers have even advocated to add motor considerations into autism screening and diagnosis (Zampella et al., 2021). Similar to language challenges, motor concerns are some of the earliest parent-reported signs of autism (Guinchat et al., 2012), and may serve as useful diagnostic clues (Harris, 2017). Examples of motor-related features are embedded into both core diagnostic domains: “difficulty using or understanding gestures for communication” falls under communication challenges, and “stereotyped or repetitive motor movements” is included under restricted and repetitive behaviors. Thus, although language and motor differences are not core diagnostic criteria, examining them both independently and in relation to one another may offer important insights into the underlying neurobiological mechanisms, clinical presentation, and lived experiences in autism.

Key Terminology

The present dissertation explores the intersection between speech, language, and communication for autistic individuals, and their potential relationship to motor system differences. For the purposes of the present work, explicit working definitions of these terms are provided below.

Communication, Language, Speech, and Prosody

In this dissertation, *communication* is used as a broad term that encompasses both speech and language, particularly as they occur within the context of social interactions (**Figure 1**).

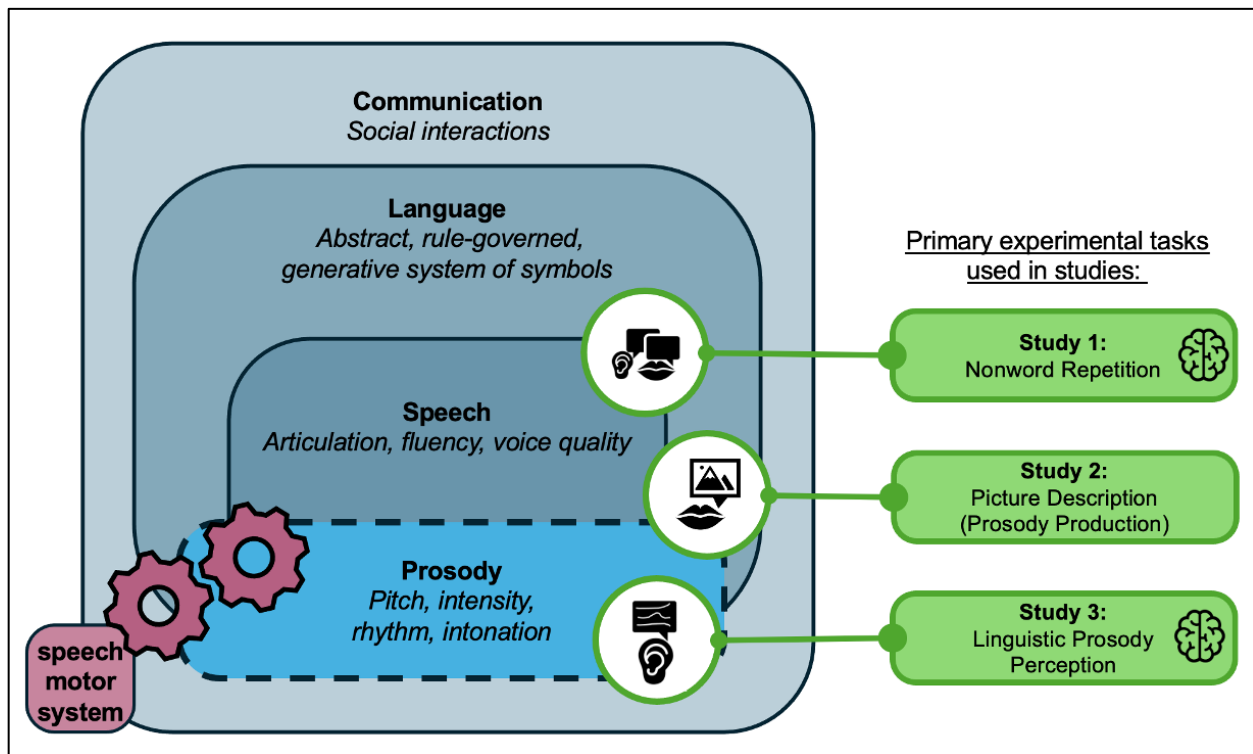
Communication is defined as the “active process of exchanging information and ideas,” involving the expression and comprehension of a message in a dynamic and reciprocal exchange (Beukelman & Miranda, 2013).

Language is an abstract, rule-governed system of generative symbols. Bloom & Lahey’s (1978) classic definition of language described three components: form, content, and use. *Form* encompasses phonology (the segmental features, such as phonemes and syllables, as well as suprasegmental features, including intonation, stress, and pauses), morphology (the smallest units of sound that carry meaning, such as words and inflection markers), and syntax (how the words and inflections are organized into sentences; Bloom, 1980). *Content* refers to semantic information, including knowledge of objects and their relations. Finally, *use* pertains to the pragmatic functions for which an individual communicates (e.g., commenting, requesting), and the contexts in which communication occurs.

Speech, on the other hand, is sometimes referred to as the “spoken mode of language” (Chenauksy & Tager-Flusberg, 2022), and involves the use of the vocal tract to articulate a linguistic utterance (Mermelstein, 1973). Speech includes articulation (how sounds are produced with the mouth, lips, and tongue), voice (how the vocal folds and breath are used to create sound), and fluency (the rhythm of speech, including use of pauses and hesitations). As shown in **Figure 1.1**, speech is one modality for expressing language within communicative settings. Notably, other strategies beyond speech (e.g., writing, manual signs, non-verbal gestures, facial expressions) may also support communication. While these other modalities of communication are not the focus of the present dissertation, their contributions to communication, and relation to autism, are important considerations for future work.

Figure 1.1.

Visualization of the Intersections Between Communication, Language, Speech, and Prosody, as Used in the Present Dissertation.



Note: A primary goal of this dissertation is to understand the role of the speech production motor system (illustrated in purple) across the levels of speech, language, and communication in autism. Prosody spans across these levels (illustrated in bright blue). The experimental tasks used in the dissertation to probe these processes include nonword repetition, picture description, and prosody perception (illustrated in green).

Prosody – often referred to as the melody of speech – spans and intersects with speech, language, and communication. Prosody is often considered to be a component of *language* (within the category of “*Form*” described above) that is conveyed via acoustic features of *speech*, primarily via voice and fluency. These features are used to *communicate* emotional and/or linguistic information. Prosody can be examined at multiple levels of granularity: the specific suprasegmental features (e.g., pitch, duration, volume, aspects of timing), what function these cues signal based on their placement and context (e.g., emphasis, emotion), and how the cues are interpreted by the listener (Dahan, 2015). Additionally, the acoustic features of prosody can be examined across different scales of communication. *Global* prosodic features refer to those

acoustic features of speech that can be extracted across an utterance or narrative, while *local* prosodic features refer to the acoustic features of speech that are specific to individual segments of speech (e.g., emphasis on a particular speech sound, rising pitch at the end of an utterance to convey a question). In the present dissertation, speakers' production of global prosody (**Chapter 3**) and listeners' perception of local prosody (**Chapter 4**) are examined.

The Speech Production Motor System

The act of speaking is one of the most motorically complex behaviors that humans regularly perform. Speech production is supported by the complex orchestration of many cortical and subcortical regions, including those supporting the integration of auditory, somatosensory, motoric, and linguistic functions. Many models of speech production have been proposed to account for speech production, starting with Paul Broca's discovery that damage to the left inferior frontal cortex (i.e., "Broca's area" or Brodmann areas 44 and 45), comprising the pars opercularis and pars triangularis, results in speech production difficulties (Broca, 1861). A more complex model of speech production that gained traction in the field of psycholinguistics is the Levelt model (1999). This model walks through stages from *conceptual preparation* of a lexical concept to *articulation* of a word. According to this model, speech production involves the following: preparation of a lexical concept to selection of a "lemma" (word), encoding of the morphemes (smallest unit of meaningful sound, translated into syllables) in that word, encoding the phonetics of that word and the corresponding phonetic gestures that are necessary to produce the target word, and generating the muscle activations that are necessary to articulate the word. The final two stages of this model, phonetic gesture encoding and articulation, are the key motor processes examined in the present dissertation.

Researchers continue to refine these early models of speech production, including through the investigation of the neurobiological mechanisms underlying speech production. A prominent neural network-based model of speech motor control, with supporting evidence from both simulations and neurophysiological data, is the Directions into Velocities of Articulators (DIVA) model (Tourville & Guenther, 2011). At its core, the DIVA model proposes that speech production relies on both feedforward and feedback motor control systems. These motor control systems include a series of maps, which represent a set of neurons in the brain that are believed to be largely left-lateralized. According to the DIVA model, speech production begins with the activation of a specific sound in the “speech sound map,” corresponding to the left ventral premotor cortex [vPMC]) in the brain. This speech sound map then sends corresponding motor commands to the motor cortex via a *feedforward control system* which generates motor-commands and a *feedback control system* which monitors ongoing speech. The *feedforward control system* first initiates the motor program with temporal precision via a cortico-basal ganglia loop with the supplementary motor area (SMA). Once initiated, the learned motor program itself is activated, generating the commands for actually *producing* the speech sounds, via the articulator map in the ventral primary motor cortex. Additionally, the *auditory feedback* and *somatosensory feedback* systems (in the posterior auditory cortex and ventral somatosensory cortex, respectively) detect differences (i.e., errors) between expected and actual auditory and tactile/proprioceptive information from the articulators. If an error is detected, these error maps send a signal to the feedback control map (vPMC) to modify the motor commands and correct the error in real time.

The “Motor System” and “Motor Processes”

The present dissertation remains intentionally agnostic regarding the specific stage(s) or locations in the speech production system at which motor differences may arise, as current research

is insufficient to make these stronger claims. Therefore, throughout the dissertation, the “speech production motor system” (sometimes shortened to just the “motor system”) and “motor processes” are described to generally reference the network of regions that support speech production. The aim here is to investigate the possibility that differences exist in these speech production motor regions in the first place, thus forming the foundation for future research in this area.

Motor Speech Disorders

Throughout the dissertation, reference is made to *motor speech disorders*, and whether these disorders may be related to prosody differences in autism. Motor speech disorders are caused by damage to the systems that control the production of speech, with the two primary motor speech disorders being apraxia and dysarthria. Apraxia of speech (also known as childhood apraxia of speech [CAS]) is caused by challenges with planning and execution of motor commands (American Speech and Hearing Association [ASHA], 2007). CAS can be caused by brain injury (particularly in speech motor planning regions of the brain) or as part of a genetic disorder. In contrast to CAS, dysarthria is marked by muscle weakness, leading to imprecise and/or unintelligible articulation, due to damage to lower or upper motor neurons (ASHA, *n.d.*). Differential diagnosis of motor speech disorders can be challenging, though efforts have been made to improve diagnostic agreement among clinicians (Iuzzini-Seige et al., 2022).

Given the observed overlap between CAS and autism (described in more detail below), it is likely that if an autistic participant in the present studies has a motor speech disorder, it would be CAS. That said, the motor speech screening approaches used in the present dissertation are insufficient for providing a differential diagnosis. Thus, as with the “motor processes” described

above, reference to “motor speech disorders” remains intentionally vague, though CAS is the most likely speech motor disorder under consideration.

Major Themes in Speech and Language Research in Autism

Like other features of autism, language profiles of autistic individuals vary considerably: some individuals present with no functional language differences, most individuals with ASD present with borderline or impaired language abilities, and up to a third develop limited or no speech (Kjelgaard & Tager-Flusberg, 2001). Of note, the studies comprising the present dissertation do not include any individuals with limited or no speech, and thus results may have limited generalizability to this historically understudied subgroup of autistic individuals.

Speech Sound Disorders, Motor Speech Disorders, and Autism

Do autistic individuals differ in their production of the segmental features of speech, such as speech sounds and syllables? A landmark study by Kjelgaard & Tager-Flusberg (2001) characterized the speech and language profiles of 89 autistic children, primarily including participants who used spoken language to communicate. The authors reported substantial variation in language skills across participants, but found that speech sound articulation skills – as measured by the *Goldman-Fristoe Test of Articulation* (Goldman & Fristoe, 1986) – were in the normal range across *all* language ability subgroups. However, articulation errors have been reported in autistic individuals. Shriberg et al. (2012) observed a higher number of speech errors in a group of autistic children (ages 6-9 years) compared to population estimates. Still, articulation skills have received limited attention in autism, likely due its characterization as an area of relative strength and limited functional impact.

Beyond articulation at the single word level, researchers have examined the co-occurrence of childhood apraxia of speech (CAS) in autism. CAS is a neurological speech sound disorder that

affects the *planning* and *programming* of motor sequences for speech. According to criteria set forth by the ASHA, CAS is diagnosed based on inconsistent speech errors, lengthened transitions between segmental speech features, and atypical prosody (ASHA, 2007). In a large-scale study of children with CAS ($N = 375$), 34% also had an autism diagnosis, though co-occurring autism did not predict the severity of CAS (Chenauksy et al., 2022). Rates of co-occurring CAS in autistic individuals have been less established. The CAS-ASD hypothesis posits that praxic deficits in autism may explain why some autistic children do not develop speech (the “strong” form of the hypothesis), and why others on the spectrum have atypical prosody and/or voice (the “weak” form of the hypothesis) (Shriberg et al., 2012). Shriberg (2012) did not find perceptual or acoustic evidence to support the “weak” form of the hypothesis in autistic children who use speech to communicate; the story may differ in children with minimal or no speech (per the “strong” form of the hypothesis). Research including minimally-verbal and low-verbal autistic children provides evidence for subgroups of minimally verbal children with CAS (Chenauksy et al., 2019; Chenauksy et al., 2021). A conceptual review of the literature identified the need for increased “deep speech phenotyping” of neurodevelopmental conditions to better understand speech development across these conditions (Chenauksy & Tager-Flusberg, 2022)

Finally, related work has focused more broadly on oral motor function (i.e., mouth movements) in autism. Indeed, a scoping review of studies investigating oral motor functioning in autistic adults found that 81% of studies reviewed ($N = 108$) reported significant oromotor differences in autistic individuals (including speech production, nonspeech oral motor skills, and feeding) (Maffei et al., 2023).

Taken together, this body of work suggests the need for further investigation of the speech production-related motor differences in autism, how these profiles may differ in autistic

individuals with and without spoken language, and how these speech motor differences may impact communication more broadly.

Language and Autism

A very brief overview of the literature related to language in autism is provided here to contextualize the more specific language foci in Chapters 2, 3, and 4. Investigations of language and autism have largely been focused on “structural” (i.e., syntax, grammar) and “pragmatic” (i.e., social-linguistic skills) components of language. Broadly speaking, pragmatic language challenges have been found more robustly across autistic individuals, while structural language differences are more variable across individuals. Notably, structural language skills are strongly associated with pragmatic skills in autistic children (Reindal et al., 2023).

Structural Language. Given the variability of language abilities in autism, a common goal amongst autism language researchers has been to investigate subgroups of language abilities, using both standardized (e.g., Kjelgaard & Tager-Flusberg, 2001, described above) and non-standardized measures (Wittke et al., 2017). A moderately-sized study ($N = 51$) of 5 year-old autistic children used sensitive naturalistic markers of language impairment via a spontaneous speech sample (Wittke et al., 2017). This study identified three sub-groups of structural language profiles in autistic children who used speech to communicate: “Language Normal” (26%), “Grammatical Impairment” (21%), and “Language Impairment” (16%), and a fourth group of children who were minimally or non-speaking at age 5 years. The Grammatical Impairment group was generally matched to the Language Normal group in nonverbal IQ and vocabulary, but had significant differences in its morphosyntactic production (errors in verb, noun, and pronoun morphemes). In contrast, the Language Impairment Group was notable in its lower scores across language metrics (shorter mean length utterances, smaller vocabulary, fewer morphemes produced), as well as non-

verbal IQ. Work related to semantics (i.e., word meanings) has also shown variability among autistic individuals. Some studies have identified specific contexts in which word learning is challenging for autistic children, including in learning words with multiple meanings (Floyd et al., 2021) and extending the meanings of words in unusual ways (Cuneo et al., 2022). Of note, this work has largely been conducted in children. The heterogeneity of structural language profiles in autism underscores the importance of careful characterization of language in autism research to better understand for whom specific findings are most applicable, and to avoid overgeneralizing results across a diverse population.

Pragmatic Language. Pragmatic language refers to the appropriate *use* of language within social settings (Matthews et al., 2018), as defined by particular social or cultural contexts. A substantial amount of work has focused on pragmatic components of language, given the relevance to the social communicative diagnostic challenges of autism. Unsurprisingly, the pragmatic language of autistic children has been shown to be distinct from nonautistic peers (Baixauli-Fortea et al., 2019) as well as individuals with language disorders (e.g., Developmental Language Disorder). A systematic review of 13 studies comparing pragmatic language in children with autism and developmental language disorder identified key differences in pragmatic language between the two conditions (Andreou et al., 2022). While both groups demonstrated similarities in certain pragmatic areas (e.g., reduced coherence when producing narratives, stereotypical language, reduced inference skills), the autistic individuals showed additional challenges (e.g., significant reduction in use of context during narration, increases in atypical communicative initiation, atypical nonverbal communication strategies). Pragmatic differences may be particularly notable during conversations. A systematic review of studies examining conversational skills in autism revealed that autistic individuals showed greater difficulty with

providing novel and pertinent information within a conversation, a reduction in initiation and responses during conversations, and tendency to prefer particular topics of conversation (Ying Sng et al., 2018).

Notably, there is a growing body of evidence supporting “The Double Empathy Problem” (Milton, 2012, Milton et al., 2022), which proposes that communication breakdowns between autistic and nonautistic individuals are the result of bidirectional misunderstandings on both parts, rather than the result of pragmatic “impairments” in autistic people. This literature suggests that autistic individuals can engage in higher quality interactions and more effectively share information with other autistic compared to nonautistic people (Crompton et al., 2020; Morrison et al., 2020). Consideration of these bidirectional misunderstandings is important for better understanding and ameliorating the communication breakdowns experienced by many autistic individuals.

Phonological Working Memory, Prosody, and Autism

The present dissertation uses measures of phonological working memory (as measured by nonword repetition) and prosody (production and perception) as the experimental tasks for understanding the role of the motor system in speech and language in autism. Chapter 2 provides a thorough overview of phonological working memory and nonword repetition, Chapter 3 provides an in-depth discussion of the behavioral literature related to prosody and autism, and Chapter 4 provides a summary of the neuroimaging literature related to prosody and autism.

The Language and Motor Systems in Autism

Growing evidence suggests that *motor* differences are a clinically meaningful feature of autism, including within the domain of *communication* (Craig et al., 2021; Zampella et al., 2021). In fact, speech and language (expressive and receptive) abilities in autistic individuals are

correlated with oromotor, manual motor, fine motor, and gross motor skills (Gernsbacher et al., 2008; Bal et al., 2020; Bhat et al., 2023; Mody et al., 2017; Bedford et al., 2016). Despite this apparent overlap, language and motor systems in the brain have largely been studied in isolation until recently. Here, the distinct literatures related to the language and motor systems in autism are briefly reviewed separately, and then considered in relation to each other.

The Language System in Autism

In decades of autism neuroimaging work, one of the most robust and replicated findings pertaining to the language system has been that of atypical lateralization. Group and individual-level analyses have shown that autistic individuals have stronger right hemisphere activation in language regions during language processing, including across paradigms and neuroimaging methods (Eyler et al., 2012; Kleinhans et al., 2008; Jouravlev et al., 2020). This finding has been extended to also demonstrate that autism-traits in nonautistic individuals (as measured by the Autism Quotient) were correlated with reduced leftward lateralization (Jouravlev et al., 2020). A large-scale meta-analysis of structural connectivity of the language network in autism ($N = 831$ autistic individuals) also revealed a leftward bias in atypical white matter connectivity, with greater reductions in white matter integrity in language associated tracts in the left hemisphere than the right (Li et al., 2022). This meta-analysis also supported the “under-connectivity” hypothesis of autism, based on widespread markers of atypical microstructure in language tracts. A systematic review of functional connectivity of the language system in autism highlighted local over-connectivity of language networks, and global under-connectivity of language networks with other networks (Larson et al., 2025). Taken as a whole, these results suggest disruption to the general organization of the language network in the brain that is characterized by differences in activation as well as structural and functional connectivity.

The Motor System in Autism

Motor delays and atypical movement development (e.g., stereotyped movement) have been identified as some of the earliest signs of autism, which often emerge even before the classical social communication differences (West, 2019; Posar & Visconti, 2022). What differences in the brain underlie these motor challenges? A systematic review by Stillesjö et al. (2025) of fMRI studies investigating action execution and observation in autism found functional differences (both hyper- and hypo-activation) in several key motor regions during action execution (e.g., precentral cortex, inferior frontal gyrus [IFG], middle temporal gyrus [MTG], and cerebellum) and action observation (e.g., bilateral IFG, right precentral gyrus, MTG). Beyond action execution and observation, other studies provide evidence for the disruption of the functional organization of the primary motor cortex in autism, potentially reflecting a delay in functional specialization of these regions (Nebel et al., 2014). Structurally, there is evidence for increased white matter volume in the motor cortex, and that this volume is related to motor difficulties (Mostofsky et al., 2007). Taken together, these results point to atypical structure and function of motor regions involved in both motor planning and execution.

The Overlap Between the Language and Motor Systems in Autism

While the language and motor systems have largely been studied in siloes, recent behavioral and imaging work has attempted to bridge across these domains, given the interconnected nature of speech and motor skills observed in autism. Indeed, difficulties with oral-motor and manual-motor skills have been shown to relate to later speech and language abilities in autistic individuals (e.g., Bhat, 2023). A large-scale analysis of motor, speech, and language abilities in autistic children ($n=1781$) revealed that fine motor skills were positively correlated with expressive and receptive language, and gross motor skills were positively correlated with

receptive language (Mody et al., 2017). Taken together, this emerging literature supports a connection between motor skills and communication in autistic individuals, and indeed, there have been calls for an interdisciplinary investigation of these domains (e.g., Mody & McDougle, 2019).

Beyond behavior, there have been a few studies that investigate the brain bases of the motor act of speech production; these studies show brain-based differences between autistic and nonautistic individuals during basic and complex oromotor and speech production tasks. A magnetoencephalography study of the neural dynamics during phonemic production tasks revealed significant differences in magnitude and latency of activation in several motor regions between autistic and nonautistic children (Pang et al., 2016). A diffusor tensor imaging study in autistic adults revealed a weaker connection between the ventral premotor cortex and the supplementary motor area (SMA) compared to neurotypical participants (Peeva et al., 2013).

The Current Dissertation

The present dissertation aims to weave together threads from the speech, language, and motor literatures to explore the role of the motor system in speech, language, and communication challenges in autism. The link between the speech motor and language systems remains a broad and open question in the field, and the current work represents an early step toward integrating these domains to build a more cohesive understanding. The research questions, aims, and hypotheses related to each chapter are outlined in **Table 1.1**.

Table 1.1

Research questions, aims and hypotheses for each study comprising the dissertation

Chapter 2	<i>Title:</i> Atypical Motor System Engagement in Autistic Children During Nonword Repetition <i>Research Question:</i> What are the brain bases of phonological working memory impairments in autism?
------------------	--

	<i>Study Aims:</i> (1) Identify functional brain differences during nonword repetition in autistic vs. neurotypical children in the three key brain networks involved in nonword repetition: speech perception, working memory, and speech production. (2) Investigate the specificity of these differences to autism by comparing neural responses to the same task in children with a reading disability. <i>Study Hypotheses:</i> (1) Multivariate pattern analysis (decoding accuracy) and univariate analysis will reveal the brain regions that are most consistent or distinctive in response to more challenging phonological tasks in the autism group, and whether specific brain regions are being recruited to compensate during challenging phonological processing. (2) These differences will be specific to autism (i.e., not appearing in the reading disability group), and will be related to performance on language tasks and communication measures.
Chapter 3	<i>Title:</i> Acoustic and Self-Reported Prosodic Differences in Autistic Adults <i>Research Question:</i> How do autistic adults experience prosody production and perception, what global prosodic features distinguish between autistic and nonautistic adults in a spontaneous speech task, and how do these subjective and objective features relate to autism traits and diagnosis? <i>Study Aims:</i> (1) Characterize the self-reported experiences of prosody in autistic and nonautistic adults. (2) Quantify and compare acoustic features (pitch, rate, pauses) of speech during a spontaneous picture description task between autistic and nonautistic adults. (3) Explore the relationship between self-reported prosody experiences, prosodic acoustic features, autism traits, nonverbal IQ, and language (measured via vocabulary size). <i>Study Hypotheses:</i> (1) Autistic adults will report greater difficulty with prosody perception and production. (2) There will be a combination of acoustic features of global prosody that distinguish between the autistic and nonautistic groups.
Chapter 4	<i>Title:</i> Precision fMRI Reveals Engagement of Speech Motor Regions During Prosody Perception and Differences in Autism

	<i>Research Question:</i> Are speech-production motor brain regions engaged differently during language tasks in autistic compared to neurotypical adults? <i>Study Aims:</i> (1) Investigate how nonautistic individuals engage their individually localized speech production system during a passive prosody perception task. (2) Examine if this engagement differs from that of autistic individuals. (3) Explore whether this engagement is related to behavioral performance on a prosody perception task outside of the scanner. <i>Study Hypotheses:</i> (1) Nonautistic individuals will engage some motor regions more when listening to sentences with a dynamic prosody contour than a flat contour. (2) Autistic individuals will demonstrate reduced activation of these regions. (3) Activation in these regions will be related to performance on a prosody perception task outside of the scanner.
--	---

Overview of Chapters

Atypical Motor System Engagement in Autistic Children During Nonword Repetition

Chapter 2 establishes a rationale for the line of inquiry into the intersection of the motor network with speech and language in autism. In this chapter, engagement of three key language-relevant brain networks – speech perception, working memory, and speech production – are investigated in relation to nonword repetition difficulties in autistic children. Nonword repetition is a commonly used clinical task that strongly correlates with language skills and engages multiple language subprocesses, making it an ideal lens for studying the brain bases of language challenges in autism. Univariate and multivariate analysis approaches are used to understand which of the three networks of interest are related to nonword repetition challenges in autism. This chapter lays the foundation for further investigations into the role of the speech motor system in communication challenges in autism.

Acoustic and Self-Reported Prosodic Differences in Autistic Adults

Chapter 3 examines how motor-system differences may impact speech, language, and communication for autistic adults beyond nonword repetition, with a focus on *prosody*. Prosody—the use of pitch, metrical structure, duration, loudness, and voice quality—plays a key role in conveying emotion, emphasizing key points, and expressing pragmatic intent. While atypical prosody has been noted in autism since its earliest descriptions, prior research has largely overlooked spontaneous prosody (especially in adults), and has given limited attention to timing and pauses. Further, no study to date has examined autistic individuals’ lived experiences of prosody. To address these gaps, this chapter examines prosody in autism through two complementary approaches: (1) a novel subjective self-report measure capturing autistic adults’ experiences with prosody perception and production, and (2) an objective acoustic analysis of prosody during a spontaneous picture description task. This chapter establishes groundwork for further investigating these motor system differences in the brain.

Precision fMRI Reveals Engagement of Speech Motor Regions During Prosody Perception and Differences in Autism

Chapter 4 explores the motor-based brain differences related to prosody challenges in autism. Research in nonautistic individuals, including imaging and lesion studies, has suggested the importance of key motor regions in both the production and perception of linguistic prosody. However, the specific contributions of these regions remain poorly understood in both autistic and nonautistic adults. This chapter presents novel insights from an fMRI study that precisely characterizes the speech production motor regions in autistic and non-autistic adults, and examines their contribution to prosody perception. First, each individual’s speech production motor regions were localized. Then, these individualized regions were used to extract activation during a passive

linguistic prosody task. Overall, this chapter highlights the neural connections between speech production and prosody perception in the brain.

This dissertation reframes and examines language and communication challenges in autism through the lens of speech motor function, offering new insights into how motor differences may contribute to prosody and speech production challenges. By bridging brain and behavior and incorporating autistic individuals' lived experiences, these findings open new avenues for understanding and supporting communication in autism.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (DSM-5®). American Psychiatric Pub.
- American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [Technical report].
- American Speech-Language-Hearing Association. (n.d.). *Dysarthria*. <https://www.asha.org/public/speech/disorders/dysarthria/>
- Andreou, G., Lymperopoulou, V., & Aslanoglou, V. (2022). Developmental Language Disorder (DLD) and Autism Spectrum Disorder (ASD): similarities in pragmatic language abilities. A systematic review. *International Journal of Developmental Disabilities*, 70(5), 777–791. <https://doi.org/10.1080/20473869.2022.2132669>
- Baixauli-Forte, I., Miranda Casas, A., Berenguer-Forner, C., Colomer-Diago, C., & Roselló-Miranda, B. (2019). Pragmatic competence of children with autism spectrum disorder. Impact of theory of mind, verbal working memory, ADHD symptoms, and structural language. *Applied Neuropsychology*. Child, 8(2), 101–112. <https://doi.org/10.1080/21622965.2017.1392861>
- Bal, V. H., Fok, M., Lord, C., Smith, I. M., Mirenda, P., Szatmari, P., Vaillancourt, T., Volden, J., Waddell, C., Zwaigenbaum, L., Bennett, T., Duku, E., Elsabbagh, M., Georgiades, S., Ungar, W. J., & Zaidman-Zait, A. (2020). Predictors of longer-term development of expressive language in two independent longitudinal cohorts of language-delayed preschoolers with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 61(7), 826–835. <https://doi.org/10.1111/jcpp.13117>
- Bedford, R., Pickles, A., & Lord, C. (2016). Early gross motor skills predict the subsequent development of language in children with autism spectrum disorder. *Autism Research*, 9(9), 993–1001. <https://doi.org/10.1002/aur.1587>
- Beukelman, D. & Mirenda, P. (2013). *Augmentative and Alternative Communication: Supporting Children & Adults with Complex Communication Needs 4th Edition*. Baltimore: Paul H. Brookes Publishing.
- Bhat, A., Tulskey, D. S., & Boulton, A. J. (2023). Cross-replicating findings on unique motor impairments of children with ASD using confirmatory factor analysis and a novel SPARK study sample. *Autism Research*, 16(5), 967–980. <https://doi.org/10.1002/aur.2904>
- Bloom, L., & Lahey, M. (1978). *Language Development and Language Disorders*. Wiley.
- Bloom, L. (1980). Language development, language disorders, and learning disabilities: LD3. *Bulletin of the Orton Society*, 30(1), 113–133. <https://doi.org/10.1007/BF02653712>
- Broca, P. (1861). Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. *Bulletins de la Société d'Anthropologie de Paris*, 2, 235–238.

Centers for Disease Control and Prevention. (2024). *Data and statistics on autism spectrum disorder*. <https://www.cdc.gov/autism/data-research/index.html>

Chenauksy, K., Brignell, A., Morgan, A., & Tager-Flusberg, H. (2019). Motor speech impairment predicts expressive language in minimally verbal, but not low verbal, individuals with autism spectrum disorder. *Autism & Developmental Language Impairments*, 4. <https://doi.org/10.1177/2396941519856333>

Chenauksy, K. V., Brignell, A., Morgan, A. T., Norton, A. C., Tager-Flusberg, H. B., Schlaug, G., & Guenther, F. H. (2021). A modeling-guided case study of disordered speech in minimally verbal children with autism spectrum disorder. *American Journal of Speech-Language Pathology*, 30(3S), 1542–1557. https://doi.org/10.1044/2021_AJSLP-20-00121

Chenauksy, K., & Tager-Flusberg, H. (2022). The importance of deep speech phenotyping for neurodevelopmental and genetic disorders: a conceptual review. *Journal of Neurodevelopmental Disorders*, 14(1), 36–14. <https://doi.org/10.1186/s11689-022-09443-z>

Craig, F., Crippa, A., Ruggiero, M., Rizzato, V., Russo, L., Fanizza, I., & Trabacca, A. (2021). Characterization of autism spectrum disorder (ASD) subtypes based on the relationship between motor skills and social communication abilities. *Human Movement Science*, 77, 102802. <https://doi.org/10.1016/j.humov.2021.102802>

Crompton, C. J., Ropar, D., Evans-Williams, C. V., Flynn, E. G., & Fletcher-Watson, S. (2020). Autistic peer-to-peer information transfer is highly effective. *Autism*, 24(7), 1704–1712. <https://doi.org/10.1177/1362361320919286>

Cuneo, N., Floyd, S., & Goldberg, A. E. (2024). Word meaning is complex: Language-related generalization differences in autistic adults. *Cognition*, 244, 105691. <https://doi.org/10.1016/j.cognition.2023.105691>

Dahan, D. (2015). Prosody and language comprehension. *Wiley Interdisciplinary Reviews. Cognitive Science*, 6(5), 441–452. <https://doi.org/10.1002/wcs.1355>

Davison, K.E., Liu, T., Belisle, R., Perrachione, T., Qi, Z., Gabrieli, J.D.E., Tager-Flusberg, H., & Zuk, J. (*In press*). Speech timing in relation to white matter organization in autistic and non-autistic children.

Eyler, L. T., Pierce, K., & Courchesne, E. (2012). A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain: A Journal of Neurology*, 135(Pt 3), 949–960. <https://doi.org/10.1093/brain/awr364>

Floyd, S., Jeppsen, C., & Goldberg, A. E. (2021). Brief Report: Children on the Autism Spectrum are Challenged by Complex Word Meanings. *Journal of Autism and Developmental Disorders*, 51(7), 2543–2549. <https://doi.org/10.1007/s10803-020-04687-x>

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill Goldsmith, H. (2008).

Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(1), 43–50.

Goldman, R. & Fristoe, M. (1986). *Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service.

Guinchat, V., Chamak, B., Bonniau, B., Bodeau, N., Perisse, D., Cohen, D., & Danion, A. (2012). Very early signs of autism reported by parents include many concerns not specific to autism criteria. *Research in Autism Spectrum Disorders*, 6(2), 589–601. <https://doi.org/10.1016/j.rasd.2011.10.005>

Harris, S. R. (2017). Early motor delays as diagnostic clues in autism spectrum disorder. *European Journal of Pediatrics*, 176(9), 1259–1262. <https://doi.org/10.1007/s00431-017-2951-7>

Iuzzini-Seigel, J., Allison, K. M., & Stoeckel, R. (2022). A tool for differential diagnosis of childhood apraxia of speech and dysarthria in children: A tutorial. *Language, Speech, and Hearing Services in Schools*, 53(4), 926–946. https://doi.org/10.1044/2022_LSHSS-21-00164

Jouravlev, O., Kell, A. J. E., Mineroff, Z., Haskins, A. J., Ayyash, D., Kanwisher, N., & Fedorenko, E. (2020). Reduced Language Lateralization in Autism and the Broader Autism Phenotype as Assessed with Robust Individual-Subjects Analyses. *Autism Research*, 13(10), 1746–1761. <https://doi.org/10.1002/aur.2393>

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, 20(4), 442–462. <https://doi.org/10.1177/1362361315588200>

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16(2–3), 287–308. <https://doi.org/10.1080/01690960042000058>

Kleinhans, N. M., Müller, R. A., Cohen, D. N., & Courchesne, E. (2008). Atypical functional lateralization of language in autism spectrum disorders. *Brain Research*, 1221, 115–125. <https://doi.org/10.1016/j.brainres.2008.04.080>

Larson, C., Thomas, H. R., Crutcher, J., Stevens, M. C., & Eigsti, I.-M. (2025). Language Networks in Autism Spectrum Disorder: A systematic review of connectivity-based fMRI studies. *Review Journal of Autism and Developmental Disorders*, 12(1), 110–137. <https://doi.org/10.1007/s40489-023-00382-6>

Levelt W. J. (1999). Models of word production. *Trends in Cognitive Sciences*, 3(6), 223–232. [https://doi.org/10.1016/s1364-6613\(99\)01319-4](https://doi.org/10.1016/s1364-6613(99)01319-4)

Li, M., Wang, Y., Tachibana, M., Rahman, S., & Kagitani-Shimono, K. (2022). Atypical structural connectivity of language networks in autism spectrum disorder: A meta-analysis of diffusion tensor imaging studies. *Autism Research*, 15(9), 1585–1602. <https://doi.org/10.1002/aur.2789>

- Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., Jones, E. J. H., Jones, R. M., Pickles, A., State, M. W., Taylor, J. L., & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1), 5. <https://doi.org/10.1038/s41572-019-0138-4>
- Maffei, M. F., Chenausky, K. V., Gill, S. V., Tager-Flusberg, H., & Green, J. R. (2023). Oromotor skills in autism spectrum disorder: A scoping review. *Autism Research: Official Journal of the International Society for Autism Research*, 16(5), 879–917. <https://doi.org/10.1002/aur.2923>
- Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: a systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review*, 34(1), 73–86. <https://doi.org/10.1016/j.cpr.2013.11.002>
- Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual differences in children’s pragmatic ability: A review of associations with formal language, social cognition, and executive functions. *Language Learning and Development*, 14(3), 186–223. <https://doi.org/10.1080/15475441.2018.1455584>
- Mermelstein P. (1973). Articulatory model for the study of speech production. *The Journal of the Acoustical Society of America*, 53(4), 1070–1082. <https://doi.org/10.1121/1.1913427>
- Milton, D. E. M. (2012). On the ontological status of autism: The ‘double empathy problem’. *Disability & Society*, 27(6), 883–887.
- Milton, D., Gurbuz, E., & Lopez, B. (2022). The 'double empathy problem': Ten years on. *Autism*, 26(8), 1901–1903. <https://doi.org/10.1177/13623613221129123>
- Miranda, A., Berenguer, C., Baixauli, I., & Roselló, B. (2023). Childhood language skills as predictors of social, adaptive and behavior outcomes of adolescents with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 103, 102143. <https://doi.org/10.1016/j.rasd.2023.102143>
- Mody, M., & McDougle, C. J. (2019). Getting the Word “Out”: A Role for the Motor System in Autism Spectrum Disorder. *Perspectives of the ASHA Special Interest Groups*, 4(6), 1221–1228. https://doi.org/10.1044/2019_PERSP-19-00126
- Mody, M., Shui, A. M., Nowinski, L. A., Golas, S. B., Ferrone, C., O'Rourke, J. A., & McDougle, C. J. (2017). Communication Deficits and the Motor System: Exploring Patterns of Associations in Autism Spectrum Disorder (ASD). *Journal of Autism and Developmental Disorders*, 47(1), 155–162. <https://doi.org/10.1007/s10803-016-2934-y>
- Morrison, K. E., DeBrabander, K. M., Jones, D. R., Faso, D. J., Ackerman, R. A., & Sasson, N. J. (2020). Outcomes of real-world social interaction for autistic adults paired with autistic compared to typically developing partners. *Autism*, 24(5), 1067–1080. <https://doi.org/10.1177/1362361319892701>

- Mosner, M. G., Kinard, J. L., Shah, J. S., McWeeny, S., Greene, R. K., Lowery, S. C., Mazefsky, C. A., & Dichter, G. S. (2019). Rates of co-occurring psychiatric disorders in autism spectrum disorder using the mini international neuropsychiatric interview. *Journal of Autism and Developmental Disorders*, 49(9), 3819–3832. <https://doi.org/10.1007/s10803-019-04090-1>
- Mostofsky, S. H., Dubey, P., Jerath, V. K., Jansiewicz, E. M., Goldberg, M. C., & Denckla, M. B. (2006). Developmental dyspraxia is not limited to imitation in children with autism spectrum disorders. *Journal of the International Neuropsychological Society*, 12(3), 314–326. <https://doi.org/10.1017/s1355617706060437>
- Mostofsky, S. H., Burgess, M. P., & Gidley Larson, J. C. (2007). Increased motor cortex white matter volume predicts motor impairment in autism. *Brain: A Journal of Neurology*, 130(Pt 8), 2117–2122. <https://doi.org/10.1093/brain/awm129>
- Nebel, M. B., Joel, S. E., Muschelli, J., Barber, A. D., Caffo, B. S., Pekar, J. J., & Mostofsky, S. H. (2014). Disruption of functional organization within the primary motor cortex in children with autism. *Human Brain Mapping*, 35(2), 567–580. <https://doi.org/10.1002/hbm.22188>
- Pang, E. W., Valica, T., MacDonald, M. J., Taylor, M. J., Brian, J., Lerch, J. P., & Anagnostou, E. (2016). Abnormal brain dynamics underlie speech production in children with autism spectrum disorder. *Autism Research*, 9(2), 249–261. <https://doi.org/10.1002/aur.1526>
- Posar, A., & Visconti, P. (2022). Early motor signs in autism spectrum disorder. *Children* (Basel, Switzerland), 9(2), 294. <https://doi.org/10.3390/children9020294>
- Peeva, M. G., Tourville, J. A., Agam, Y., Holland, B., Manoach, D. S., & Guenther, F. H. (2013). White matter impairment in the speech network of individuals with autism spectrum disorder. *NeuroImage: Clinical*, 3, 234–241. <https://doi.org/10.1016/j.nicl.2013.08.011>
- Reindal, L., Nærlund, T., Weidle, B., Lydersen, S., Andreassen, O. A., & Sund, A. M. (2023). Structural and pragmatic language impairments in children evaluated for autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, 53(2), 701–719. <https://doi.org/10.1007/s10803-020-04853-1>
- Richards, M., Mossey, J., & Robins, D. L. (2016). Parents' concerns as they relate to their child's development and later diagnosis of autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics: JDBP*, 37(7), 532–540. <https://doi.org/10.1097/DBP.0000000000000339>
- Shriberg, L. D., Paul, R., Black, L. M., & van Santen, J. P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 41(4), 405–426. <https://doi.org/10.1007/s10803-010-1117-5>
- Stillesjö, S., Hjærtström, H., Johansson, A., Rudolfsson, T., Säfsström, D., & Domellöf, E. (2025). Action execution and observation in autistic adults: A systematic review of fMRI studies. *Autism Research*, 18(2), 238–260. <https://doi.org/10.1002/aur.3291>

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981. <https://doi.org/10.1080/01690960903498424>

West K. L. (2019). Infant Motor Development in Autism Spectrum Disorder: A Synthesis and Meta-analysis. *Child Development*, 90(6), 2053–2070. <https://doi.org/10.1111/cdev.13086>

Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., & Naigles, L. R. (2017). Grammatical language impairment in autism spectrum disorder: Exploring language phenotypes beyond standardized testing. *Frontiers in Psychology*, 8, 532. <https://doi.org/10.3389/fpsyg.2017.00532>

World Health Organization. *International Classification of Diseases (ICD)*. <https://www.who.int/standards/classifications/classification-of-diseases>

Ying Sng, C., Carter, M., & Stephenson, J. (2018). A systematic review of the comparative pragmatic differences in conversational skills of individuals with autism. *Autism & Developmental Language Impairments*, 3, <https://doi.org/10.1177/2396941518803806>

Zampella, C. J., Wang, L. A. L., Haley, M., Hutchinson, A. G., & de Marchena, A. (2021). Motor skill differences in autism spectrum disorder: A clinically focused review. *Current Psychiatry Reports*, 23(10), 64. <https://doi.org/10.1007/s11920-021-01280-6>

Chapter 2:

Altered Engagement of the Speech Motor Network is Associated with Reduced Phonological Working Memory in Autism

This chapter was previously published in *Neuroimage: Clinical*. It contains the following changes from the original article: Figures have been renumbered to be consistent with the rest of the dissertation, formatting of numbers and references have been revised to be consistent with APA style and the remainder of the dissertation.

O'Brien, A. M., Perrachione, T. K., Wisman Weil, L., Sanchez Araujo, Y., Halverson, K., Harris, A., Ostrovskaya, I., Kjelgaard, M., Wexler, K., Tager-Flusberg, H., Gabrieli, J. D. E., & Qi, Z. (2023). Altered engagement of the speech motor network is associated with reduced phonological working memory in autism. *NeuroImage: Clinical*, 37, 103299. <https://doi.org/10.1016/j.nicl.2022.103299>

**Altered engagement of the speech motor network is associated with reduced phonological
working memory in autism**

Amanda M. O'Brien^{1,2}, Tyler K. Perrachione³, Lisa Wisman Weil⁴, Yoel Sanchez Araujo⁵, Kelly
Halverson⁶, Adrienne Harris⁷, Irina Ostrovskaya⁸, Margaret Kjelgaard⁹, Kenneth Wexler¹⁰,
Helen Tager-Flusberg¹¹, John D.E. Gabrieli², Zhenghan Qi¹²

¹Program in Speech and Hearing, Bioscience and Technology, Harvard University

²McGovern Institute for Brain Research, Massachusetts Institute of Technology

³Department of Speech, Language, and Hearing Sciences, Boston University

⁴Department of Communication Sciences and Disorders, Emerson College

⁵Princeton Neuroscience Institute, Princeton University

⁶Department of Clinical Psychology, University of Houston

⁷The Carolina Institute for Developmental Disabilities, University of North Carolina School of
Medicine

⁸Haverhill Public Schools, Massachusetts

⁹Department of Communication Sciences and Disorders, Bridgewater State University

¹⁰Department of Brain and Cognitive Sciences & Department of Linguistics and Philosophy,
Massachusetts Institute of Technology

¹¹Department of Psychological & Brain Sciences, Boston University

¹²Department of Communication Sciences and Disorders & Department of Psychology,
Northeastern University

Abstract

Nonword repetition, a common clinical measure of phonological working memory, involves component processes of speech perception, working memory, and speech production. Autistic children often show behavioral challenges in nonword repetition, as do many individuals with communication disorders. It is unknown which subprocesses of phonological working memory are vulnerable in autistic individuals, and whether the same brain processes underlie the transdiagnostic difficulty with nonword repetition. We used functional magnetic resonance imaging (fMRI) to investigate the brain bases for nonword repetition challenges in autism. We compared activation during nonword repetition in functional brain networks subserving speech perception, working memory, and speech production between neurotypical and autistic children. Autistic children performed worse than neurotypical children on nonword repetition and had reduced activation in response to increasing phonological working memory load in the supplementary motor area. Multivoxel pattern analysis within the speech production network classified shorter vs. longer nonword-repetition trials less accurately for autistic than neurotypical children. These speech production motor-specific differences were not observed in a group of children with reading disability who had similarly reduced nonword repetition behavior. These findings suggest that atypical function in speech production brain regions may contribute to nonword repetition difficulties in autism.

Keywords: autism spectrum disorder, phonological working memory, nonword repetition, motor cortex, speech production

Introduction

During language development, children have a remarkable ability to rapidly learn to recognize and produce a multitude of new words (Bloom, 1973). Transforming *language heard* into *language said* depends on a sophisticated orchestration among the cognitive and neural systems that support speech perception, working memory, and speech production. The ability to briefly hold speech information in mind during this transformation is known as *phonological working memory* and is thought to play a key role in language development (Baddeley, Gathercole, & Papagno, 1998). In turn, this faculty is thought to be disrupted in many developmental communication disorders (e.g., William et al., 2012; Archibald, 2017; Gray et al., 2019).

Clinicians and researchers frequently use *nonword repetition* as a measure of phonological working memory, which has high sensitivity for developmental language disorders (Archibald & Joanisse, 2009; Estes et al., 2007; Ellis Weismer et al., 2000) and reading disorders (Ehrhorn et al., 2021; Melby-Lervag & Lervag, 2012). Nonword repetition is also atypical in many autistic³ children (Whitehouse et al., 2008; Gabig, 2008; Williams et al., 2013; Nadig & Mulligan, 2017). It is currently unknown why children with autism spectrum disorder (ASD) have difficulty with nonword repetition, particularly given the heterogenous language profiles of autistic individuals (Kjelgaard & Tager-Flusberg, 2001). While nonword repetition performance is classically operationalized as an assessment of phonological working memory, this multifaceted behavior involves several subprocesses related to speech perception (e.g., encoding), short-term working memory (e.g., storage), and speech production (e.g., planning and articulation) (Hickock, 2009; Majerus, 2013). During the task of nonword repetition, it is likely that a distributed set of brain

³ Many autistic people prefer identity-first language (“autistic person”) to person-first language (“person with autism”), while some members within the community prefer person-first language (Lei et al., 2021). These preferences guide our semantic language choices; we use identity-first and person-first language interchangeably throughout the manuscript.

networks associated with each of these cognitive processes is engaged (Acheson et al., 2011; Fiez, 2016; Indefrey & Levelt, 2004; Scott & Perrachione, 2019). As such, each of these process-related networks may operate differently in autistic children, relating to different possible sources for the nonword repetition difficulties in this population. Here, we aimed to identify how functional response in each of these three brain networks differ in autism during nonword repetition, thereby shedding light on the neural bases of phonological working memory challenges in autistic individuals.

Phonological working memory is related to a wide range of language-related skills including vocabulary acquisition, sentence and discourse processing, and reading development (Adams & Gathercole, 1996; Baddeley et al., 1998; Gathercole et al., 1997; Gathercole, 2006). Several behavioral studies have found reduced phonological working memory in autistic individuals regardless of articulation skills (Gabig, 2008; Habib et al., 2019; Macizo et al., 2016; Williams et al., 2006). Further, reduced phonological working memory has been identified in autistic children who have lower scores on assessments of vocabulary, and higher-order syntax and semantics (Kjelgaard & Tager-Flusberg, 2001). Phonological working memory challenges are also seen in other populations with language difficulties, including dyslexia, dysfluency (stuttering), Down syndrome, and developmental language disorders (Estes et al., 2007; Peter et al., 2011; Melby-Lervag & Lervag, 2012; Ehrhorn et al., 2021). However, the underlying source of these difficulties may differ across diagnoses. For example, the types of errors made by autistic children during nonword repetition tend to differ from those made by children with specific language impairment (or developmental language disorders) (Nadig & Mulligan, 2017; Williams et al., 2013). These studies suggest that, while reduced phonological working memory in autistic

children may have widespread consequences in language development, the source of their difficulties may differ from those of children with other developmental communication disorders.

Despite the importance of phonological working memory in language development, and the clinical value of its measurement via nonword repetition (Archibald & Gathercole, 2006; Dollaghan & Campbell, 1998), little is known about how phonological working memory is disrupted in autistic children, or why their difficulties with nonword repetition may manifest differently from those in other developmental communication disorders. Disentangling the sources of children's nonword repetition difficulties via behavioral assays has proved challenging, as behavior alone cannot distinguish between difficulties perceiving, remembering, or reproducing the target speech. Neurally, however, the brain networks that support speech perception, working memory, and speech production are substantially distinct (Hickok & Poeppel, 2007), such that localized differences in brain activation may help inform which specific processes are disrupted. In this study, we investigated activation within each of these three brain networks to improve our understanding of their relationship to nonword repetition difficulties in autism.

Our primary aim in this research was to identify functional brain differences during nonword repetition in autistic versus neurotypical (NT) children. Previous research has documented sensory, cognitive, and motor differences in autism, such that each of the three primary sub-components of nonword repetition may be the source of impaired phonological working memory: Correspondingly, **(1)** activation differences in the speech perception network during nonword repetition would be consistent with evidence for perception-based differences in autism (e.g., Frith, 1989; Mottron & Burack, 2001; Pellicano & Burr, 2012); **(2)** activation differences in the working memory network during nonword repetition would be in-line with evidence for executive functioning differences in autism (e.g., Ozonoff et al., 1991; Joseph and Tager-Flusberg, 2004;

Pellicano, 2012; Leung et al., 2016); and (3) activation differences in the speech production network during nonword repetition would be consistent with evidence for motor-based differences in autism (e.g., Whyatt & Craig, 2013; Bhat et al., 2011). Upon identifying brain regions of difference between the autistic and NT groups, we then tested the specificity of any difference in a group of non-autistic group of children with reading disability who also exhibited impaired nonword repetition.

Neural systems supporting nonword repetition

Encoding: The speech perception network

Speech perception involves encoding acoustic speech signals and mapping them onto abstract linguistic representations (Samuel, 2011; Poeppel, 2015). Speech perception tasks engage a network of widely distributed brain regions, including primary and association auditory areas (e.g., bilateral superior temporal gyrus (STG) and superior temporal sulcus (STS)) involved in spectrotemporal and phonological processing, respectively. Sensorimotor transformations are believed to occur via a posterior/dorsal auditory stream, encompassing planum temporale and parietal operculum, with articulatory representations and actions encoded in left inferior frontal gyrus and ventral (pre)motor cortex, and in parietal operculum (Hickok & Poeppel, 2007).

Various sensory-based hypotheses of autism have posited that autistic individuals may make enhanced use of low-level sensory features at the expense of higher-level gestalt representations (Mottron & Burack, 2001; Happe & Frith, 2006). Indeed, some behavioral evidence has borne out this hypothesis in the domain of speech and language. Some autistic individuals pay greater attention to lower-level acoustic cues at the cost of higher-level linguistic information (Järvinen-Pasley et al., 2008; Soulières et al., 2007). Further, a functional magnetic resonance imaging (fMRI) meta-analysis study of the brain bases of speech perception revealed

that activation in temporal-lobe language areas (e.g., STG), was common to both autistic and NT groups, but activation in frontal language areas (e.g., superior frontal gyrus, left medial frontal gyrus, and right inferior frontal gyrus) was found exclusively in the NT group (Tryfon et al., 2018).

Storage: The working memory network

Working memory refers to the temporary storage and manipulation of mental representations of information (e.g., speech sounds and sequences) on the timescale of seconds. In their influential phonological loop theory, Baddeley and Hitch (1974) proposed that short-term storage of verbal information occurs through interactions between a phonological buffer and subvocal rehearsal process (Baddeley, 1992; Baddeley & Hitch, 1974; Baddeley, 2003; Hitch & Baddeley, 1976). However, this psychological model does not posit specific brain regions associated with it. Classical fMRI and PET studies using working memory tasks from cognitive psychology (e.g., digit span, n-back) identified a network of frontal and parietal regions as important for working memory (e.g., Bushbaum & D’Esposito, 2019). However, studies that have used tasks focused on immediate repetition of nonword tend to find a network comprising mostly speech perception and speech production regions, such as STG, IFG, supplementary motor area (SMA), posterior middle frontal gyrus, and ventral motor cortex (Scott & Perrachione, 2019; Perrachione et al., 2017; Strand et al., 2008). To this end, it remains an open question whether and how nonword repetition (and therefore phonological working memory) requires the participation of domain-general (or multiple demand) working memory areas such as dorsolateral prefrontal cortex and parietal lobe (Jacquemot & Scott, 2006; Majerus, 2013; Hickok, 2009; Fiez, 2016.) Furthermore, it is unknown whether nonword repetition difficulties in developmental communication disorders like ASD place differential demands on speech vs. domain-general (multiple demand) working memory networks for performing this task.

There is evidence to suggest that reduced executive functioning – including working memory – is associated with many challenges observed in autistic individuals (e.g., Pellicano, 2012; Leung et al., 2016). Meta-cognitive skills such as working memory and planning predicted social skills in autistic but not NT children (Leung et al., 2016). Behavioral research investigating spatial, verbal, and nonverbal working memory in autistic individuals has shown functional challenges, including reduced accuracy and slower performance (for a meta-analysis, see: Wang et al., 2017; Rabiee et al., 2017; Sachse et al., 2013; Cui et al., 2010). Conflicting research, however, has shown intact verbal working memory as measured by *n*-back letter tasks (Williams et al., 2005) and intact visual working memory as measured by a visual change detection task (Lynn et al., 2022). Multiple studies have reported reduced *n*-back performance and altered patterns of brain activation in autistic children and adults (Vogan et al., 2018; Yeung et al., 2019; Koshino et al., 2005; Barendse et al., 2013).

Planning and articulation: The speech production network

Speech production involves the planning and execution of articulatory movements, as well as the continuous monitoring of the auditory and somatosensory experience of one's own speech. The computational and neuroanatomical Directions into Velocities of Articulators (DIVA) model provides a unified theory of speech production and implicates a network of cortical and subcortical regions, including motor regions (primary motor cortex, supplementary motor area (SMA), basal ganglia) and sensory regions (e.g., STG) (Tourville & Guenther, 2011). Neuroimaging studies have shown bilateral activation of motor cortex, premotor cortex, somatosensory cortex, and SMA to be most reliably associated with overt articulation, with a rough somatotopic organization of articulators in the motor cortex (Conant et al., 2013; Long et al., 2016; Guenther, 2016). The

primary sensory regions play a key role in the auditory monitoring of one's speech and articulation by mapping auditory targets, auditory states, and errors.

Despite the diagnostic emphasis on social and communication skills in autism, motor impairments have been increasingly identified as an underdiagnosed, yet clinically meaningful, feature of autism (Craig et al., 2021; Whyatt & Craig, 2013; Zampella et al., 2021). Motor delays and atypical movement development (e.g., stereotyped movement) have been identified as some of the earliest signs of autism, which often emerge even before the classical social communication differences (West, 2018; Posar & Visconti, 2022). Speech and language (expressive and receptive) abilities in autistic individuals have been shown to be correlated with oromotor, visuomotor, manual motor, fine motor, and gross motor skills (Gernsbacher et al., 2008; Bhat et al., 2022; Mody et al., 2017; Bedford et al., 2016). Behaviorally, studies of the motor speech profiles of autistic individuals have identified atypical speech production execution, including difficulty with complex syllable production tasks (Adams, 1998) and imprecise articulation (Wynn et al., 2021). Despite the mounting evidence suggesting motor differences in autism, including fine motor control related to speech production, the research that investigates the behavioral and brain bases of speech production in autism remains highly limited. A magnetoencephalography (MEG) study examining neural dynamics during phonemic production tasks revealed significant differences between ASD and NT in magnitude and latency of activation in several brain regions including primary motor cortex, motor planning areas (e.g., SMA), sensorimotor integration areas (e.g., orbitofrontal cortex), and the dorsolateral prefrontal cortex (Pang et al., 2016). A diffusor tensor imaging study in adults with autism revealed a weaker connection between the ventral premotor cortex and the SMA (Peeva et al., 2013). These outcomes suggest brain and behavioral differences between autistic and NT groups during basic and complex oro-motor and phonemic tasks.

Specificity of phonological working memory difficulties in autism

An important goal in studying the brain bases of nonword repetition is to better understand how the functional substrates of phonological working memory impairments are shared versus distinct across autism and other neurodevelopmental disorders (cf. Lu et al., 2016). For example, individuals with reading disabilities often demonstrate behavioral difficulties with nonword repetition (Baird et al., 2011), as well as its associated processes (e.g., speech perception: Mannis et al., 1997; Joanisse et al., 2000; Boets et al., 2011; working memory: Smith-Spark et al., 2005; Jeffries & Everatt, 2004; speech production: Catts, 1989; Lambrecht Smith et al., 2010; Carroll et al., 2014). It is possible that any brain differences observed during nonword repetition in autism are not specific to autism but rather underlie nonword repetition and/or communication challenges more broadly. A previous neuroanatomical study found that shared atypicality in the white-matter microstructural connectivity was related to the degree of phonological working memory impairment for both children with reading disabilities (RD) and those with autism (Lu et al., 2016). Testing whether non-autistic populations with parallel challenges in nonword repetition show similar differences in brain functions as autistic children will shed light on the debate about the relationship between language or reading impairment and autism.

Current study

The primary aim of this study was to identify the brain bases of phonological working memory impairments in autism. Children with and without autism performed a nonword repetition task in the scanner while we measured their neural activation using fMRI. With complementary univariate and multivariate pattern analyses, we examined whether the atypical phonological working memory abilities seen in autism are associated with functional differences in each of the three key brain networks thought to be involved in nonword repetition: speech perception, working memory,

and speech production. We also used targeted analyses to examine whether children with reading disability also demonstrate the same brain differences identified in the comparison between autistic and NT children.

Materials and Methods

Participants

Children (age 5 to 18 years) with and without autism were recruited from the greater Boston area of the United States for this study. From an initial sample of $n = 216$ participants, a total of $n = 69$ children met criteria for inclusion in the neurotypical (NT, $n = 28$) or autism ($n = 23$) groups. We also included a control group of participants with a reading disability (RD, $n = 16$) for targeted secondary analyses. Inclusion criteria for all participants required them to be native speakers of American English, right-handed, and born after 32 weeks gestational age. All included participants had parent-reported normal hearing and had nonverbal cognitive ability within normal limits, as measured by nonverbal IQ scores of ≥ 80 on the *Kaufman Brief Intelligence Test* (KBIT; Kaufman & Kaufman, 2004). Included participants had no history of head injury, comorbid psychiatric or neurological conditions, or any genetic disorders associated with autism (e.g., Fragile X syndrome). Written informed consent was obtained from a guardian of all participants. Children provided informed written assent and received monetary compensation for their time. All procedures were approved and overseen by the Committee on the Use of Humans as Experimental Subjects (COUHES) at MIT and conducted in accordance with the Declaration of Helsinki.

Behavioral assessments

Participants completed a brief battery of standardized clinical measures of cognitive, language, and reading ability, including the *Kaufman Brief Intelligence Test* (KBIT) Expressive Vocabulary, Definitions, and Matrix subtests (Kaufman & Kaufman, 2004); the *Clinical*

Evaluation of Language Fundamentals (CELF-4) core language subtests (Semel, Wiig, & Secord, 2003); the *Comprehensive Test of Phonological Processing* (CTOPP) Elision, Blending Words, and Nonword Repetition subtests (Bruno & Walker, 1999); the *Children's Test of Nonword Repetition* (CNRep) (Gathercole et al., 1994); the *Woodcock Reading Mastery Test – Revised Normative Update* (WRMT) Word Identification (WI) and Word Attack (WA) subtests (Woodcock, 2011); and the *Test of Word Reading Efficiency* (TOWRE-2) Sight Word Efficiency and Phonemic Decoding subtests (Tarar et al., 2015; Torgesen et al., 2012). The CTOPP and CNRep are tests of phonological working memory and are typically used to identify children at risk for developmental language disorders (Gathercole, 1995; Archibald & Joanisse, 2009).

Diagnostic confirmation

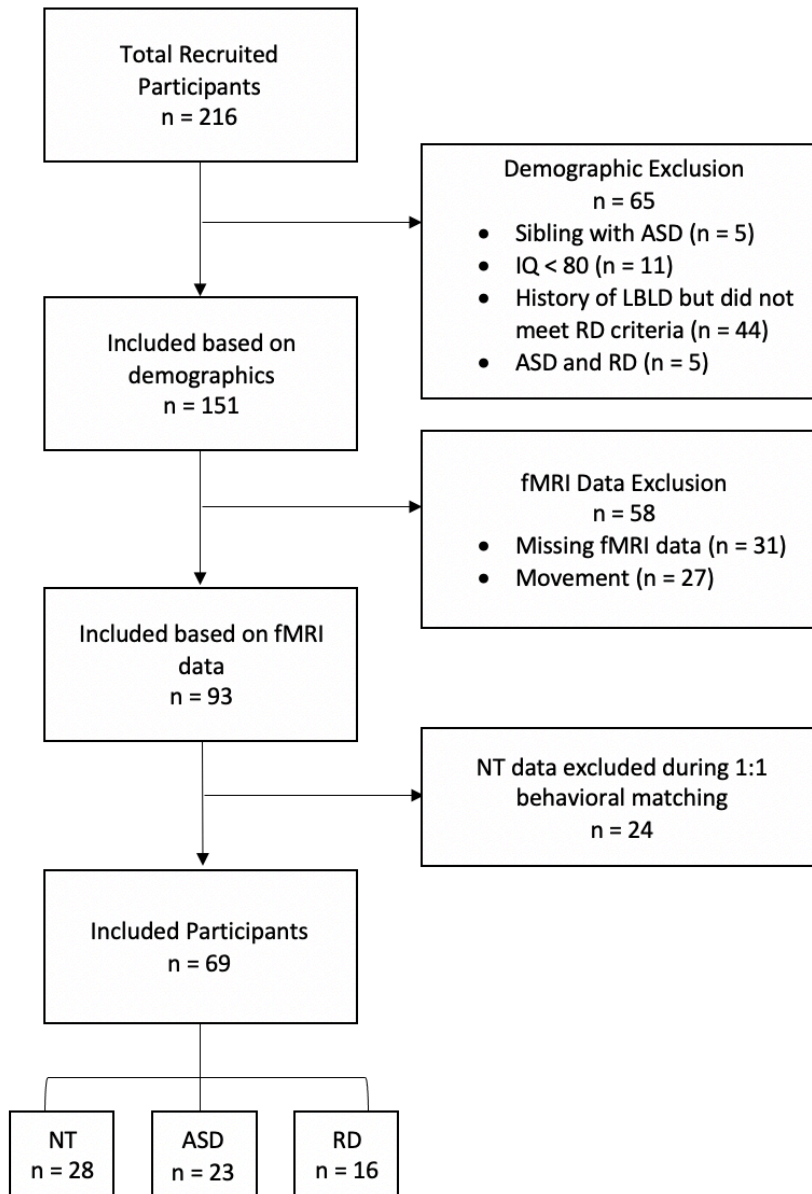
Participants were included in the autism group if they had a clinical diagnosis of autism, which was also confirmed by trained research staff using an *Autism Diagnostic Observation Schedule* (ADOS/ADOS-2) Module 3 or 4 (Lord et al., 2000; 2012). A score of 7 or higher was deemed consistent with an autism diagnosis. To quantify the degree of the autism symptomatology, we converted participants' ADOS scores to autism calibrated severity scores (CSS) ranging from 1-10 (Gotham et al. 2006; Hus & Lord, 2014). Participants in the NT group scored within normal limits on the ADOS-2 and had no first-degree relatives with autism or reading disabilities. Participants in the RD group had standard scores below 90 (below the 25th percentile) on at least two of the four subtests of timed or untimed word or nonword reading from the WRMT and TOWRE. Participants in the RD group also scored within normal limits on the ADOS (**Table 2.1**). See **Appendix A** for more information on ADOS diagnostic confirmation and ADOS scores.

Included matched sample

After screening for demographic inclusion criteria (see 2.1), data quality (see MRI data pre-processing, below), and matching for demographic characteristics, 69 children were included in this study. See **Figure 1.1** for a breakdown of reasons for exclusion. All matching was done using the package in R (Hansen et al., 2006), and a summary of the demographic, cognitive, language, and reading characteristics of all participants is shown in **Table 1.1**. **Appendix A** provides a breakdown of the racial and ethnic makeup of the participants. In matching all participants, we used the pair matching function of the *optmatch* tool in R. The pair matching function works by minimizing the mean paired Euclidean distance between designated variables (i.e., scores). To verify our matching, we used the *MatchIt* package (Ho et al., 2011) in R to calculate propensity scores for our included groups. All group comparisons had variance ratios within an acceptable range (0.5 - 2), as established by the literature (Kover et al., 2013). The language-matched sample had a variance ratio of 1.56 for the variables of age, sex, IQ, and CELF core language score. The NT vs ASD groups had a variance ratio of 0.83 for the variables of age, sex, and IQ. The NT vs RD groups had a variance ratio of 1.71 for the variables of age, sex, and IQ. We also conducted supplementary analyses with an additional control group of autistic and neurotypical children matched on the Core Language Score of the CELF. See **Appendix A** for additional information on the language-matched groups. Hypotheses related to our primary aim (neurofunctional differences during nonword repetition) were tested using groups of autistic (n = 23) and NT (n = 28) children who were matched by age, sex, and non-verbal IQ. Hypotheses related to our secondary aim (altered neural responses' specificity to autism) were tested by comparing the two groups above to a third group of children with reading disability (RD; n = 16) and by comparing language-matched autistic vs. NT subgroups.

Figure 2.1

Flowchart depicting exclusionary criteria and numbers of included participants in each group



Note. ASD = autism spectrum disorder; IQ = intelligence quotient; LBLD = language-based learning disorder; RD = reading disability; fMRI = functional magnetic resonance imaging; NT = neurotypical

Table 2.1.

Phenotypic cognitive, behavioral, language, and reading characteristics in the primary analysis group (NT and ASD), as well as for the RD groups

		NT n = 28 (8 F)		ASD n = 23 (5F)		ASD vs NT	RD n = 16 (7F)		RD vs NT	RD vs ASD
		Mean	SD	Mean	SD	p-value	Mean	SD	p-value	p-value
Demographics	Age	11.9	3.2	12.2	3.4	.389	11.5	3.1	.345	.266
IQ	KBIT	112.8	14.2	110.6	16.2	.302	104.3	12.3	.022	.088
Language	CELF (CLS)	116.2	9.8	103.2	19.2	.004	97.3	17.9	.001	.166
NWR	In-Scanner	87.7	8.8	75.0	22.7	<.001	73.8	21.2	.000	.262
	CTOPP	10.0	2.7	7.7	2.1	.001	8.4	2.7	.035	.199
	CNRep (Raw)	34.7	3.2	30.1	6.4	.002	30	6.2	.005	.475
Reading	TOWRE					.001			<.001	<.001
	(SWE)	108.3	9.9	97.5	10.5		83.1	7.2		
	TOWRE					.012			<.001	<.001
	(PDE)	110.3	13.3	101.7	12.4		83.7	5.0		
	WRMT (WI)	113.0	9.8	102.7	15.0	.004	88.1	9.7	<.001	.001
	WRMT (WA)	108.7	12.0	101.9	14.1	.038	91.1	6.9	<.001	.002
	WJ	114.3	11.6	104.6	17.2	.015	83.7	8.7	<.001	.001
Autism	ADOS CSS	1.19	1.28	10.18	4.48	<.001	2.08	2.64	.05	.001

Note. All *p*-values in this table are from one-tailed *t*-tests (ASD < NT, RD < NT, and RD < ASD for age and all assessments, except for ADOS where ASD > NT and ASD > RD). NT= neurotypical; ASD = autism spectrum disorder; F = female; RD = reading disability; CELF (CLS)= Core Language Score (Clinical Evaluation of Language Fundamentals, Fourth edition); NWR = nonword repetition; CNRep = Children's Test of Nonword Repetition; CTOPP = Comprehensive Test of Phonological Processing 2; KBIT = Kaufman Brief Intelligence Test 2 (Matrix subtest); WRMT = Woodcock Reading Mastery Tests; WI = Word Identification; WA = Word Attack; TOWRE = Test of Word Reading Efficiency; SWE = Sight Word Efficiency; PDE = Phonemic Decoding Efficiency. WJ = Woodcock-Johnson III Tests of Achievement (Sentence Reading Fluency subtest); ADOS = Autism Diagnostic Observation Schedule; CSS = Calibrated Severity Score (averaged across modules 3 and 4).

Stimuli

We used the nonword (i.e., pseudoword) stimuli that were previously generated and described (Perrachione et al., 2017). There were 96 nonwords total, balanced across syllable lengths (2, 3, 4, and 5 syllables), with 24 nonwords per syllable-length condition. The positional phoneme and biphone phonotactic probabilities of the stimuli did not vary significantly across syllable length (both $F(1, 133) < 0.23$, $p > .61$). The nonwords were generated to closely parallel the structural and statistical properties of real English words. A full list of the nonword stimuli can be found in **Appendix A**.

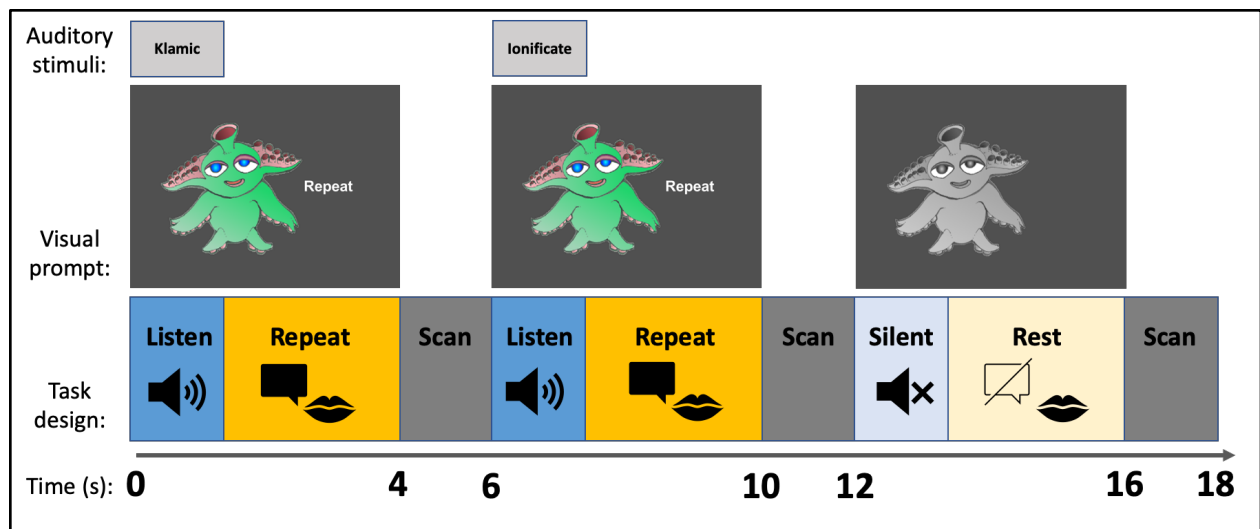
Recordings were made from the speech of a female native speaker of standard American English. Before making the audio recordings of the nonword stimuli, the speaker was extensively familiarized with the nonwords to ensure natural, correct pronunciation. The speaker read the stimuli in citation format (individually and in isolation) and the speech was digitally recorded using a SM58 microphone (Shure Inc., Niles, IL) and Edirol UA-25EX sound card (Roland Corp., Los Angeles, CA), sampling at 44.1 kHz (as described by Perrachione et al., 2017). Each token was normalized for root-mean-square amplitude to 70 dB using Praat (Boersma, 2001).

Experimental session

This experiment was part of a larger project that included other fMRI and MRI acquisition protocols. Participants first engaged in a practice scan to become familiarized with the equipment and behavioral tasks (see **Appendix A** for more information). Children completed the nonword repetition task during a sparse-sampling fMRI paradigm, which allowed the auditory stimuli to be presented, and verbal responses to be uttered, during silent periods in between functional volume acquisitions. Children were instructed to listen to the “alien language” and repeat the words. An image of a colorful alien was used as a visual prompt for the child to repeat the word during the repeating trials, and a gray alien image signaled the child to remain silent during the resting trials (**Figure 2.2**). The participants did not hear or say anything during the rest trials.

Figure 2.2.

Task Design and Accompanying Visual Prompt



Auditory stimuli were presented binaurally using Sensimetrics S-14 earphones (Sensimetrics Corporation, Woburn, MA) at a comfortable listening level. Stimulus presentation was controlled by the software PsychoPy (Peirce et al., 2019). The task was conducted in three runs. Each run included 32 trials, including 8 trials of each syllable load (two, three, four, five syllables) and 8 trials of rest. Within each run, the trials were presented in pairs of each condition (e.g., the participant completed two trials of four syllables, then two trials of two syllables, then two trials of rest). These condition-pairs were pseudo-randomized in each run, such that all participants were presented with the stimuli and rest conditions in the same order, by run. Each trial lasted 6 s and consisted of 4 s of silence, during which the auditory stimulus was presented, and overt verbal repetition recorded, followed by a 2 s fMRI volume acquisition (**Figure 2.2**). Rest trials also lasted 6 s, during which the participant saw the grayed alien image and remained silent.

Neuroimaging data collection and analysis

Scanning parameters

Data were acquired on a Siemens Trio 3T scanner (Siemens AG, Berlin and Munich, Germany) with a 32-channel phased array head coil. A whole-head, high-resolution T1-weighted, multi-echo magnetization-prepared rapid gradient-echo (ME-MPRAGE) anatomical volume was

obtained (acquisition parameters: repetition time [TR] = 2530 mms, echo time [TE] = {1.64, 3.44, 5.24, 7.04 ms}, flip angle = 7.0°, inversion time [TI] = 1400 ms, voxel resolution = 1.0 mm³, field of view [FOV] = 220, 220 mm, 176 sagittal slices).

Three functional runs containing 108 volumes each were collected using sparse-sampled T2*-weighted gradient-echo planar imaging (EPI) scans (acquisition parameters: TR = 6000 ms, acquisition time [TA] = 2000 ms, TE = 30 ms, flip angle = 90°, voxel resolution = 3.0 mm³, FOV = 192x192 mm, and 32 transverse slices acquired parallel to the anterior commissure–posterior commissure [AC-PC] plane, providing whole-brain coverage). Each functional run was preceded by two additional TRs during which no data were recorded to allow for stabilization of longitudinal magnetization. Each run was 4 minutes, 12 seconds long. Sparse-sampling (Hall et al., 1999) was chosen to present auditory stimuli in silence, as this is the case when nonword repetition is administered clinically. The sparse scanning protocol allows the stimuli to be presented, and overt verbal responses to occur, during silent intervals between EPI volume acquisitions, precluding motion artifacts from children’s head motion, as well as from structural changes in oral and nasal cavity, during speech production. The reduction in acoustic stimulation during scanning improves not only dynamic range in auditory brain responses (Gaab et al., 2006), but has been shown to reduce spurious activation in executive regions related to the cognitive demands of noise exclusion (Pelle, 2014; Gaab et al., 2008). A sparse-sampling rate of TR = 6 s, with four seconds of sparse delay, was selected to allow sufficient time for children to listen to and repeat the nonwords in silence while still sampling BOLD signal at the peak of the hemodynamic response at each trial (Perrachione & Ghosh, 2013). The delay in volume acquisition leads to substantially improved contrast-to-noise ratio (CNR) for each sparse volume relative to analogous continuously-sample EPI volume acquisition. Furthermore, we used an optimized procedure for modelling sparse data

using convolved HRF, which allows detection of neural activation to events across discontinuous volume acquisitions (see Perrachione & Ghosh, 2013, for evidence that even slow sparse designs can capture event-related functional activation through HRF convolution).

MRI data pre-processing

The fMRI DICOM data were converted to NIFTI files following Brain Imaging Data Structure (BIDS) formatting using HeuDiConv (<https://github.com/nipy/heudiconv>; Gorgolewski et al., 2017; Yarkoni et al., 2019). Anatomical and functional images were preprocessed using fMRIPrep (v.1.4.1), including T1 correction, brain extraction, normalization to the fsaverage5 template, tissue segmentation, and motion correction procedures [(Esteban et al., 2019); fMRIPrep Available from: 10.5281/zenodo.852659, RRID:SCR_016216], a Nipype-based tool (Gorgolewski et al., 2017); Nipype available from: doi:10.5281/zenodo.581704, RRID:SCR_002502). Slice timing correction was not applied due to the sparse data collection method that was used. See **Appendix A** for additional details on the use of fMRIPrep.

Framewise displacement (FD; Power et al., 2014), together with other commonly used confounds, such as motion parameters, was calculated for each functional run using the implementation in Nipype. FD and spatial s.d. of the data after temporal differencing (known as ‘DVARs’) are calculated for each functional run, both with their implementations in Nipype (Esteban et al., 2018). Volumes with an $FD \geq 2\text{mm}$ were regressed out from the analysis (Siegel et al., 2014) and participants with $\geq 33\%$ of motion outlier volumes were removed from the analyses (See **Appendix A** for the group-average number of removed volumes and the total number of removed participants due to motion). The average number of removed volumes and average FD by group are reported in **Appendix A**. The two groups were not statistically different on either the number of outliers or motion parameters (**Appendix A**).

fMRI analyses

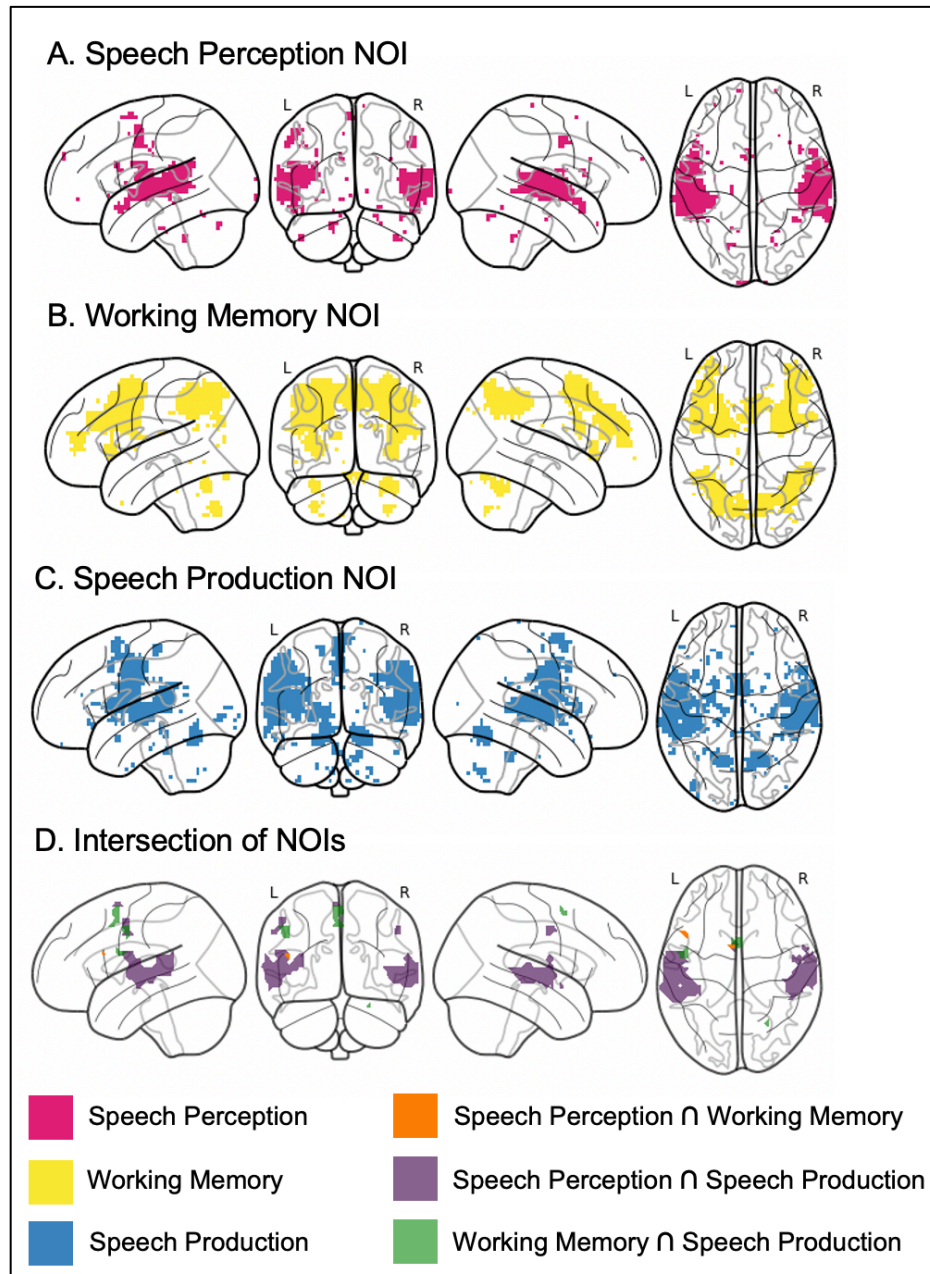
Analyses and hypotheses for this project were pre-registered using Open Science Framework (<https://osf.io/xerpc>).

Networks of interest

We conducted the group-level analyses within hypothesis-driven networks of interest (NOIs) representing the three underlying cognitive processes involved in nonword repetition, as described in the introduction (**Figure 1.3**). The NOIs were regions associated with the following functional terms on Neurosynth.org (<http://neurosynth.org>): “speech perception” (derived from a meta-analysis of 97 studies), “working memory” (derived from a meta-analysis of 1091 studies), and “speech production” (derived from a meta-analysis of 107 studies). The use of a priori NOIs derived from meta-analyses is useful in reducing the effects of multiple comparisons. Univariate analyses were performed within a union (i.e., combination) mask of the three NOIs to examine within- and between-group differences. Multivariate pattern analysis (MVPA) were separately conducted in each NOI, as well as the union and conjunction (i.e., intersection of at least 2 NOIs) regions. To confirm that we did not miss possible effects of condition or group occurring in regions outside of the a priori NOIs, and to examine possible compensatory activity in the autistic group, additional univariate and multivariate analyses in the complement regions to the NOIs (i.e., those outside the networks) were subsequently performed after the main analyses. The location and extent of the Neurosynth-derived NOIs were compared to significant activation in the NT and ASD groups for nonword repetition of all syllable lengths > rest ($p < 0.001$) in the current study (**Appendix A**), and were found to be generally inclusive of the task-based activation patterns.

Figure 2.3.

Networks of interest (NOI) derived from Neurosynth.org used for the univariate and multivariate analyses



Note: Regarding overlap between NOIs, the speech perception network has a 58.7% with the speech production network and a 1% overlap with the working memory network. The working memory network has a 0.4% overlap with the speech perception network and a 2.4% overlap with the speech production network. The speech production network has a 30.5% overlap with the perception network and a 3% overlap with the working memory network. For coordinates and automated anatomical labeling (based on the Automated Anatomical Labeling Atlas) of the regions included within each of the networks of interest, see **Appendix A**.

Univariate analyses

We used SnPM (version 13.1.08; <http://niso.org/Software/SnPM13/>) to perform the non-parametric univariate analyses within the combination of the (1) speech perception, (2) working memory, and (3) speech production regions (Nichols & Holmes, 2002). We used within- and between-group cluster-wise nonparametric t -tests (5000 permutations, cluster-forming threshold $p < 0.01$, cluster-level FWE-corrected $p < 0.05$) to compare activation between task and rest (i.e., all nonword repetition syllable lengths > rest) and to compare activation as a function of phonological working memory-load (i.e., five > two-syllables). These analyses were followed by a more stringent threshold (cluster-forming threshold $p < 0.001$, cluster-level FWE-corrected $p < 0.05$) to compare activation as a function of phonological working memory-load (i.e., five syllable > two syllable). Given the wide age-range of our participants, we also conducted the above univariate analysis with age and motion as a covariate, to ensure that age was not driving the observed effects. Univariate analyses were also conducted in the union and complement of the phonological working memory NOIs. See **Appendix A** for methods used to identify significant anatomical brain regions.

To investigate whether the univariate differences between NT and ASD were also seen in the RD group, we extracted each participant's beta values from the voxels in which we detected a primary group difference (NT ($n=28$) > ASD ($n=23$)). We conducted independent samples one-tailed t -test to compare of the beta values for (1) RD vs. ASD and (2) RD vs. NT.

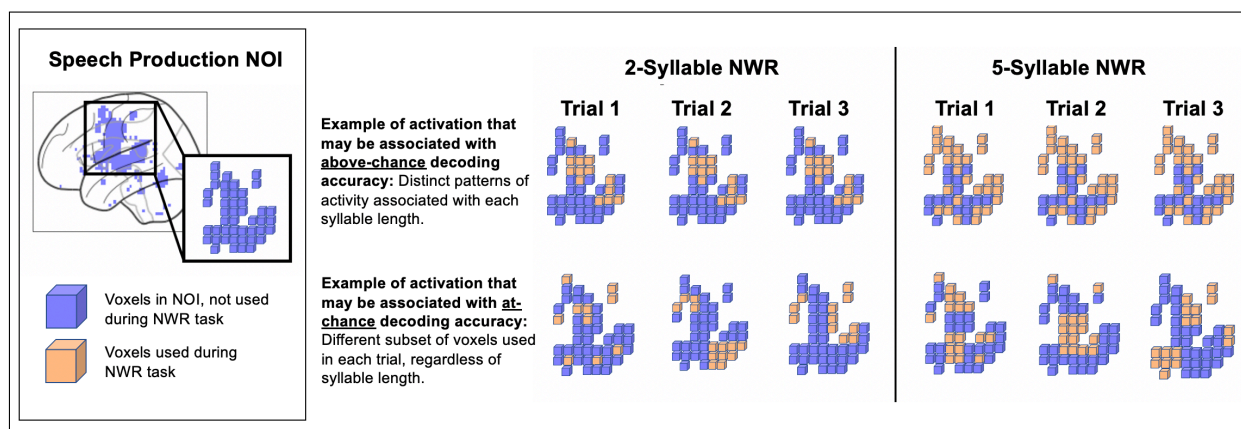
Multivariate pattern analysis (MVPA)

We used MVPA as an information-based decoding approach, to ask how the decoding accuracy of information content (i.e., voxel-wise brain activity patterns associated with two vs. five syllable nonword repetition) varies between autistic and NT children (Bae et al., 2020). MVPA is used to explore brain activation distinctions in processing between stimulus features and was chosen due to its usefulness in analyzing distributed patterns of activity that differentiate between

multiple experimental conditions within typical and clinical populations (Kamitani & Tong, 2005; Bray, 2009). The accuracy of a decoder can be used as a measure of neural specificity or reliability for a particular task (in this study, nonword repetition of varying syllable lengths; Park et al., 2010; Haynes & Rees, 2006; Rissman & Wagner, 2012). Given the likely increased variability within children, a within-subject MVPA may yield additional insight regarding neural function and compensatory processing strategies at an individual level, beyond that of a traditional univariate approach (Hanke, 2009). By comparing the average decoding accuracy of each group, we can better understand the underlying neural specificity during nonword repetition for each group. **Figure 2.4** provides a schematic representation of how patterns of activity may be decoded between two conditions.

Figure 2.4.

A Schematic Representation of How Patterns of Activity May be Decoded Between Two Conditions (Two Syllable NWR and Five Syllable NWR) Within the Speech Production NOI



Note: The first example depicts distinct patterns of activation associated with each syllable length and would likely be associated with above chance decoding accuracy. The second example depicts an inconsistent pattern of voxel activation across syllable lengths, and would likely be associated with at-chance decoding accuracy.

We used the BIDS app version of PyMVPA for running the MVPA analysis (Sajjadtorabian, 2018; <https://github.com/BIDS-Apps/PyMVPA>). Using leave-one-out support vector machine classification, the classifier decoded between five-syllable and two-syllable word

activation (i.e., between high and low phonological working memory load) for each participant. The contrast of five syllables vs. two syllables was selected to examine how the autistic and NT groups use neural information to represent increased phonological working memory load in each NOI. The classifier was iterated through each NOI to reveal how well the pattern of activation in each NOI differentiates between high and low phonological working memory and whether there are group differences (similar methods used in: Koster-Hale et al., 2013).

All included participants had at least 67% of two and five syllable trials included to ensure sufficient trials for the analysis (**Appendix A**). There were no significant differences between groups or conditions in the number of included trials for decoding. Given our relatively low number of runs ($n = 3$) and trials, we generated estimates per individual trials rather than per run, which allowed for more, albeit noisier, estimates. We used feature-wise, chunk-wise z-scoring of the data, which normalizes the dataset by scaling features into approximately the same range and removing the mean.

Average decoding accuracy was calculated for the ASD and NT groups within each of the three phonological working memory NOIs (**Figure 2.2**), as well as the combination and conjunction areas. In both the autistic and NT groups, decoding accuracy for each network was compared against chance. An ANOVA was used to compare decoding accuracy between NT and ASD across the three primary NOIs: speech perception, working memory, and speech production. Post-hoc independent one-tailed t-tests were used to compare between decoding accuracy for ASD and NT within each of the three NOIs. The p -values for these t-tests were FDR-corrected for each between-group test conducted.

We then completed a targeted follow-up analysis by comparing MVPA decoding accuracy within the network in which we ultimately detected group differences in our primary group

analysis (i.e., speech production network). Independent one-tailed t-tests were used to compare decoding accuracy between (1) RD vs. ASD, and (2) RD vs. NT.

Data Availability Statement

Pre-processed z-maps have been uploaded to GitHub as an anonymized data set. Additionally, one anonymized subject's data has been uploaded to GitHub with an accompanying tutorial to reproduce our MVPA results at a single-subject level.

Results

In-scanner behavioral results

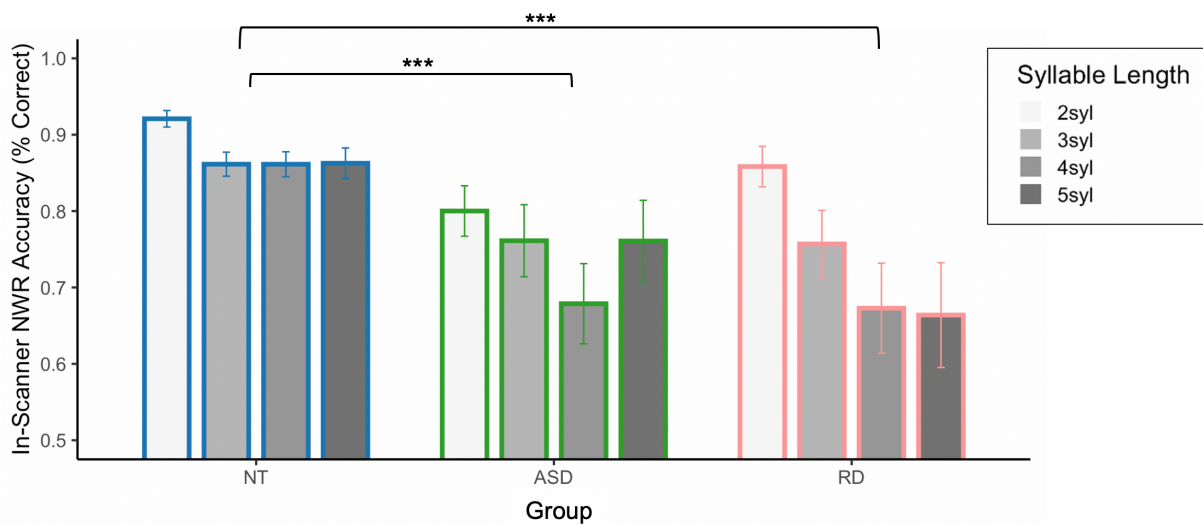
We used an analysis of variance (ANOVA) to test the effects of diagnosis and syllable length on nonword repetition accuracy among the RD, NT, and ASD groups (**Figure 2.5**). The nonword repetition recordings from one participant in the RD group were missing due to technical issues. The ANOVA revealed significant main effects of groups ($F(1, 256) = 30.32, p < 0.0001$) and syllable length ($F(3, 256) = 4.71, p = .003$), but no interaction between diagnosis and syllable length ($F(3, 256) = 1.28, p = .283$).

A comparison between the NT and ASD groups showed a main effect of diagnosis ($F(1, 196) = 29.89, p < 0.0001$) with the ASD group performing worse, a marginal effect of syllable length ($F(3, 196) = 2.43, p = .066$), with better performance at shorter lengths, and no interaction between diagnosis and syllable length ($F(3, 196) = 0.699, p = .55$). A comparison between the NT and RD groups showed a main effect of diagnosis with the RD group performing worse ($F(1, 164) = 40.29, p < 0.0001$), a main effect of syllable length ($F(3, 164) = 5.629, p = .0011$), with better performance at shorter lengths, and a marginal interaction between diagnosis and syllable length ($F(3, 164) = 2.284, p = .081$), suggesting a steeper decrease of accuracy with increasing syllable length in RD than NT. A comparison between the ASD and RD groups showed no effect

of diagnosis ($F(1, 144) = 0.115, p = .735$), a main effect of syllable length ($F(3, 144) = 3.099, p = .0288$), and no interaction between diagnosis and syllable length ($F(3, 144) = 0.787, p = .503$). Within each group, the effect of syllable-length was significant in the NT ($F(3, 88) = 3.378, p = .021$) and RD ($F(3, 56) = 3.029, p = .037$) groups, but not in the ASD group ($F(3, 88) = 1.17, p = .325$).

Figure 2.5.

In-scanner Performance on Nonword Repetition Task.



Note: Mean accuracy of nonword repetition within each condition for the NT, ASD, and RD groups. The NT had significantly higher overall performance than the ASD and the RD groups. There was no significant difference between the RD and ASD group. Error bars represent standard errors. Statistical significance marked by * is based on one-tailed T-tests on group differences. *** $p < 0.001$.

Univariate comparison of group differences in task-based brain activation

We investigated group differences in activation during nonword repetition > rest between the NT and ASD groups by conducting nonparametric one-sample (within) and two-sample (between) t -tests (cluster-forming threshold $p < 0.01$, cluster-level FWE-corrected threshold $p < 0.05$). Both groups showed significant clusters of activation in locations known to be involved in nonword repetition, such as bilateral STG and SMA (task vs. rest contrast; **Appendix A**). No between-group

differences were detected in the complement non-phonological working memory NOI regions or within the union of the three NOIs (**Figure 2.3**), suggesting that similar brain regions are involved in nonword repetition in both groups of children.

Univariate comparison of group differences in brain activation as a function of nonword length

We examined the effect of syllable length (five syllables > two syllables) in each group using nonparametric one-sample *t*-tests (cluster-forming non-parametric threshold $p < 0.01$, cluster-level FWE-corrected threshold $p < 0.05$) within the union of the three NOIs. The NT group showed significantly greater activation in bilateral STG and left STS for longer vs. shorter nonwords (**Table 2.2, Figure 2.6A**). The ASD group also showed significantly greater activation in bilateral STG for the five > two syllables contrast.

The NT group had significantly greater difference between the two- and five-syllable conditions in pre-SMA and SMA-proper (peak coordinates: -3, 15, 55; **Table 2.2, Figure 2.6A**) compared to the ASD group. The reverse contrast of ASD > NT did not yield any significant clusters. The group difference remained consistent with use of a more stringent cluster-forming threshold ($p < 0.001$) and cluster-level FWE-corrected threshold of 0.05, but with smaller cluster sizes ($k = 19$; peak coordinates: -3, 15, 44), and also remained significant when age and motion were added as covariates. Notably, the SMA in the NT group had greater activation to longer than shorter nonwords ($t(27) = 2.24$, $p = 0.017$), as measured by the beta weight from two and five syllables extracted from the SMA ROI. In contrast, the SMA in the ASD group had greater activation during shorter than longer nonwords ($t(22) = 2.99$, $p = .003$). No within-group syllable-load effect nor between-group differences were found in the complement regions outside of the phonological working memory NOIs.

To better understand the relationship between SMA activation and phonological working memory abilities in the ASD and NT groups, we correlated each participant's load-dependent activation difference (beta weight from the contrast of five > two syllables) extracted from the SMA ROI with their overall in-scanner nonword repetition accuracy. Using a nonparametric Spearman's rank correlation in each group, we found a significant positive correlation for the NT group between SMA activation and in-scanner performance ($\rho = 0.32$, one-tailed $p < 0.05$) (**Figure 2.6B**). In contrast, the ASD group did not have a significant correlation between SMA activation and in-scanner performance ($\rho = -0.23$, one-tailed $p = .15$), which was a significantly weaker association compared to that in the NT group ($z = 1.89$, one-tailed $p = .03$).

We performed targeted follow-up analyses of activation within the SMA for two control comparisons (1) a smaller subset of the full ASD and NT groups who were additionally matched on CELF Core Language scores (*language-matched groups*; See **Appendix A** for details) and (2) the RD group, who did not differ from the ASD group on nonword repetition performance either in or out of the scanner. The language-matched NT group continued to show significantly greater SMA activation with increased nonword syllable load than the language-matched ASD group ($t(33.3) = 3.89$, one-tailed $p < 0.001$). The RD group showed significantly greater SMA activation during increased phonological working memory load (five > two syllables) than the ASD group ($t(29) = 2.34$, one-tailed $p = .02$), but did not differ from the NT group ($t(31) = 0.52$) (**Figure 2.6C**).

Table 2.2.

Peak Coordinates and Anatomical Locations of Significant Clusters for Within and Between Group Analyses for Five-Syllable > Two-Syllable Contrast

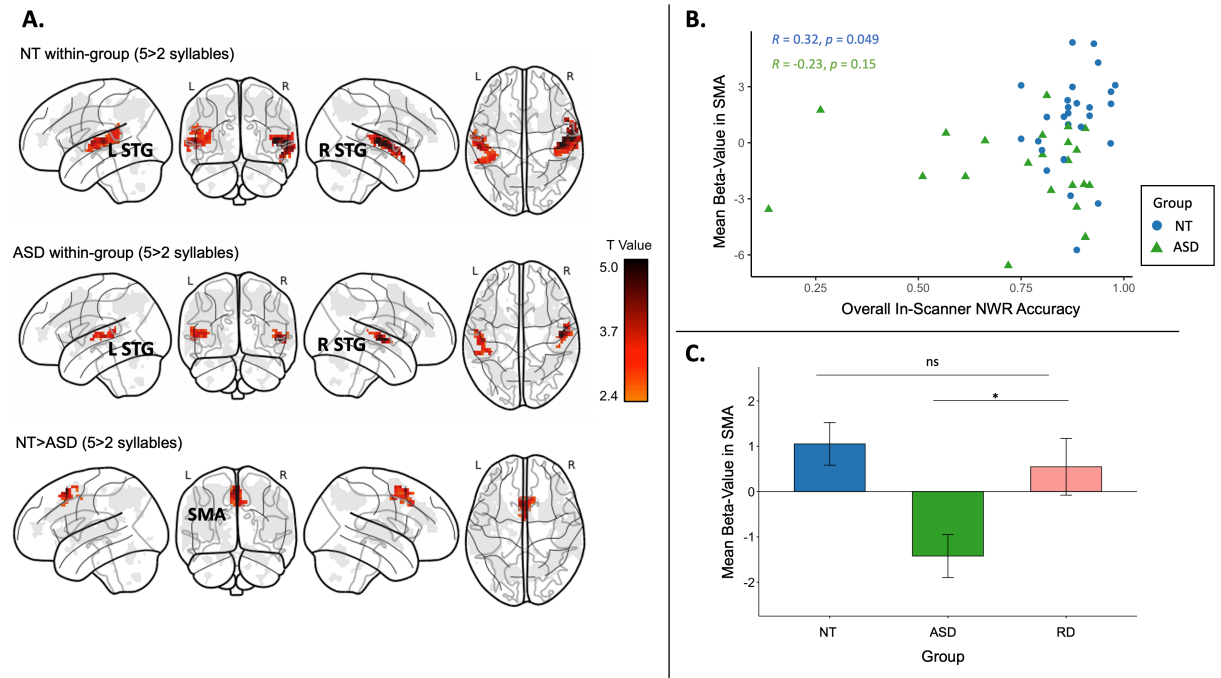
Comparison	Cluster-level				Voxel-level				Anatomical location
	p(FWE-corr)	p(uncorr)	k	T	p(uncorr)	x	y	z	

NT (within group)	.004	.0008	201	5.47	.0002	45	-21	5	R STG
	.047	.0126	46	4.29	.0002	54	-42	1	R STS
ASD (within group)	.006	.0011	178	4.19	.0002	-54	-24	4	L STG
	.011	.0019	58	5.09	.0002	54	-9	1	R STG
NT > ASD	.007	.0012	66	3.52	.0010	-51	-27	5	L STG
	.014	.0019	135	4.43	.0002	-3	15	55	SMA

Note: Cluster defining threshold of $p < 0.01$ and FWE-corrected $p < 0.05$.

Figure 2.6.

Effects of Syllable Length in fMRI



Note: **A.** Two vs. five-syllable effect within the neurotypical (NT) and autism spectrum disorder (ASD) groups and the difference between the two groups (cluster-forming nonparametric $p < 0.01$, cluster-level FWE-corrected $p < 0.05$). **B.** Activation responses at SMA were positively correlated with in-scanner nonword repetition performance in the NT group, but not in the ASD group. **C.** The reading disability (RD) group did not vary significantly in SMA activation magnitude from the NT group but was significantly higher than the ASD group. Error bars represent standard errors. Spearman's Rank correlation p -values are reported as one-tailed values. * $p < 0.05$; ** $p < 0.01$

Multivariate group comparison of neural decoding of nonword length

We used the accuracy of multivoxel trial-wise decoding between the two- and five-syllable conditions as a measure of how consistently each of the NOIs represented phonological working

memory load within each individual. A between-group comparison of decoding accuracy within each NOI for the ASD and NT groups, using a group by NOI ANOVA, showed a main effect of diagnosis, with significantly lower decoding accuracy in the ASD group compared to the NT group ($F(1, 147) = 7.65, p < 0.01$), which remained significant after controlling for in-scanner performance ($F(1, 146) = 7.34, p < 0.01$).

There was above-chance decoding of phonological working memory load in all NOIs in the NT group, as revealed by a one-sample t-test against chance. In contrast, there was not above-chance decoding in any of the tested networks in the ASD group, including the combination, union, or complement networks (**Table 2.3**). After we controlled for multiple comparisons, a significant group difference was only observed only in the speech-production NOI, shown as higher decoding accuracy in the NT group than the ASD group ($t(47) = 2.5$, FDR-corrected $p < 0.05$). Further, even for the two- vs. four-syllable contrast where more robust behavioral differences were observed in the ASD group (See **Figure 2.5**), neural decoding of these two conditions in the ASD group did not differ significantly from the classification of two- vs. five-syllable load (**Appendix A**).

To examine whether the differences in decoding accuracy in the speech production NOI were being driven by brain regions common to the speech production and speech perception networks, or by those common to the speech production and working memory networks, we conducted two pairwise t-tests to compare decoding accuracy within these network intersections between ASD and NT. Neither group comparison was significant (**Table 2.3**), suggesting that the differences in the speech production network arise within core speech-production specific regions, rather than the overlapping regions between the networks.

To determine whether decoding accuracy within the speech production network was associated with behavioral characteristics in the ASD group, we conducted a series of correlations

between decoding accuracy and several behavioral measures (in-scanner nonword repetition accuracy, Social Responsiveness Scale (SRS-2; Constantino & Gruber, 2012; Frazier et al., 2014), Social Communication Questionnaire (SCQ; Rutter et al., 2003), and ADOS Severity Score). Within the autism group, decoding accuracy from the speech production network was not associated with the in-scanner nonword repetition performance (Spearman's $\rho < 0.08$). Instead, we observed modest negative correlations (Spearman's ρ 's > -0.30) between the decoding accuracy of the speech-production network and all three autism-related measures (**Appendix A**), with lower decoding accuracy correlating with higher autism traits across all measures.

As we did for the univariate analyses, we again conducted targeted follow-up analyses of MVPA decoding accuracy within the speech production network for the two control groups. As seen with the full group, the significantly higher decoding accuracy was found for the language-matched NT group vs. the language-matched ASD group ($t(32) = 1.86, p = .036$). In contrast, there were no significant differences in decoding accuracy between the RD and ASD group ($p = .72$) or between the RD and NT group ($p = .123$).

Table 2.3.

Decoding accuracy for NT and ASD groups within the NOIs, intersections of interest, complement regions, and combination of all networks.

	NT		ASD				NT vs. ASD
	Mean	SD	Against Chance FDR corrected p-value	Mean	SD	Against Chance FDR corrected p-value	FDR corrected p-value
Networks of Interest							
Speech Perception Network	56.2	8.8	.001	52.5	8.7	.142	.110
Working Memory Network	54.3	8.8	.009	52.3	7.4	.142	.194
Speech Production Network	57.7	7.1	.000	52.6	7.3	.141	.024
Perception $\cap$ Production	55.5	8.6	.002	51.6	9.4	.248	.122
Working Memory $\cap$ Production	53.8	8.8	.015	51.0	7.8	.279	.143
Complement Regions	55.8	8.9	.002	53.3	7.1	.142	NA
Union of all three NOIs	56.4	7.8	.0005	52.2	6.5	.142	NA
Intersection of all three NOIs	55.5	8.8	.002	51.6	9.4	.248	NA

Note: Between group-comparisons were conducted for the three NOIs. $\cap$: Intersection of two NOIs.

Discussion

To better understand the brain bases of phonological working memory challenges in autism, we investigated how brain activation during nonword repetition differed in autistic versus neurotypical children across three networks of interest known to support speech perception, working memory, and speech production. Children with autism exhibited reduced nonword repetition accuracy both in and out of the scanner. Group differences in activation during nonword repetition were found in speech-production regions, the SMA and pre-SMA, which showed greater modulation of activation for longer versus shorter nonwords in NT than in ASD. Multivoxel pattern analysis within the speech production network classified shorter versus longer nonwords less accurately for autistic than neurotypical children. The specificity for the speech production difference in autism was supported by lack of an activation difference in children with RD. Further, these univariate and multivariate findings *were* replicated in a comparison between NT and ASD children who were matched on language skills, indicating that the brain differences were likely related to the autism phenotype rather than reflecting differences in core language abilities.

Overall, these results suggest that phonological working memory challenges in autism may be due to differential engagement of speech-motor areas, including especially SMA, as seen here during nonword repetition. In contrast, our findings suggest that nonword challenges in autism may *not* be related to differences in perceptual processing, as activation in superior temporal regions was not atypical in this group. Instead, the results suggest typical phonological representation in autistic children, consistent with some previous neuroimaging and behavioral research (Ceponiene et al., 2003; Pomper et al., 2019).

The univariate analyses revealed that while the ASD and NT groups recruited similar, canonical phonological working memory networks during the nonword repetition task overall, they differed in how this activation was modulated by increased phonological working memory

load. Although the ASD group showed typical sensitivity to increased syllable length within bilateral STG, they differed specifically in their recruitment of the pre-SMA and SMA proper – brain regions associated with planning and controlling motor and speech actions (Tremblay & Gracco, 2009; Nachev et al., 2007; Alario et al., 2006). SMA and pre-SMA in particular have been implicated in sequencing and planning of speech (Guenther & Hickok, 2016). SMA also showed parametric responses for nonword syllable load in purely receptive tasks (Perrachione et al., 2017; Strand et al., 2008), suggesting that engagement of SMA can signal a covert articulatory rehearsal process and does not require explicit speech production.

Beyond the SMA, studies of nonword repetition in neurotypical adults have shown increased load-associated activation in several regions associated with speech production including left precentral gyrus, left inferior frontal gyrus (IFG), and right and left cerebellum (Scott & Perrachione, 2019). A study of nonword repetition in neurotypical adults showed task-specific increases in the functional connections between pre-SMA and left dorsal premotor cortex (Hartwigsen et al., 2013). According to the GODIVA neurocomputational model of speech processing (Bohland et al., 2010), IFG and SMA/pre-SMA subserve parallel phonological buffers, with the former encoding phonemic sequences and the latter representing abstract syllable frames and/or premotor sequencing of speech (Tremblay & Gracco, 2006; Alario et al., 2006). Evidence also suggested a diverse range of activations in both regions across individual participants (Scott & Perrachione, 2019). In the current study, the differences in SMA, but not in IFG, suggests potential intact phonemic processing during nonword repetition in autistic children, but abnormal motor sequencing, which independent of the specific phonemic or phonological content of the stimuli.

Not only was there less activation of the SMA in the ASD group, but the ASD group showed a fundamentally different pattern of activation in the SMA. Unlike the NT group who showed

greater activation for longer than shorter nonwords, the ASD group showed the reverse pattern with greater activation for shorter than longer nonwords. This suggests there may be less consistent or reliable motor sequencing for longer nonwords in ASD. A diffusion tensor imaging study in autistic adults with average nonverbal IQ revealed a weaker connection in the speech production network between the SMA and ventral premotor cortex compared to neurotypical controls (Peeva et al., 2013). It is possible that SMA functions less efficiently and coherently with other regions in the speech motor system when children with ASD are faced with longer syllable sequences.

The MVPA approach provided evidence supporting altered speech motor function beyond SMA. We used MVPA as an information-based decoding approach, to understand how the decoding accuracy of information content (i.e., voxel-wise brain activity patterns associated with two vs. five syllable nonword repetition) varied between ASD and NT children. Above chance decoding accuracy indicates consistent patterns of activation associated with each syllable length, while at-chance decoding suggests an indiscriminate pattern of voxel activation across syllable lengths. Across speech perception, working memory, and speech production networks, only the speech production network had reliably lower decoding accuracy in the ASD group compared to the NT group. Further, lower decoding accuracy within the speech production network was modestly correlated with more severe autism-based symptoms in the ASD group. This functional alteration in the speech motor network remained even when matching language skills across the ASD and NT groups.

Together, these complementary findings support a speech-motor atypicality account for reduced nonword repetition performance in autism. However, the differences in SMA and speech production network do not imply that differences only took place at the last temporal phase of nonword repetition. The fMRI task, involving immediate repetition after listening, does not allow differentiation of temporal components of the behavior: speech perception, maintenance, and

speech production. In addition, our fMRI nonword repetition task does not differentiate between verbal short-term recall and working memory, two associated cognitive processes captured by the same paradigm (Oberauer et al., 2012; Archibald & Gathercole, 2007). Future research investigating the brain bases of phonological working memory should include additional behavioral tasks to validate these findings (e.g., serial recall of lists of monosyllabic nonwords).

Our findings add to a growing literature showing broad motor system differences in ASD, including motor differences related to speech and language. In fact, motor differences in autism have been found to be more closely linked to the diagnostic features of ASD, social communication and restricted and repetitive behaviors, than cognitive ability (Ketcheson et al., 2021). Related to speech and language, better fine motor skills in autistic children are associated with stronger receptive and expressive language skills (Mody et al., 2017). Visuomotor, general coordination, and fine motor skills have been shown to predict language delays in autistic children (Bhat et al., 2022). Similarly, oromotor, manual motor skills, and fine motor skills have been shown to be strong indicators of speech development (Gernsbacher et al., 2008; Bal). Notably, Krishnan et al. (2013) demonstrated that NT children's ability to imitate and reproduce nonlinguistic oromotor sequences predicted their performance in a nonword repetition task. This relationship is independent of language skills and general cognitive abilities, suggesting an involvement of domain-general motor skills and planning underlying phonological working memory, which is consistent with the present findings across autistic children with and without reduced concurrent language skills. These findings, together with the neuroimaging evidence from the current study, suggest that atypical motor function in autism might not be a byproduct of concurrent language challenges, but a possible indicator for further risks in language development due to its association with reduced phonological working memory.

Atypical activation of SMA occurred in the ASD group but not in an RD group that had similar behavioral impairment in nonword repetition. ASD and RD, however, may share some brain differences in relation to nonword repetition, such as atypical structural connectivity (Lu et al., 2016). Perhaps atypical functional organization in speech perception (Boets et al., 2013) and working memory networks (Hachmann et al., 2020; Jeffries & Everatt, 2004; Smith-Spark & Fiske, 2007) may better explain the challenges of RDs in nonword repetition. Although SMA activation in RD was similar to NT during nonword repetition, future research with a larger sample size is necessary to investigate other potential mechanisms related to speech production in RD, such as phonemic/phonological representation as opposed to sequencing, and whether they are related to nonword repetition deficits given the known difficulties in phonological processing in reading disabilities (Kovelman et al., 2012; Van Bergen et al., 2014). Behavioral research has documented some evidence of speech production difficulties in RD, including reduced speech rate and increased speech errors (e.g., Catts, 1989; Lambrecht Smith et al., 2010; Carroll et al., 2014). Some researchers connect speech errors with phonological awareness (Cabbage et al., 2018), while others suggest a complex relationship between speech sound disorder, reading disability, and language impairment (Hayiou-Thomas et al., 2017). Therefore, research in other language disorders will provide greater clarification for the specific and shared neural substrates underlying a common behavioral hallmark (cf. Pigdon et al., 2020; Liegeois et al., 2011).

The present findings should be considered within the context of several methodological limitations. Due to the scarcity of pediatric neuroimaging studies, the networks of interest for our analyses were based on meta-analyses from primarily adult participants. However, there is evidence that the core speech perception, speech production, and cognition networks are also implicated and selective in children during development (Hiersche et al., 2022). Despite this, we demonstrated a substantial overlap between the Neurosynth networks of interest and the activation

patterns during nonword repetition observed in our pediatric sample (**Appendix A**). An alternative approach for future studies is to generate task-specific individual-specific functional ROIs using a separate speech production task. These individualized ROIs would be developmentally more appropriate and might improve decoding accuracy across all participant groups.

Secondly, a replication study with a larger sample size is warranted. Due to the heterogeneity of the populations with autism and reading disabilities, the current study is limited in its power to detect more subtle group differences that may exist in other brain networks or brain regions. In fact, none of the networks of interest in either the ASD group or the RD group reliably classified syllable length above chance. Therefore, the MVPA results may also reveal a global neural inconsistency phenomenon that is neither unique to autism nor unique to the speech production network (e.g., Centanni et al., 2021). In addition, a larger sample size is necessary to uncover reliable brain-behavior correlations. Despite the small sample size, however, within the context of this paper, the targeted investigation with rigorous control for multiple comparisons allowed us to test the specificity of the brain and behavioral differences to autism.

Thirdly, there was a wide range of participants' age in each group. Groups were carefully matched on data quality and age. However, because we used one standard MNI template across all groups, the degree of spatial warping for younger brains was greater than older brains during normalization, which might lead to additional age-related confound during preprocessing. Although our main neuroimaging findings do not appear to be associated with age (**Supplementary Table 6**), a larger sample size in future studies will allow us to examine the developmental changes of motor speech functions in age-specific anatomical spaces. Different developmental trajectory might still exist between autistic and neurotypical groups.

2. Conclusions and future directions

Autistic children - even those with core language skills matched to NT children - may perform and function differently during nonword repetition tasks due to atypical engagement of the speech production network and especially the SMA. The vulnerability in phonological working memory even for language-matched autistic children could lead to downstream language and language-learning challenges (Baddeley, Gathercole, & Papagno, 2017), which will require future longitudinal samples to verify and characterize. Speech production, and its underlying neural mechanisms, is an important part of phonological working memory that has been largely neglected in previous phonological working memory research. Future research should investigate how motor functions in both the verbal and nonverbal domains predict language learning outcomes in not only autistic children, but also children with other communication disorders who have concomitant motor and working memory difficulties.

Acknowledgements

We thank the participating families for their contribution to our study; Satrajit Ghosh and Anila D'Mello for advice on data analysis; Michelle Han, Sara Beach, Abigail Cyr, Clara Baron-Hyppolite, Colleen Buckless, Katalina Sher, Calvin Goetz, and Mengyuan Xu for help on data collection; and Atsushi Takahashi, Sheeba Arnold Anteraper, and Steven Shannon at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, Massachusetts Institute of Technology, for technical support.

Funding

This work was funded by National Institutes of Health Grant No. R01 DC011339 (to JDEG, HTF, and KW); TKP was supported by NIH grant R03 DC014045 and a NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation; AMO was funded by NIH Training

Grant: T32 DC000038 and the Hock E. Tan and K. Lisa Yang Center for Autism Research. ZQ was supported by NIH grant R21DC017576.

Competing interests

None.

References

- Acheson, D. J., Hamidi, M., Binder, J. R., & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience* (Vol. 23, Issue 6, pp. 1358–1367). <https://doi.org/10.1162/jocn.2010.21519>
- Adams, L. (1998). Oral-motor and motor-speech characteristics of children with autism. In *Focus on Autism and Other Developmental Disabilities*, 13(2), 108–112. <https://doi.org/10.1177/108835769801300207>
- Alho, J., Bharadwaj, H., Khan, S., Mamashli, F., Perrachione, T. K., Losh, A., ... & Kenet, T. (2021). Altered maturation and atypical cortical processing of spoken sentences in autism spectrum disorder. *Progress in Neurobiology*, 203, 102077. <https://doi.org/10.1016/j.pneurobio.2021.102077>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.
- Adams, A.M., & Gathercole, S. E. (1996). Phonological working memory and speech production in preschool children. *Journal of Speech, Language, and Hearing Research*, 38(2), 403–414.
- Alario, F. X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain Research*, 1076(1), 129–143. <https://doi.org/10.1016/j.brainres.2005.11.104>
- Archibald, L., & Gathercole, S. (2007). Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? *Applied Psycholinguistics*, 28(4), 587–606. <https://doi.org/10.1017/S0142716407070324>
- Archibald, L. M. D., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research: JSLHR*, 52(4), 899–914. [https://doi.org/10.1044/1092-4388\(2009/08-0099\)](https://doi.org/10.1044/1092-4388(2009/08-0099))
- Archibald. (2017). Working memory and language learning: A review. *Child Language Teaching and Therapy*, 33(1), 5–17. <https://doi.org/10.1177/0265659016654206>
- Baddeley A. (1992). Working memory. *Science* (New York, N.Y.), 255(5044), 556–559. <https://doi.org/10.1126/science.1736359>
- Baddeley, A., & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (pp. 47–89). [https://doi.org/10.1016/s0079-7421\(08\)60452-1](https://doi.org/10.1016/s0079-7421(08)60452-1)
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173. <https://doi.org/10.1037/0033-295x.105.1.158>

- Bae, G.-Y., Leonard, C. J., Hahn, B., Gold, J. M., & Luck, S. J. (2020). Assessing the information content of ERP signals in schizophrenia using multivariate decoding methods. *NeuroImage: Clinical*, 25, 102179. <https://doi.org/10.1016/j.nicl.2020.102179>
- Baird, G., Slonims, V., Simonoff, E., & Dworzynski, K. (2011). Impairment in non-word repetition: a marker for language impairment or reading impairment?. *Developmental medicine and Child Neurology*, 53(8), 711–716. <https://doi.org/10.1111/j.1469-8749.2011.03936.x>
- Barendse, E. M., Hendriks, M. P., Jansen, J. F., Backes, W. H., Hofman, P. A., Thoonen, G., Kessels, R. P., & Aldenkamp, A. P. (2013). Working memory deficits in high-functioning adolescents with autism spectrum disorders: Neuropsychological and neuroimaging correlates. *Journal of Neurodevelopmental Disorders*, 5(1), 14. <https://doi.org/10.1186/1866-1955-5-14>
- Bhat, A. N., Landa, R. J., & Galloway, J. C. (2011). Current perspectives on motor functioning in infants, children, and adults with autism spectrum disorders. *Physical Therapy*, 91(7), 1116–1129. <https://doi.org/10.2522/ptj.20100294>
- Bishop D. V. (2010). Overlaps between autism and language impairment: phenomimicry or shared etiology?. *Behavior genetics*, 40(5), 618–629. <https://doi.org/10.1007/s10519-010-9381-x>
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341–345.
- Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., & Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Research in Developmental Disabilities*, 32(2), 560–570. <https://doi.org/10.1016/j.ridd.2010.12.020>
- Bray, S. (2009). Applications of multivariate pattern classification analyses in developmental neuroimaging of healthy and clinical populations. In *Frontiers in Human Neuroscience* (Vol. 3). <https://doi.org/10.3389/neuro.09.032.2009>
- Bruno, R. M., & Walker, S. C. (1999). Comprehensive Test of Phonological Processing (CTOPP). In *Diagnostique* (Vol. 24, Issues 1-4, pp. 69–82). <https://doi.org/10.1177/153450849902401-408>
- Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: from loop to convolution. *Journal of Cognitive Neuroscience*, 20(5), 762–778. <https://doi.org/10.1162/jocn.2008.20501>
- Centanni, T., Beach, S. D., Ozernov-Palchik, O., May, S., Pantazis, D., & Gabrieli, J. D. E. (2021). Categorical perception and influence of attention on neural consistency in response to speech sounds in adults with dyslexia. *Annals of Dyslexia*. <https://doi.org/10.1007/s11881-021-00241-1>
- Buchsbaum, B. R., & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. *Cortex*, 112, 134–148. <https://doi.org/10.1016/j.cortex.2018.11.010>

Cabbage, Farquharson, K., Iuzzini-Seigel, J., Zuk, J., & Hogan, T. P. (2018). Exploring the overlap between dyslexia and speech sound production deficits. *Language, Speech & Hearing Services in Schools*, 49(4), 774–786. https://doi.org/10.1044/2018_LSHSS-DYSLC-18-0008

Carroll, J. M., Mundy, I. R., & Cunningham, A. J. (2014). The roles of family history of dyslexia, language, speech production and phonological processing in predicting literacy progress. *Developmental Science*, 17(5), 727–742. <https://doi.org/10.1111/desc.12153>

Catts H. W. (1989). Speech production deficits in developmental dyslexia. *The Journal of speech and hearing disorders*, 54(3), 422–428. <https://doi.org/10.1044/jshd.5403.422>

Christophel, Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J.-D. (2016). The distributed nature of working memory. *Trends in Cognitive Sciences*, 21(2), 111–124. <https://doi.org/10.1016/j.tics.2016.12.007>

Conant, Bouchard, K. E., & Chang, E. F. (2013). Speech map in the human ventral sensory-motor cortex. *Current Opinion in Neurobiology*, 24(1), 63–67. <https://doi.org/10.1016/j.conb.2013.08.015>

Constantino, J. N. & Gruber C. P. (2012). *Social Responsiveness Scale–Second Edition (SRS-2)*. Torrance, CA: Western Psychological Services.

Cui, J., Gao, D., Chen, Y., Zou, X., & Wang, Y. (2010). Working Memory in Early-School-Age Children with Asperger’s Syndrome. *Journal of Autism and Developmental Disorders*, 40(8), 958–967. <https://doi.org/10.1007/s10803-010-0943-9>

Diehl, J. J., Friedberg, C., Paul, R., & Snedeker, J. (2015). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Development and Psychopathology*, 27(3), 867–884. <https://doi.org/10.1017/S0954579414000741>

Ehrhorn, A. M., Adlof, S. M., Fogerty, D., & Laing, S. (2021). Probing Phonological Processing Differences in Nonword Repetition for Children with Separate or Co-Occurring Dyslexia and Developmental Language Disorder. *Scientific Studies of Reading*, 25(6), 486–503. <https://doi.org/10.1080/10888438.2020.1849223>

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, 16(1), 111–116.

Estes, K. G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. In *Journal of Speech, Language, and Hearing Research* (Vol. 50, Issue 1, pp. 177–195). [https://doi.org/10.1044/1092-4388\(2007/015\)](https://doi.org/10.1044/1092-4388(2007/015))

Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194.

Fedorenko, E., Morgan, A., Murray, E., Cardinaux, A., Mei, C., Tager-Flusberg, H., Fisher, S. E., & Kanwisher, N. (2016). A highly penetrant form of childhood apraxia of speech due to deletion of 16p11.2. *European Journal of Human Genetics: EJHG*, 24(2), 302–306. <https://doi.org/10.1038/ejhg.2015.149>

Fiez, J. A. (2016). Neural basis of phonological short-term memory. *Neurobiology of Language*, 855–862) <https://doi.org/10.1016/b978-0-12-407794-2.00068-7>

Frazier, T. W., Ratliff, K. R., Gruber, C., Zhang, Y., Law, P. A., & Constantino, J. N. (2014). Confirmatory factor analytic structure and measurement invariance of quantitative autistic traits measured by the Social Responsiveness Scale-2. *Autism*, 18(1), 31–44. <https://doi.org/10.1177/1362361313500382>

Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford: Blackwell.

Frith, C. (2003). What do imaging studies tell us about the neural basis of autism? *Novartis Foundation Symposium*, 251, 149–166; discussion 166–176, 281–297.

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. *Human Brain Mapping*, 28(8), 721–732. <https://doi.org/10.1002/hbm.20299>

Gaab, N., Gabrieli, J. D. E., & Glover, G. H. (2008). Resting in peace or noise: Scanner background noise suppresses default-mode network. *Human Brain Mapping*, 29(7), 858–867. <https://doi.org/10.1002/hbm.20578>

Gabig, C. S. (2008). Verbal working memory and story retelling in school-age children with autism. *Language, Speech, and Hearing Services in Schools*, 39(4), 498–511.

Gargan, C. E. & Andrianopoulos, M. V. (2021). Receptive and expressive lexical stress in adolescents with autism. *International Journal of Speech Language Pathology*, 24(6), 636–646. <https://doi.org/10.1080/17549507.2021.2008006>

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, 27(4), 513–543.

Gathercole, S. E., Hitch, G. J., Service E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology*, 33(6), 966–979.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The Children's Test of Nonword Repetition: a test of phonological working memory. *Memory* (Hove, England), 2(2),

103–127. <https://doi.org/10.1080/09658219408258940>

Gilbert, Meuwese, J. D. ., Towgood, K. J., Frith, C. D., & Burgess, P. W. (2009). Abnormal functional specialization within medial prefrontal cortex in high-functioning autism: a multi-voxel similarity analysis. *Brain*, 132(4), 869–878. <https://doi.org/10.1093/brain/awn365>

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill Goldsmith, H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(1), 43–50. <https://doi.org/10.1111/j.1469-7610.2007.01820.x>

Gorgolewski, K. J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M. M., Churchill, N. W., Cohen, A. L., Craddock, R. C., Devenyi, G. A., Eklund, A., Esteban, O., Flandin, G., Ghosh, S. S., Guntupalli, J. S., Jenkinson, M., Keshavan, A., Kiar, G., Liem, F., ... Poldrack, R. A. (2017). BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS Computational Biology*, 13(3), e1005209. <https://doi.org/10.1371/journal.pcbi.1005209>

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2006). The Autism Diagnostic Observation Schedule (ADOS). *Journal of Autism and Developmental Disorders*, 37(4), 613–627. <https://doi.org/10.1007/s10803-006-0280-1>

Gray, S., Fox, A. B., Green, S., Alt, M., Hogan, T. P., Petscher, Y., & Cowan, N. (2019). Working memory profiles of children with dyslexia, developmental language disorder, or both. *Journal of Speech, Language, and Hearing Research*, 62(6), 1839–1858. https://doi.org/10.1044/2019_JSLHR-L-18-0148

Guenther. (2016). *Neural Control of Speech*. The MIT Press.

Habib, A., Harris, L., Pollick, F., & Melville, C. (2019). A meta-analysis of working memory in individuals with autism spectrum disorders. *PloS one*, 14(4), e0216198. <https://doi.org/10.1371/journal.pone.0216198>

Hachmann, W. M., Cashdollar, N., Postiglione, F., & Job, R. (2020). The relationship of domain-general serial order memory and reading ability in school children with and without dyslexia. *Journal of Experimental Child Psychology*, 193, 104789. <https://doi.org/10.1016/j.jecp.2019.104789>

Hanke, M. (2009). PyMVPA: a unifying approach to the analysis of neuroscientific data. In *Frontiers in Neuroinformatics* (Vol. 3). <https://doi.org/10.3389/neuro.11.003.2009>

Hansen, B. B., & Klopfer, S. O. (2006). Optimal full matching and related designs via network flows. *Journal of Computational and Graphical Statistics*, 15(3), 609–627.

Hartwigsen, G., Saur, D., Price, C. J., Baumgaertner, A., Ulmer, S., & Siebner, H. R. (2013). Increased facilitatory connectivity from the pre-SMA to the left dorsal premotor cortex during

pseudoword repetition. *Journal of Cognitive Neuroscience*, 25(4), 580-594.

Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C., & Snowling, M. J. (2017). When does speech sound disorder matter for literacy? The role of disordered speech errors, co-occurring language impairment and family risk of dyslexia. *Journal of Child Psychology and Psychiatry*, 58(2), 197–205. <https://doi.org/10.1111/jcpp.12648>

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7(7), 523–534.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, 6(3), 121–143. <https://doi.org/10.1016/j.plrev.2009.06.001>

Hiersche, K. J., Schettini, E., Li, J., & Saygin, Z. M. (2022). The language network is selective and distinct from other cognition in both function and connectivity in early childhood. *bioRxiv*. <https://doi.org/10.1101/2022.08.11.503597>

Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *The Quarterly Journal of Experimental Psychology*, 28(4), 603–621. <https://doi.org/10.1080/14640747608400587>

Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1–28.

Hus, V., & Lord, C. (2014). The Autism Diagnostic Observation Schedule, Module 4: Revised Algorithm and Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, 1–17. <https://doi.org/10.1007/s10803-014-2080-3>

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, 92(1-2), 101–144.

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*, 10(11), 480–486.

Järvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P. (2008). Enhanced perceptual processing of speech in autism. *Developmental Science*, 11(1), 109–121.

Jeffries, S., & Everatt, J. (2004). Working memory: its role in dyslexia and other specific learning difficulties. *Dyslexia (Chichester, England)*, 10(3), 196–214. <https://doi.org/10.1002/dys.278>

Joanisse, M., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*, 77(1), 30–60. <https://doi.org/10.1006/jecp.1999.2553>

Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and*

Psychopathology, 16(1), 137–155. <https://doi.org/10.1017/s095457940404444x>

Kamitani, Y. & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5), 679–685. <https://doi.org/10.1038/nn1444>

Kasdan, A. V., Burgess, A. N., Pizzagalli, F., Scartozzi, A., Chern, A., Kotz, S. A., Wilson, S. M., & Gordon, R. L. (2022). Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic review. *Neuroscience and Biobehavioral Reviews*, 136, 104588. <https://doi.org/10.1016/j.neubiorev.2022.104588>

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test: KBIT 2 ; Manual*.
Ketcheson, L. R., Pitchford, E. A., & Wentz, C. F. (2021). The relationship between developmental coordination disorder and concurrent deficits in social communication and repetitive behaviors among children with autism spectrum disorder. *Autism Research*, 14(4), 804–816.

Key, A. P. & D'Ambrose Slaboch, K. (2021). Speech processing in autism spectrum disorder: an integrative review of auditory neurophysiology findings. *Journal of Speech, Language, and Hearing Research*, 64(11), 4192–4212. https://doi.org/10.1044/2021_JSLHR-20-00738

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: implications for genetic subgroups. *Language and Cognitive Processes*, 16(2-3), 287–308.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, 30(3), 859–873.

Koshino, H., Carpenter, P. A., Minshew, N. J., Cherkassky, V. L., Keller, T. A., & Just, M. A. (2005). Functional connectivity in an fMRI working memory task in high-functioning autism. In *NeuroImage* (Vol. 24, Issue 3, pp. 810–821). <https://doi.org/10.1016/j.neuroimage.2004.09.028>

Koster-Hale, J., Saxe, R., Dungan, J., & Young, L. L. (2013). Decoding moral judgments from neural representations of intentions. *Proceedings of the National Academy of Sciences*, 110(14), 5648–5653. <https://doi.org/10.1073/pnas.1207992110>

Kovelman, I., Norton, E. S., Christodoulou, J. A., Gaab, N., Lieberman, D. A., Triantafyllou, C., ... & Gabrieli, J. D. (2012). Brain basis of phonological awareness for spoken language in children and its disruption in dyslexia. *Cerebral Cortex*, 22(4), 754–764. <https://doi.org/10.1093/cercor/bhr094>

Kover, S. T., Atwood, A. K. (2013). Establishing equivalence: methodological progress in group-matching design and analysis. *American Journal of Intellectual Developmental Disabilities*, 118(1):3-15. <https://doi.org/10.1352/1944-7558-118.1.3>

Krishnan, S., Alcock, K. J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words: children's oromotor skills predict nonword repetition abilities. *Journal of Speech, Language, and Hearing Research*, 56(6), 1800–1812.

[https://doi.org/10.1044/1092-4388\(2013/12-0206\)](https://doi.org/10.1044/1092-4388(2013/12-0206))

Lambrecht Smith, Roberts, J. A., Locke, J. L., & Tozer, R. (2010). An exploratory study of the development of early syllable structure in reading-impaired children. *Journal of Learning Disabilities*, 43(4), 294–307. <https://doi.org/10.1177/0022219410369094>

Liégeois, F., Morgan, A.T., Connelly, A., & Vargha-Khadem, F. (2011). Endophenotypes of FOXP2 : Dysfunction within the human articulatory network. *European Journal of Paediatric Neurology*, 15(4), 283–288. <https://doi.org/10.1016/j.ejpn.2011.04.006>

Lipkin, B. et al. (in prep) LanA (Language Atlas): A probabilistic atlas for the language network based on fMRI data from >800 individuals.

Lei, J., Jones, L., & Brosnan, M. (2021). Exploring an e-learning community's response to the language and terminology use in autism from two massive open online courses on autism education and technology use. *Autism: The International Journal of Research and Practice*, 25(5), 1349–1367.

Leung, R. C., Vogan, V. M., Powell, T. L., Anagnostou, E., & Taylor, M. J. (2016). The role of executive functions in social impairment in autism spectrum disorder. *Child Neuropsychology*, 22(3), 336–344. <https://doi.org/10.1080/09297049.2015.1005066>

Long, Katlowitz, K., Svirsky, M., Clary, R., Byun, T., Majaj, N., Oya, H., Howard, M., & Greenlee, J. . (2016). Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation. *Neuron*, 89(6), 1187–1193. <https://doi.org/10.1016/j.neuron.2016.01.032>

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr, Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223. <https://doi.org/10.1023/A:1005592401947>

Lord, C., DiLavore, P. C., Gotham, K., Guthrie, W., Luyster, R., Risi, S., & Rutter, M. (2012). *Autism diagnostic observation schedule: ADOS-2* (2nd Edition). Western Psychological Services. <http://www.worldcat.org/title/autism-diagnostic-observation-schedule-ados-2/oclc/851410387>

Lu, C., Qi, Z., Harris, A., Weil, L. W., Han, M., Halverson, K., Perrachione, T. K., Kjelgaard, M., Wexler, K., Tager-Flusberg, H., & Gabrieli, J. D. E. (2016). Shared neuroanatomical substrates of impaired phonological working memory across reading disability and autism. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 1(2), 169–177.

Macizo, P., Soriano, M. F., & Paredes, N. (2016). Phonological and visuospatial working memory in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 46(9), 2956–2967. <https://doi.org/10.1007/s10803-016-2835-0>

Manis, F., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen,

- A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology*, 66(2), 211–235. <https://doi.org/10.1006/jecp.1997.2383>
- Majerus, S. (2013). Language repetition and short-term memory: An integrative framework. *Frontiers in Human Neuroscience*, 7, Article 357. <https://doi.org/10.3389/fnhum.2013.00357>
- Marvel, C. L., & Desmond, J. E. (2010). The contributions of cerebro-cerebellar circuitry to executive verbal working memory. *Cortex*, 46(7), 880–895. <https://doi.org/10.1016/j.cortex.2009.08.017>
- McGettigan, Warren, J. E., Eisner, F., Marshall, C. R., Shanmugalingam, P., & Scott, S. K. (2011). Neural correlates of sublexical processing in phonological working memory. *Journal of Cognitive Neuroscience*, 23(4), 961–977. <https://doi.org/10.1162/jocn.2010.21491>
- McPartland, J. C., Wu, J., Bailey, C. A., Mayes, L. C., Schultz, R. T., & Klin, A. (2011). Atypical neural specialization for social percepts in autism spectrum disorder. *Social Neuroscience*, 6(5-6), 436–451. <https://doi.org/10.1080/17470919.2011.586880>
- Mody, M., Shui, A. M., Nowinski, L. A., Golas, S. B., Ferrone, C., O'Rourke, J. A., & McDougle, C. J. (2017). Communication deficits and the motor system: exploring patterns of associations in autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, 47(1), 155–162. <https://doi.org/10.1007/s10803-016-2934-y>
- Mottron, L. & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), *The Development of Autism: Perspectives from Theory and Research* (pp. 131–148). Lawrence Erlbaum Associates Publishers.
- Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *NeuroImage*, 36 Suppl 2(3-3), T155–T163. <https://doi.org/10.1016/j.neuroimage.2007.03.034>
- Nadig, A., & Mulligan, A. (2017). Intact non-word repetition and similar error patterns in language-matched children with autism spectrum disorders: A pilot study. *Journal of Communication Disorders*, 66, 13–21.
- Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human Brain Mapping*, 15(1), 1–25. <https://doi.org/10.1002/hbm.1058>
- Norrelgen, F., Fernell, E., Eriksson, M., Hedvall, Å., Persson, C., Sjölin, M., Gillberg, C., & Kjellmer, L. (2015). Children with autism spectrum disorders who do not develop phrase speech in the preschool years. *Autism: The International Journal of Research and Practice*, 19(8), 934–943. <https://doi.org/10.1177/1362361314556782>
- Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working

memory: An interference model of complex span. *Psychonomic Bulletin & Review*, 19, 779–819. <https://doi.org/10.3758/s13423-012-0272-4>

Ozonoff, S., Rogers, S. J., & Pennington, B. F. (1991). Asperger's syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 32(7), 1107–1122. <https://doi.org/10.1111/j.1469-7610.1991.tb00352.x>

Pang, E. W., Valica, T., MacDonald, M. J., Taylor, M. J., Brian, J., Lerch, J. P., & Anagnostou, E. (2016). Abnormal brain dynamics underlie speech production in children with autism spectrum disorder. *Autism Research: Official Journal of the International Society for Autism Research*, 9(2), 249–261. <https://doi.org/10.1002/aur.1526>

Park, J., Carp, J., Hebrank, A., Park, D. C., & Polk, T. A. (2010). Neural specificity predicts fluid processing ability in older adults. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(27), 9253–9259. <https://doi.org/10.1523/JNEUROSCI.0853-10.2010>

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>

Peelle, J. E. (2014). Methodological challenges and solutions in auditory functional magnetic resonance imaging. *Frontiers in Neuroscience*, 8, 253. <https://doi.org/10.3389/fnins.2014.00253>

Peeva, Tourville, J. A., Agam, Y., Holland, B., Manoach, D. S., & Guenther, F. H. (2013). White matter impairment in the speech network of individuals with autism spectrum disorder. *NeuroImage Clinical*, 3(C), 234–241. <https://doi.org/10.1016/j.nicl.2013.08.011>

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. <https://doi.org/10.1016/j.tics.2012.08.009>

Pellicano E. (2012). The development of executive function in autism. *Autism Research and Treatment*, 146132. <https://doi.org/10.1155/2012/146132>

Perrachione, T. K., & Ghosh, S. S. (2013). Optimized design and analysis of sparse-sampling fMRI experiments. *Frontiers in Neuroscience*, 7, 55. <https://doi.org/10.3389/fnins.2013.00055>

Perrachione, T. K., Ghosh, S. S., Ostrovskaya, I., Gabrieli, J. D. E., & Kovelman, I. (2017). Phonological working memory for words and nonwords in cerebral cortex. *Journal of Speech, Language, and Hearing Research*, 60(7), 1959–1979. https://doi.org/10.1044/2017_JSLHR-L-15-0446

Pickles, A., Anderson, D. K., & Lord, C. (2014). Heterogeneity and plasticity in the development of language: a 17-year follow-up of children referred early for possible autism. In *Journal of Child Psychology and Psychiatry* (Vol. 55, Issue 12, pp. 1354–1362). <https://doi.org/10.1111/jcpp.12269>

- Pigdon, L., Willmott, C., Reilly, S., Conti-Ramsden, G., Liegeois, F., Connelly, A., & Morgan, A. T. (2020). The neural basis of nonword repetition in children with developmental speech or language disorder: An fMRI study. *Neuropsychologia*, 138, 107312. <https://doi.org/10.1016/j.neuropsychologia.2019.107312>
- Poeppel, D. (2015). Speech Perception. In *Brain Mapping* (pp. 429–434). <https://doi.org/10.1016/b978-0-12-397025-1.00264-5>
- Pomper, R., Ellis Weismer, S., Saffran, J., & Edwards, J. (2019). Specificity of Phonological Representations for Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(8), 3351–3363. <https://doi.org/10.1007/s10803-019-04054-5>
- Posar, A., & Visconti, P. (2022). Early Motor Signs in Autism Spectrum Disorder. *Children (Basel, Switzerland)*, 9(2), 294. <https://doi.org/10.3390/children9020294>
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142–2154. <https://doi.org/10.1016/j.neuroimage.2011.10.018>
- Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, 84, 320–341. <https://doi.org/10.1016/j.neuroimage.2013.08.048>
- Rabiee, A., Vasaghi-Gharamaleki, B., Samadi, S. A., Amiri-Shavaki, Y., Alaghband-Rad, J., Seyedin, S., & Hosseini, S. (2018). Impaired nonverbal working memory in high-functioning autism spectrum disorder. *Medical Journal of the Islamic Republic of Iran*, 32, 107. <https://doi.org/10.14196/mjiri.32.107>
- Rissman, J., & Wagner, A. D. (2012). Distributed representations in memory: insights from functional brain imaging. *Annual Review of Psychology*, 63, 101–128. <https://doi.org/10.1146/annurev-psych-120710-100344>
- Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services.
- Sachse, M., Schlitt, S., Hainz, D. et al. Executive and Visuo-motor Function in Adolescents and Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 43, 1222–1235 (2013). <https://doi.org/10.1007/s10803-012-1668-8>
- Sajjadtorabian. (2018, August 10). sajjadtorabian/PyMVPA: PyMVPA BIDS App version v1.0.0 (Version v1.0.0). Zenodo. <http://doi.org/10.5281/zenodo.1343531>
- Samuel, A. G. (2011). Speech Perception. *Annual Review of Psychology*, 62(1), 49–72. <https://doi.org/10.1146/annurev.psych.121208.131643>
- Scott, T. L., & Perrachione, T. K. (2019). Common cortical architectures for phonological working

memory identified in individual brains. *NeuroImage*, 202, 116096. <https://doi.org/10.1016/j.neuroimage.2019.116096>

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical evaluation of language fundamentals, fourth edition—Screening test* (CELF-4 screening test). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain and Language*, 164, 77–105. <https://doi.org/10.1016/j.bandl.2016.10.004>

Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory*, 15(1), 34–56. <https://doi.org/10.1080/09658210601043384>

Soulières, I., Mottron, L., Saumier, D., & Larochelle, S. (2007). Atypical categorical perception in autism: Autonomy of discrimination? *Journal of Autism and Developmental Disorders*, 37(3), 481–490. <https://doi.org/10.1007/s10803-006-0172-4>

Strand, F., Forssberg, H., Klingberg, T., & Norrelgen, F. (2008). Phonological working memory with auditory presentation of pseudo-words — An event related fMRI Study. *Brain Research*, 1212(30 May), 48–54. <https://doi.org/10.1016/j.brainres.2008.02.097>

Szatmari, P., Georgiades, S., Duku, E., Bennett, T. A., Bryson, S., Fombonne, E., Mirenda, P., Roberts, W., Smith, I. M., Vaillancourt, T., Volden, J., Waddell, C., Zwaigenbaum, L., Elsabbagh, M., Thompson, A., & Pathways in ASD Study Team. (2015). Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. *JAMA Psychiatry*, 72(3), 276–283. <https://doi.org/10.1001/jamapsychiatry.2014.2463>

Tager-Flusberg, H., & Kasari, C. (2013). Minimally verbal school-aged children with autism spectrum disorder: the neglected end of the spectrum. *Autism Research: Official Journal of the International Society for Autism Research*, 6(6), 468–478. <https://doi.org/10.1002/aur.1329>

Tarar, J. M., Meisinger, E. B., & Dickens, R. H. (2015). Test Review: Test of Word Reading Efficiency—Second Edition (TOWRE-2) by Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. In *Canadian Journal of School Psychology* (Vol. 30, Issue 4, pp. 320–326). <https://doi.org/10.1177/0829573515594334>

Torgesen, J. K., Wagner, R., & Rashotte, C. (2012). *Test of Word Reading Efficiency: (TOWRE-2)*.

Tourville, J. A., & Guenther, F. H. (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes*, 26(7), 952–981. <https://doi.org/10.1080/01690960903498424>

Tremblay, P., Gracco, V. L. (2009) Contribution of the pre-SMA to the production of words and

non-speech oral motor gestures, as revealed by repetitive transcranial magnetic stimulation (rTMS). *Brain Research*, 1268, 112-124. <https://doi.org/10.1016/j.brainres.2009.02.076>

Tryfon, A., Foster, N. E. V., Sharda, M., & Hyde, K. L. (2018). Speech perception in autism spectrum disorder: An activation likelihood estimation meta-analysis. *Behavioural Brain Research*, 338, 118–127. <https://doi.org/10.1016/j.bbr.2017.10.025>

Van Bergen, E., van der Leij, A., & de Jong, P. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience*, 8, 346–346. <https://doi.org/10.3389/fnhum.2014.00346>

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychological Review*, 121(4), 649–675. <https://doi.org/10.1037/a0037665>

Vogan, V. M., Francis, K. E., Morgan, B. R., Smith, M. L., & Taylor, M. J. (2018). Load matters: neural correlates of verbal working memory in children with autism spectrum disorder. *Journal of Neurodevelopmental Disorders*, 10(1), 19. <https://doi.org/10.1186/s11689-018-9236-y>

Wang, Y., Zhang, Yb., Liu, Ll. et al. A Meta-Analysis of Working Memory Impairments in Autism Spectrum Disorders. *Neuropsychol Review*, 27, 46–61 (2017). <https://doi.org/10.1007/s11065-016-9336-y>

Weismer, S. E., Bruce Tomblin, J., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword Repetition Performance in School-Age Children With and Without Language Impairment. In *Journal of Speech, Language, and Hearing Research* (Vol. 43, Issue 4, pp. 865–878). <https://doi.org/10.1044/jslhr.4304.865>

West, K. L. (2019). Infant motor development in autism spectrum disorder: A synthesis and meta-analysis. *Child Development*, 90(6), 2053-2070. <https://doi.org/10.1111/cdev.13086>

White, S., Frith, U., Milne, E., Rosen, S., Swettenham, J., & Ramus, F. (2006). A double dissociation between sensorimotor impairments and reading disability: A comparison of autistic and dyslexic children. *Cognitive Neuropsychology*, 23(5), 748–761. <https://doi.org/10.1080/02643290500438607>

Whitehouse, Barry, J. G., & Bishop, D. V. (2008). Further defining the language impairment of autism: Is there a specific language impairment subtype? *Journal of Communication Disorders*, 41(4), 319–336. <https://doi.org/10.1016/j.jcomdis.2008.01.002>

Whyatt, C., & Craig, C. (2013). Sensory-motor problems in Autism. *Frontiers in Integrative Neuroscience*, 7, 51. <https://doi.org/10.3389/fnint.2013.00051>

Williams, D. L., Goldstein, G., Carpenter, P. A., & Minshew, N. J. (2005). Verbal and spatial working memory in autism. *Journal of Autism and Developmental Disorders*, 35(6), 747–756.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). The profile of memory function in children with autism. *Neuropsychology*, 20(1), 21–29. <https://doi.org/10.1037/0894-4105.20.1.21>

Williams, D., Payne, H., & Marshall, C. (2013). Non-word repetition impairment in autism and specific language impairment: evidence for distinct underlying cognitive causes. *Journal of Autism and Developmental Disorders*, 43(2), 404–417.

Woodcock, R.W. (2011). *Woodcock reading mastery tests, third edition*. Pearson.

Yarkoni, T., Markiewicz, C. J., de la Vega, A., Gorgolewski, K. J., Salo, T., Halchenko, Y. O., McNamara, Q., DeStasio, K., Poline, J.-B., Petrov, D., Hayot-Sasson, V., Nielson, D. M., Carlin, J., Kiar, G., Whitaker, K., DuPre, E., Wagner, A., Tirrell, L. S., Jas, M., ... Blair, R. (2019). PyBIDS: Python tools for BIDS datasets. *Journal of Open Source Software*, 4(40). <https://doi.org/10.21105/joss.01294>

Yeung, M. K., Lee, T. L., & Chan, A. S. (2019). Right-lateralized frontal activation underlies successful updating of verbal working memory in adolescents with high-functioning autism spectrum disorder. *Biological Psychology*, 148, 107743.

Zampella, C. J., Wang, L. A. L., Haley, M., Hutchinson, A. G., & de Marchena, A. (2021). Motor skill differences in autism spectrum disorder: A clinically focused review. *Current Psychiatry Reports*, 23(10), 64.

Zhou, Ma, W., & Zhan, L. (2020). A deficit in using prosodic cues to understand communicative intentions by children with autism spectrum disorders: An eye-tracking study. *First Language*, 40(1), 41–63. <https://doi.org/10.1177/0142723719885270>

Chapter 3:

Acoustic and Self-Reported Prosodic Differences in Autistic Adults

Author contributions: **Conceptualization & methodology** (Amanda M. O’Brien; John D. E. Gabrieli; Annie Cardinaux; Cindy Li); **data acquisition and curation** (Amanda M. O’Brien; Cindy Li; Annie Cardinaux; Luwei Quan); **statistical and experiment analysis** (Amanda M. O’Brien; Annie Cardinaux; Luwei Quan); **interpretation** (Amanda M. O’Brien; Annie Cardinaux; John D. E. Gabrieli), **writing—original draft** (Amanda M. O’Brien; Annie Cardinaux), **writing—revisions** (John D. E. Gabrieli; Cindy Li); **funding** (John D. E. Gabrieli), **supervision** (John D. E. Gabrieli)

Funding: Research reported in this chapter was supported by The Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT. Amanda M. O’Brien was additionally supported by an NIDCD Institutional Training Grant (T32DC000038) and Fellowship (F31DC020864).

Additional acknowledgements: The authors extend their gratitude to The Simons Foundation Autism Research Initiative (SFARI) & Simons Foundation Powering Autism Research (SPARK) for facilitating recruitment for this study, and for all of the participants who contributed their time to make this work possible. The authors also thank the focus group consultants who supported the development of the VOX-Q measure. Additionally, the authors are grateful to feedback provided by Jennifer Zuk, Helen Tager-Flusberg, Zhenghan Qi, and Stefanie Shattuck-Hufnagel on various aspects of the study design and interpretation. Finally, the authors thank the students who provided valuable support across several components of data collection, curation, annotation, and scoring: Autumn Rambus, Simon Radhakrishnan, Sabrina He, Samantha Holler, Eva Naik, Kendall Rhee, and Tim Smith.

Abstract

Prosody—the use of pitch, metrical structure, duration, loudness, and voice quality—plays a key role in conveying emotion, emphasizing key points, and expressing pragmatic intent. While atypical prosody has been noted in autism since its earliest descriptions, prior research has largely overlooked spontaneous prosody (especially in adults), and has given limited attention to timing and pauses. Further, no study to date has examined autistic individuals’ lived experiences of prosody. To address these gaps, this chapter examines prosody in autism through two complementary approaches: (1) a novel subjective self-report measure capturing autistic adults’ experiences with prosody perception and production, and (2) an objective acoustic analysis of prosody during a spontaneous picture description task. Self-report data indicated that autistic individuals experience greater challenges with prosody, which strongly correlate with autism traits in both autistic and nonautistic participants. Acoustic analyses showed that the autistic participants had atypical speech timing, with longer pauses and a lower rate of speech. Importantly, the groups did not differ in diadochokinetic rate, a common speech-motor assessment, suggesting that these prosody differences are likely not due to an underlying speech-motor disorder (e.g., apraxia of speech), but may still be linked to broader motor system differences.

Introduction

Autism is diagnosed based on challenges with social communication and interaction (American Psychiatric Association [APA], 2013). Even among individuals on the spectrum who present with strong structural language skills, one aspect of language that contributes to communication challenges is *prosody* (DePape et al., 2012). Prosody - often referred to as the “melody” of speech - encompasses a range of features including pitch, loudness, duration, rate, and vocal quality, and is actively controlled by motor regions of the brain. Indeed, brain injuries to these motor regions lead to prosody challenges. Relevant here, motor differences are increasingly being identified as clinically meaningful in autism, including within the domain of communication, making motor involvement in prosody particularly relevant to the condition (Craig et al., 2021). While prosody differences have been widely studied in autism (for reviews see Fusaroli et al., 2017; Asghari et al., 2020), results are mixed. Moreover, most prior research focuses on children, infrequently employ spontaneous tasks which are most relevant to everyday communication, and, critically, do not consistently account for motor speech skills. In the present study, we consider the prosody of autistic adults from multiple perspectives: self-reported experiences with prosody, acoustic features (i.e., “global prosody”) during a spontaneous production task, and the relationship among these measures and autism traits, language skills, and nonverbal intelligence, while also considering motor speech skills.

Defining Prosody

Prosody manifests across three levels: (1) the acoustic cues, (2) what the cues signal based on their placement and context, and (3) how/why the cues are used by the speaker and/or interpreted by the listener. At the first level, acoustic cues can contribute to the intonation, rhythm, and quality of speech. Several key acoustic markers of prosody can be quantified from speech

samples including fundamental frequency (f_0 ; perceived as pitch), intensity (perceived as loudness), and vocal timbre (perceived as voice quality). Additionally, temporal aspects of the speech signal – including rate, duration of individual words and syllables, and the use of pauses – convey prosodic information. Prosodic cues are used concurrently with, and can extend beyond, individual phonetic segments of speech (Lehiste, 1970), and are often integrated rather than used in isolation (Dahl et al., 2023).

At the second level, the combination and placement of acoustic cues can signal and communicate multiple meanings and functions. For example, raising the pitch at the end of an utterance can change the same phrase from a statement to a question. Increasing the volume and duration of words in an utterance can help to disambiguate between multiple interpretations of the same sentence (e.g., “The RED car” vs. “The red CAR”). Simultaneous to modifying pitch and volume, temporal features can be used for emphasis, saliency, and clarity; for example, rate can convey emotion (e.g., boredom or urgency) and pauses can be used to disambiguate between multiple meanings. In these ways, a speaker’s prosody conveys more meaning than their literal words alone. The third level of prosody refers to how the cues are used within social contexts (e.g., why the speaker produces these features and how the listener interprets them). The interactions between the use, meaning, and interpretation of prosodic cues adds complexity to communication contexts that are already challenging for some individuals (e.g., those on the autism spectrum).

Finally, prosodic cues can also be analyzed and interpreted at different scales of communication (narrative, utterance, phrase, word, or phoneme). In the present study, we largely focus on “global” prosody - or the acoustic features that are extracted from the speech signal at the scale of the full narrative. This provides a broad big-picture overview of these features, but cannot be used to interpret use of these features at a finer-grain scale.

Prosody and Autism

Differences in prosody production have been noted as a perceptually salient feature in autistic speech since its earliest descriptions (e.g., Kanner, 1943). Differences in prosody production in autism are described in the DSM-V: “nonverbal communicative behaviors used for social interaction are manifested by...atypical use of...speech intonation.” Additionally, autistic individuals may demonstrate differences that are “subtle within individual modes (e.g., someone may have relatively good eye contact when speaking but noticeable in poor integration of...prosody...for social communication.” The prosody of autistic individuals has been identified as perceptually distinct from nonautistic individuals, both by experienced clinicians and naive listeners (de Marchena & Miller, 2017; Weed et al., 2023).

Considerations of Prosody in Current Diagnostic and Characterization Measures

The DSM-5 (APA, 2013) clearly cites prosody among the heterogeneous diagnostic criteria for autism, specifying that “deficits in nonverbal communicative behaviors used for social interaction are manifested by absent, reduced, or atypical use of [...] speech intonation” and that “impairment may be relatively subtle within individual modes (e.g. someone may have relatively good eye contact when speaking) but noticeable in poor integration of eye contact, gesture, body posture, *prosody*, and facial expression for social communication.” However, of the commonly used autism trait measures, few quantify aspects of prosody. Prosody is referenced in the gold-standard behavioral assessment of autism, the *Autism Diagnostic Observation Schedule-Revised* (ADOS-2, Lord et al., 1999). The ADOS-2’s code A2 (Modules 3 and 4; used in older adolescents and adults), describes “speech abnormalities that are characteristic of autism” including “speech patterns in intonation, volume, rhythm, or rate (not articulation) that are unusual, but not obviously characteristic of autism speech intonation and rhythm.” The moderate (“1”) rating specifies that

an individual exhibits “[l]ittle variation in pitch and tone; rather flat or exaggerated intonation, but not obviously peculiar, OR slightly unusual volume, AND/OR speech that tends to be somewhat unusually slow, fast, or jerky,” whereas the highest possible rating (“2”) specifies, “[s]peech that is clearly abnormal for ANY of the following reasons: slow and halting; inappropriately rapid; jerky and irregular in rhythm (other than ordinary stutter/stammer), such that there is some interference with intelligibility; odd intonation or inappropriate pitch and stress; markedly flat and toneless (“mechanical”); consistently abnormal volume” (ADOS-2 scoring booklet). While the A2 code score is not tallied into the cutoffs/totals in the originally published ADOS-2 scoring algorithm (Lord et al., 2012; Lord et al., 1999), it has since been included in the revised Module 4 scoring algorithm (Hus & Lord, 2014).

Commonly used self-report measures of autism traits do not capture intonation or other features of autistic prosody (*Autism Quotient (AQ*; Baron-Cohen et al., 2001); *Comprehensive Autistic Trait Inventory (CATI*; English et al., 2021)), with the exception of the *Social Responsiveness Scale (SRS*; Constantino, 2012), wherein two items (one related to perception and one related to production) reference prosodic intonation. Some formal assessments of prosody have been developed such as the PEP-C [Profiling Elements of Prosody in Speech-Communication] (Peppé & McCann, 2003; <http://www.peps-c.com/>) and the Social Language Development Test - Adolescents (SLDT-A; Bowers et al., 2008; <https://www.wpspublish.com/sldt-a-nu-social-language-development-testadolescent-normative-update>). However, no self-report measure of prosody experiences exists to the authors’ knowledge. A measure of prosody self-report could fill this gap across clinical populations.

Acoustic Characterization of Autistic Speech

Despite substantial efforts to quantify the nature of the perceptually distinct speech features observed in autism, researchers have had limited success in identifying “universal markers” of autistic speech. Studies have investigated a wide range of prosodic features, including pitch, timing, and intensity, measured both globally via more naturalistic tasks and locally via structured experimental tasks. Here, we focus largely on the global prosody findings relevant to pitch and timing, consistent with the present study.

Pitch differences and implications. Pitch differences have been consistently identified in autism. In fact, the primary finding of two meta-analyses of prosody-related research in autism (Fusaroli et al., 2017 and Asghari et al., 2020) was that autistic individuals had a higher mean pitch and larger pitch range than nonautistic individuals. Notably, however, these differences alone had relatively low accuracy (61-64%) in discriminating between autistic and nonautistic individuals (Fusaroli et al., 2017). Many studies have found that autistic children have a higher mean pitch than nonautistic children across multiple tasks, including in spontaneous speech (Lyakso et al., 2016; Sharda et al., 2010) and structured lexical elicitation measures (Filipe et al., 2014). with others showing null results for mean pitch (e.g., Diehl & Paul, 2013). A study in which adults engaged in a simulated social interaction also showed higher mean pitch among autistic adults (Drimalla et al., 2020). The standard deviation of pitch has also been investigated, with some studies finding that autistic children have a higher standard deviation of pitch than their nonautistic peers during spontaneous production tasks (Bonneh et al., 2011; Diehl & Paul, 2013) and picture naming (Nakai et al., 2014), with others showing no difference during imitation tasks (Diehl & Paul, 2012). Importantly, pitch-based differences can play an important social (culturally-specific) role in perception of others - one study showed that both males and females with a lower pitch were more likely to be selected as “leaders” (Klofstad et al., 2012).

Timing differences and implications. Another key prosodic feature that has been investigated in autistic speech is timing, including duration, speaking rate, articulation rate, and pauses. Duration (e.g., of lexical items, syllables, and voicing) is a “local” feature of speech prosody, and thus not examined here. In some studies with children, a slower speech rate has been reported (see Asghari et al., 2021 for a review), while other studies have not found rate differences (e.g., Nadig & Shaw, 2012). A meta-analysis by Asghari et al. (2021) did not find significant differences in rate. Three studies, all in children and adolescents, have examined pausing in autism: one finding longer pauses during social interactions (Feldstein et al., 1982), one finding fewer pragmatically motivated pauses (Thurber & Tager-Flusberg, 1993), and one finding more pauses (Morett et al., 2016). The systematic review by Fusaroli et al. (2017) reported that not enough statistical estimates were provided in the studies of timing to conduct a meta-analysis of the mixed results. Importantly, regularity of speech tempo can support listener comprehension (Hilton & Goldwater, 2021).

Multivariate and machine learning approaches. Multivariate, machine learning, and cluster-based approaches have been recommended as a potential path forward for making progress in understanding autistic prosody (Fusaroli et al., 2017; Weed et al, 2023). There has been some evidence that acoustic features of the speech signal can be useful in identification of autistic individuals within a heterogeneous group of nonautistic individuals (Briend et al., 2023), and that speech recordings alone can be useful for characterizing an autistic individuals' social communication challenges (Eni et al., 2025). Notably, however, even machine learning approaches to distinguishing autistic individuals based on prosody have been shown to have limitations in generalizability across datasets and tasks (Rybner et al., 2022).

Intersection of Autism and Motor Speech Disorders

Studies that focus on prosody in autism have rarely considered a motor-based perspective, or characterized motor speech skills among participants, making it hard to establish the extent to which motor differences do or do not underlie prosody differences in autism. This gap is especially critical given the known overlap between autism and motor speech disorders (e.g., apraxia of speech; Maffei et al., 2023), and growing evidence of motor differences being a clinically meaningful feature of autism, including within the domain of communication (Craig et al., 2021; Zampella et al., 2021). Indeed, motor delays and atypical movement development (e.g., stereotyped movement) have been identified as some of the earliest signs of autism, which often emerge even before the classical social communication differences (West, 2019; Posar & Visconti, 2022). Further, oromotor, manual motor, fine motor, and gross motor skills are correlated with speech and language (expressive and receptive) abilities in autistic individuals (Gernsbacher et al., 2008; Mody et al., 2017; Bedford et al., 2016).

Gaps in the Field

The present study combines autistic adults' self-reported prosody experiences with acoustics analyses of their spontaneous speech, addressing several key gaps in the literature. First, the examination of prosody experiences provides a novel and critical perspective for understanding prosody differences in autism. While nonautistic individuals have been asked to describe the prosody of autistic individuals (e.g., Weed et al., 2023), this study is the first to systematically ask how autistic individuals themselves perceive their own prosody, in conjunction with objective acoustic measures.

Second, most research on prosody and autism has focused on children and adolescents, with highly limited work examining the speech of autistic adults. Studying the speech of autistic adults allows for better understanding of how observed developmental differences persist or

change in adulthood, which is particularly important given the known protracted developmental trajectory of prosody perception (Filipa et al., 2022). Although research and clinical work often focuses on children, autism is a lifelong condition. An improved understanding of prosody differences and how they relate to overall communication challenges and autism traits in adulthood is important for providing meaningful support to autistic people across the lifespan.

Third, the present study analyzes speech produced during a spontaneous picture description task. While many prior studies have used structured experimental approaches to understanding prosody, picture description provides a more naturalistic window into speech production. Picture description tasks are commonly used in clinical settings), making the findings more clinically relevant (Bryant et al., 2016). Moreover, because this task does not require real-time social interactions, and participants completed the study online in the absence of an experimenter, it allows for an examination of speech production that is less confounded by social demands, which is particularly relevant for autism. Finally, compared to other spontaneous tasks such as story retelling, the picture description task places lower demands on working memory, reducing potential cognitive confounds in prosody production (these and other benefits described in Boucher et al., 2022).

Finally, prior research has included limited or no characterization of the motor speech abilities of included participants, complicating the interpretation of prosody-related findings. In the present study, we address this gap by using several approaches to screen for and exclude individuals with apraxia of speech, increasing confidence that group differences in prosody are not associated with a co-occurring primary motor speech disorder.

The Present Study

There were three aims for the present study:

- (1) Characterize the self-reported experiences of prosody in autistic and nonautistic adults;
- (2) Quantify and compare global acoustic features (pitch, rate, pauses) of speech during a spontaneous picture description task between autistic and nonautistic adults; and
- (3) Examine the relationship between self-reported prosody experiences, prosodic acoustic features, autism traits, nonverbal IQ, and language (measured via vocabulary size).

Critically, we also characterized participants' speech motor abilities in order to exclude participants with overt signs of apraxia of speech, in order to clarify that group differences in prosody are not conflated with a co-occurring motor speech disorder.

Materials and Methods

Recruitment

Participants (total $N=84$; 38 NA, 46 AUT) between ages 18-45 years old were recruited to complete the study and consented to participate voluntarily (**Table 3.1**). The final sample was derived from 118 adults that was filtered after data collection using the screening criteria listed below ($n = 15$), and further filtered due to data quality (see details below; $n = 7$) or an insufficient quantity of speech during the key experimental task (see details below; $n = 11$). Finally, one participant recruited in the AUT group was excluded due to signs of a motor speech disorder, as described below.

Prolific

Non-autistic (NA) participants were recruited through Prolific (<https://prolific.com>), a platform facilitating participant recruitment for online research. Prolific made the study available to those meeting the initial criteria (based in the U.S., no existing diagnosis of autism or ADHD, falls within the specified age range, normal vision and hearing, no colorblindness, first language of English, U.S. nationality, no medication use) using existing Prolific records. Researchers then

excluded any participants who did not meet inclusion criteria based on the study's screening questionnaire responses.

SPARK

Autistic (AUT) participants had a clinical diagnosis of autism and were recruited in a multi-step process via the Simons Foundation Powering Autism Research database (SPARK Consortium, 2018). Potential participants were identified by SPARK as eligible based on SPARK database records of a clinical diagnosis of autism, plus the criteria listed below. The SPARK database has been confirmed to have a high degree of diagnostic validity (Fombonne et al., 2022).

Enrollment Process

Potential participants were initially screened using data provided to SPARK or Prolific. The AUT participants who were eligible based on the pre-screening filtering then completed a screening questionnaire that determined their final eligibility for the present study. If eligible, they were invited to continue with other components of the study. To reduce attrition, the NA participants completed the screening questionnaire and the full study at the same time, and were filtered out later based on questionnaire responses.

The research was approved by the MIT Committee On the Use of Humans as Experimental Subjects (COUHES), in accordance with standards in the Declaration of Helsinki. Participants could opt out of the study at any time and were compensated for the portion of the study that they completed. Compensation was not linked to accuracy of task performance or diagnostic group.

Inclusion Criteria

Only participants who self-reported being exposed to English since birth (including bilingual or multilingual individuals) were included. Participants from either group were excluded if they reported any history of head trauma (resulting in concussion or loss of consciousness),

nonfebrile seizures, hearing conditions, color-blindness or other uncorrected vision difficulties, or prematurity (<37 weeks' gestation). For the NA group, participants were excluded if they reported a history of diagnosis of psychiatric or mood disorder, or medications to treat these conditions (including any antipsychotic medication), a diagnosis of a speech, language, or learning disability, or an immediate family member with an autism diagnosis.

In the NA participant group, 15 individuals were excluded following data collection, consistent with the inclusion/exclusion criteria. These individuals reported the following: history of psychiatric diagnosis ($n=11$, includes 1 participant who also reported a history of seizure disorder and prematurity), being non-native English language speakers ($n=1$), family history of autism diagnosis ($n=2$, plus 6 individuals who also had psychiatric diagnosis and were counted there), and history of a fluency disorder ($n = 1$). The consent, screening survey, and all tasks were completed online via participants' personal computers. All participants were located in the United States.

Data Collection Platforms

Participants completed the screening survey via the survey platform Qualtrics (www.qualtrics.com), and were then routed to the Gorilla platform (<https://gorilla.sc/>) to complete the experimental tasks. Upon completion of the final experimental task, they were routed back to Qualtrics to perform the characterization measures.

Characterization Measures

All participants completed the following characterization measures (see **Table 3.1** for a summary).

Matrix Reasoning/Nonverbal IQ

The Test My Brain (TMB) matrix reasoning subtest (Singh et al., 2021) is a validated

measure of nonverbal intellectual ability comparable to other standardized measures of nonverbal IQ. Participants were asked to identify the image that completes the missing **item** in a matrix based on a logical rule. Higher scores indicate better performance or a higher nonverbal IQ. TMB matrix reasoning scores are referenced in this article as “IQ.” Participants were not excluded based on IQ score.

Vocabulary Size Test

The Vocabulary Size Test (VST) (Nation & Belgar, 2007) measures receptive language knowledge in English. Participants were presented with a series of word-sentence pairs (e.g., *See-They saw it*) and four possible definitions. Participants were then asked to choose the definition that most closely matched the meaning of the word as used within the paired sentence. The original VST comprises 10 levels of difficulty, with 10 words per difficulty level. Within the present study, we removed 2 items per level of difficulty to both shorten the measure and make it more appropriate to our participants in the United States. Removed words included those that (1) had the highest discrepancy between British National Corpus values and American corpus values; (2) were based on highly specific knowledge; and/or (3) were marked as “British” in the dictionary. The modified VST used here included 80 total items, with a maximum possible score of 80 (1 point per item).

Comprehensive Autistic Trait Inventory

The Comprehensive Autistic Trait Inventory (CATI) is a self-report survey used to quantify autistic traits and behaviors in the general population (English et al., 2021), with reliability data suggesting (English et al., 2024). Participants were asked to agree or disagree with available response options on questions probing Social Interactions, Communication, Social Camouflage, Self-Regulatory Behaviours, Cognitive (In)Flexibility, and Sensory Sensitivity. Scores on the

CATI can range from 48 to 240, with higher scores indicating more reported autism traits.

Table 3.1.

Participant Characteristics by Group

Characterization	Autistic (AUT) (n=46: 26M; 20F)	Nonautistic (NA) (n=38: 20M; 18F)	t(82)	p-value
	<i>M (SD)</i>	<i>M (SD)</i>		
Age	30.11 (6.29)	31.84 (7.76)	-1.13	.261
CATI	159.46 (22.74)	108.21 (27.27)	9.40	<.001
Test My Brain	27.30 (6.27)	27.32 (4.71)	-0.01	.993
VST	58.54 (8.30)	59.61 (6.62)	-0.64	.525

Note: Student’s t-tests were used to compare statistical significance between groups. All individuals were native English speakers, and some individuals reported being at least conversationally fluent in one or two other languages (AUT: n = 9; NA: n = 6). M=Male (sex assigned at birth); F=Female (sex assigned at birth); CATI=The Comprehensive Autistic Trait Inventory; VST=Vocabulary Size Test.

Prosody Self-Report Measure

The Vocal Output eXperiences Questionnaire (VOX-Q) is a novel 42-item Speech and Listening self-response measure developed by members of this research team (O’Brien et al., 2025). The VOX-Q was developed through two focus group sessions with autistic consultants. Respondents indicate degree of agreement with items pertaining to the perception and production of prosody on a 5-point Likert scale. For more information about this open-access questionnaire, including the full measure, as well as information about subscales and scoring, visit: <https://osf.io/gx8d7/>.

Complete data from this survey including additional subscores will be published separately. In the present paper, we report the “*Composite*” score (16 items used as an overall index of challenges with prosody perception and production) and “*Characterization/ Comparison*” items (3 items that ask participants to compare features of their speech (pauses, rate, pitch) to that of others (e.g., “Compared to others, the pace/rate of my speech is [*1 = a lot slower; 2 = slower; 3 = the same; 4 = faster; 5 = a lot faster*]”).

Speech Production Tasks

Participants completed the following two tasks while their responses were recorded via their computer's microphone. This task was part of a larger project including additional behavioral tasks (prosody speech production, prosody perception) not reported here. To increase likelihood of high-quality audio recordings, participants were instructed to complete the task in a quiet location with no background voices or other noises. Participants were also required to confirm that they would complete the task with headphones and use the Google Chrome browser to increase consistency of recordings.

Diadochokinetic (DDK) Speech Tasks. To characterize participants' ability to rapidly alternate or sequence speech sounds (i.e., oral diadochokinesis), participants completed two traditional measures of oral diadochokinesis: *Sequential Motion Rate (SMR)* & *Alternating Motion Rate (AMR)*. Both of these tasks are sensitive to various motor speech disorders, including apraxia of speech (Josephs et al., 2023). Participants were instructed to repeatedly articulate syllables as quickly and accurately as possible during a 6 second interval. Following instructions and a practice trial with the microphone interface, the task consisted of one 6 second trial for each of the alternating (3, AMR) and sequential (2, SMR) repetition conditions. Participants completed one trial each of the single syllables ("puh," "tuh," and "kuh") and sequential syllable ("puh-tuh-kuh" and "buttercup") conditions. Participants viewed and heard the instruction text for that trial (e.g. Say "puh, puh, puh, puh, puh" as quickly as you can until the screen changes), along with a microphone icon showing recording status on their screen during each trial.

Picture Description Task. Participants provided a one-minute spontaneous speech sample describing a scene-based line drawing ("The Picnic Scene") from the Western Aphasia Battery-Revised (WAB-R; Kertesz, 2006). Participants were instructed to describe as much as they could

about the picture, and to continue to describe the picture for a full minute.

Analysis

Audio files were recorded via participants' computer microphones and downloaded from the Gorilla platform by the researchers at regular intervals during the data collection process (i.e., several times a week), to perform data quality checks and update recruitment goals accordingly.

Measures of Motor Speech Challenges

Three measures were used to characterize participants' speech and screen for signs of an overt motor speech disorder (i.e., apraxia of speech). These included DDK rate (number of correctly produced syllables per second) and clinical judgment of speech during a speech sample. One autistic individual was excluded from the sample based on this combination of factors.

DDK. The number of *correctly* produced syllables per second was calculated for each trial, by participant. A custom Python script identified the first uttered syllable and counted all subsequent syllables within a five second window. The script also produced visualizations with syllable markers and an output file with syllable count. A researcher reviewed visualizations alongside the corresponding audio file in Praat (www.praat.org) to verify accuracy. If the automated counts were correct, no changes were made. If discrepancies were identified, counts were manually updated by adding or subtracting syllables. In some instances, the 5 second audio frame was manually readjusted in Praat for the final syllable count. Mispronounced syllables (e.g., “Tuh” instead of “Kuh”) were excluded. Finally, all audio files were listened to by an independent researcher who coded production errors and confirmed final counts in files that contained these production errors. Any remaining uncertainties were resolved by a trained speech language pathologist. A final rate for each trial was determined by dividing correctly produced syllables by 5 seconds, and average rates for the AMR and SMR tasks were calculated for each participant.

For any participant with an average rate below 4 on either task (SMR only: $n = 3$ [NA]; AMR only: $n = 13$ [AUT $n = 8$; NA $n = 5$]; both: 6 [AUT $n = 2$; NA $n = 4$]). A speech-language pathologist reviewed the audio to assess task completion (i.e., whether the participant attempted syllable repetition for the full duration). All but one participant in the AUT group, and all participants in the NA group, demonstrated task confusion (i.e., only repeated the syllable 3 times), leading to the low rates. For these participants, only the perceived listening effort (described below) was used to evaluate motor speech function. Notably, even including these participants, rates did not differ significantly between groups for AMR ($t(76) = 0.31, p = .76$) or SMR ($t(60) = 1.25, p = .21$).

Speech sample. To complement the DDK rates—particularly given the incomplete DDK data from some participants—a trained speech-language pathologist listened to all participants’ picture description speech samples and assessed whether the speech sample was “Within Normal Limits” with regards to the following features: rate, intelligibility, volume, breath support. Participants with a clinician assessment of speech that was *not* within normal limits *and* low DDK rates (<4 syllables/second on average), were excluded from analyses (AUT: $n = 1$; NA: $n = 0$).

VOX-Q Group Comparisons

Given equal variance between groups, Student’s t -tests were used to compare the *Composite* subscore, as well as production and perception subscores, between groups. Chi-squared tests were used to compare the distributions of Likert scale responses from self-descriptions of perceived rate, pauses, and pitch variability. An ANOVA was used to examine group differences in reported self-descriptions of these three variables.

Audio File Pre-Processing

Researchers listened to each participant’s DDK and picture description audio files and

flagged files that had low audio quality, background noise, or empty/missing data. For the Picture Description task, researchers ensured that files had at least 30 seconds of speaking time (excluding large pauses); if a participant had less speaking time ($n = 8$), they were excluded from the study to ensure sufficient data for analysis. A second researcher reviewed all flagged files to confirm exclusion of files.

For the picture description task, the beginning and end points of the speaking window were each annotated with an interval in a textgrid in Praat (www.praat.org). This speaking window was used in future analysis steps for extracting acoustic measures from the appropriate time window. Researchers then used WhisperTranscribe (<https://www.whispertranscribe.com/>) to generate transcriptions of participants' picture descriptions. Researchers manually reviewed each transcription in conjunction with the audio file to ensure accuracy and added back in any missing filled pause words (e.g., “um,” “uh”) that were present in the original recording.

Acoustic Measure Extraction

Two key acoustic parameters of prosody were extracted from the picture description speech samples and compared between groups: Speech timing and pitch. Pitch and timing were examined because they are directly under speech motor control, there is existing evidence for differences in these features in autism, and differences in these features can have functional impact on communication.

Pauses. An open source Praat script was used to automatically label pauses in a separate textgrid in Praat. The default settings were maintained, such that pauses were defined by a period of at least .39 seconds of silence (<63dB to account for background noise) following the first onset of speech. Accuracy of pause detection was manually checked for a subset (20%) of files to confirm accuracy of pause detection. Information about the pauses (count, overall duration) was

then extracted from the text grid files.

Speaking and Articulatory Rate: The transcriptions (described above), combined with the timestamps generated from Whisper, were used to extract a word count for each narrative. A script in Python was used to obtain the word count, and total length of picture description (time between first and final timestamp of the transcription). Additionally, the length of pauses, calculated in the pause analysis above, was subtracted from total length to identify the total amount of speaking time. Speaking rate (word count/total narrative time, including pauses) and articulatory rates (word count/speaking time not including pauses) were computed.

Pitch. The mean and standard deviation of fundamental frequency (perceived as pitch) were extracted from each file via a script that applied Praat's auto-extraction tool.

Intensity. Intensity—which is also a primary feature of prosody—was not considered as a primary outcome measure here due to potential differences in location of microphone associated with online study setups. Notably, however, we did not find any group differences in intensity, suggesting similar levels of speech audio were analyzed for each group.

Connecting self-report of prosody challenges, acoustic speech features, and autism traits

The relationship between self-reported prosody challenges, acoustic speech features, and autism diagnosis were explored through three analyses.

Logistic Regression Model to Predict Autism Group

A logistic regression model was used to evaluate whether acoustic and self-report features could predict diagnostic group membership. Autism diagnosis (AUT vs. NA) was used as the binary outcome measure, and the following features were included: VOX-Q score, and acoustic measures from the picture description task (rate, articulation rate, sum of pauses, number of pauses, standard deviation of pitch). Mean pitch was not added to the present model given the sex-based differences

in pitch, however, adding mean pitch did not change the primary outcomes of the model. The model was estimated using the *glm()* function in R. Prior to fitting the model, predictors were examined for multicollinearity using variance inflation factors (VIF). All predictors had VIF values well below 5 (range: 1.24-2.19), suggesting acceptable levels of multicollinearity and that all features could be retained in the model.

Linear Model to Predict Prosody Challenges

Next, a linear model was run in R to examine which features predict reported prosody challenges, as measured by the VOX-Q. Here, the score on the VOX-Q was the dependent variable and the following features were used as predictors: diagnosis, CATI score, speaking rate, articulation rate, number of pauses, duration of pauses, standard deviation of pitch, nonverbal IQ, vocabulary score. All features showed low multicollinearity (range: 1.22-2.31) and were retained in the model. We re-ran the model a second time, excluding scores on the CATI, to examine the effects of the other features independently of autism traits measured by the CATI.

Cluster Analysis to Identify Sub-Groups

Recent work in the area of prosody research in autism (e.g., Weed et al., 2023; Liu et al., 2024) has shown the value of using data-driven techniques to explore clusters of participants with similar groupings of prosodic features, irrespective of diagnosis. Here, a k-means clustering approach was used to identify subgroups of participants based on the extracted acoustic features related to each of the variables described above (timing, pitch), self-reported prosody challenges, and other characterization measures.

The following speech-based features from the picture description task were used in the cluster analysis: speaking rate, articulation rate, number of pauses, total duration of pauses, mean pitch, standard deviation of pitch. Additionally, scores from the following measures were included:

VOX-Q, CATI, TMB. All features were z-scored using the *scale()* function in R to ensure equal scaling across each. To determine the optimal number of clusters, the within-cluster sum of squares (WSS) for cluster solutions ranging from 1 to 10 was computed and visualized using an elbow plot. An inflection point at $k = 3$ suggested that three clusters were optimal, which is also consistent with the number of clusters identified in previous work with autistic children and adolescents (Weed et al., 2023; Liu et al., 2024 [preprint]). As such, a k-means clustering with three centers and 25 random starts was conducted using the *kmeans()* function in R. To visualize and interpret the cluster results, a principal component analysis was performed on the standardized features using the *prcomp()* function in R, and the first two principal components were extracted.

Results

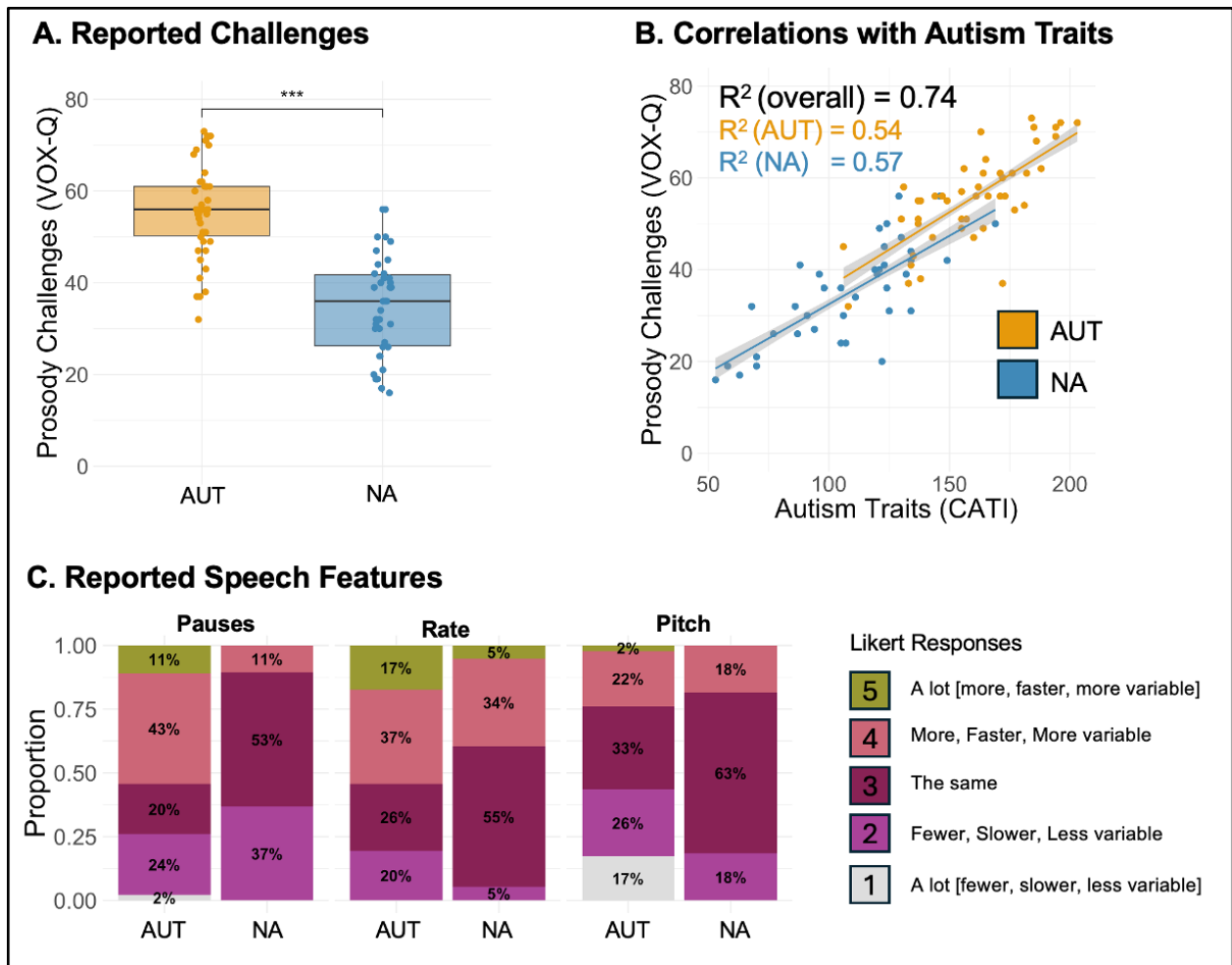
Autistic Adults Reported More Challenges with Prosody

Participants completed the VOX-Q, a self-report questionnaire, which probed their prosody-related experiences. The Composite subscore of the VOX-Q revealed that the AUT group reported greater difficulty with prosody production and perception than NA group ($t(76.6) = 9.01$, $p < .0001$; **Figure 3.1A**). An examination of items related to production and perception separately revealed increased challenges in both areas in the AUT, compared to NA, individuals (p 's $< .001$).

To better understand the relationship between reported prosody challenges as measured by the VOX-Q and autism traits as measured by the CATI, the overall scores of each were compared via Pearson's correlation. Pearson's correlations revealed that the VOX-Q Composite scores were highly correlated with the CATI overall scores across all participants ($r(82) = 0.87$, $p < 0.001$) and in both the AUT ($r(44) = 0.74$, $p < 0.001$) and NA groups ($r(36) = 0.75$, $p < 0.001$; **Figure 3.1B**).

Figure 3.1.

Self-reported prosody challenges and speech features, based on the VOX-Q.



Note: **A.** Autistic adults reported greater difficulty with prosody perception and production, as indicated by significantly higher composite scores on the VOX-Q self-report measure of prosody perception and production. **B.** Composite scores on the VOX-Q (a measure of challenges with production and perception of prosody) were highly correlated with participants' scores on the Comprehensive Autistic Trait Inventory, across all participants and within each group (autistic and nonautistic). **C.** Participants reported how they perceived features of their own speech relative to that of their peers using a five-point Likert Scale. The distribution of responses varied across all features, with NA individuals more often reporting that their features were "the same" as that of others, and AUT individuals more often reporting that the features of their speech was different from others. AUT individuals also used the extreme ends of the scale (i.e., reporting a 1 or 5) more frequently than their NA peers.

Autistic Adults Showed Differences in Their Self-Reported of Prosodic Features

Participants answered several questions asking them to use a Likert Scale to compare various features of their speech. We analyzed responses from the three items that are most relevant to the acoustic analyses in the present study: Pauses, Rate/pace, Pitch. The AUT group demonstrated significant differences in their distribution of responses, compared to the NA group,

as indicated by chi-squared tests and confirmed by fisher tests (Pauses: $\chi^2(4) = 21.30, p < .001$; Rate: $\chi^2(4) = 13.61, p < .01$; Pitch: $\chi^2(4) = 12.79, p < .05$).

More specifically, across all three features, the majority of NA individuals (>50%) indicated that their pauses, rate, and pitch were “The same” as those of others (**Figure 3.1C**). In contrast, only a third or less of individuals in the AUT group indicated that these features were “The same” as those of others. Additionally, more individuals in the AUT group indicated that their speech features aligned with one of the extreme ends of the scale (i.e., 1 or 5) across all features. The NA group on the other hand only indicated an extreme answer for one of the features: 5% of the NA participants indicated that their rate/pace was “A lot faster” than that of others.

When comparing the mean responses for each feature, the only feature with a significant group difference was number of pauses ($F(1)=10.66, p=.002$). The lack of group differences in the other features is likely due to the split in responses at both extremes of the response scale spectrum among the autistic individuals.

Autistic Adults Had Acoustic Differences in Speech Timing and Pitch

Participants completed a spontaneous speech production task during which they described a line drawing scene for one minute. Two key acoustic features of prosody were extracted from each speech sample and compared between groups: Speech timing and pitch.

A one-way ANOVA investigating the length of speaking (i.e., from the speech onset to offset) by diagnosis and sex revealed that the AUT group spoke for marginally less time than the NA group ($F(1, 82)=3.89, p=.052$). There were no significant differences in duration of speaking between males and females. A comparison of overall word count showed that the AUT group used fewer words than the NT group ($t(79, 1)=14.99, p<.001$), and word count did not differ by sex.

Speech Timing Differences

Several timing-related measures were used, including those related to pauses (number of pauses, duration of pauses) and those specific to rate (speaking rate, articulatory rate).

Pauses. There were no group differences in the total number of pauses between groups, however, autistic adults had an overall longer duration of pauses (i.e., increased mean pause length; $t(78.7)=2.43, p=.018$) than the nonautistic adults (**Figure 3.2**). All acoustic results are reported in **Table 3.2**.

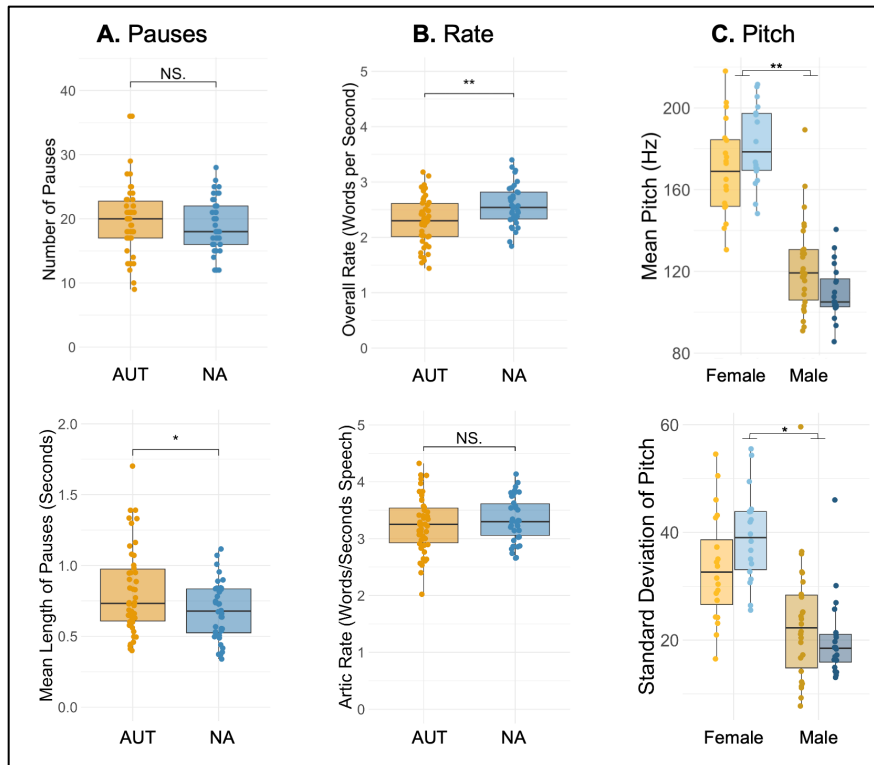
To contextualize the pause data with the cognitive and linguistic characteristics of the participants, the mean pause length was correlated with nonverbal IQ, vocabulary size, and complexity of the narrative (operationalized as type token ratio: number of unique words/number of total words). There was no significant correlation between mean pause length and IQ ($r = .07, p = .52$), suggesting longer pauses were not driven by overall differences in cognitive ability. Mean pause length was also not significantly correlated with the vocabulary size ($r = -.03, p = .79$). Mean pause length was significantly correlated with the participants' type token ratio for their narrative. The groups did not differ significantly on type token ratio, though there was a trend towards greater complexity in the AUT group ($t(81) = 1.92, p = .060$). Across groups, mean pause duration was significantly correlated with type token ratio ($r = .37, p = .0004$), with longer mean pause duration associated with a more complex narrative. In the AUT group, this relationship was significant ($r = .38, p = .009$), while in the NT group it was not ($r = .25, p = .137$). A Fisher's z -test comparing the two independent correlations, however, indicated that the difference in correlations between groups was not statistically significant ($z = 0.66, p = .51$).

Rate. Consistent with the increased pause duration, autistic adults had a lower speaking rate, as measured by words/second ($t(81.9)=-3.31, p=.001$). When examining only time spent speaking, however, autistic adults did not differ in articulatory rate. The same pattern of results

was observed when examining the number of syllables produced per second (**Table 3.2**). This suggests that overall rate differences were driven by increased pauses, rather than reduced rate of articulation.

Figure 3.2.

Primary Results From Acoustic Features Extracted From The Picture Description Task



Note: **A.** The AUT adults do not differ from the NA group in their overall number of pauses, but take longer pauses overall (increased mean length of pauses). **B.** As expected, given the increased overall duration of pauses, the AUT group has an overall decrease in their speaking rate (words per second) compared to the NA group. However, when only considering time spent speaking (total time minus duration of pauses), the groups do not differ significantly in their articulation rate. **C.** Pitch-based analysis showed an interaction between group and sex, such that AUT females had an overall lower mean pitch than NA males, and AUT males and an overall higher mean pitch than NA males. There was also a similar interaction between and group and sex for the standard deviation of pitch: AUT females had a lower standard deviation of pitch than NA females, and AUT males had a higher standard deviation of pitch than NA males.

Autistic individuals show sex-based differences in mean pitch

The mean and standard deviation of each participant's fundamental frequency (f_0), perceived as pitch, was examined by group and by sex. Given the sex-based differences in mean pitch, all pitch analyses also considered the interaction between diagnosis and sex.

Mean pitch. An examination of mean pitch revealed an interaction between diagnosis and sex, such that AUT males had a *higher* mean pitch than NA males and AUT females had a *lower* mean pitch than their NA counterparts ($F(1, 80)=8.14, p=.006$; **Figure 2B**). Given the increased rates of gender diversity in the autism group (Bungert et al., 2024), and potential impacts on pitch differences when examining sex, these analyses were repeated – and results replicated – after removing participants who reported a gender that was not male or female (i.e., gender-queer, nonbinary), who reported being transgender, or who otherwise reported a gender that differed from their sex (AUT: $n=6$; NA: $n=0$; $F(1, 74)=4.37, p=.040$).

Standard deviation of pitch. There was a small significant interaction between diagnosis and sex in standard deviation of pitch, such that while females had a greater overall variation in their pitch compared to males, but the autistic females had a lower standard deviation than NA females, while AUT males had a higher standard deviation than NA males ($F(1, 80)=4.14, p=.045$; **Figure 2B**). As with the mean pitch above, these results remained significant after removing participants who reported a gender that differed from their sex ($F(1,74)=4.36, p =0.040$), suggesting that these effects are not being driven by greater gender diversity in the AUT group.

Table 3.2.

Summary of Acoustic Features Extracted as Measures of Global Prosody From Speech Samples

		AUT	NA	Stat	Sig
		(<i>N</i> = 46)	(<i>N</i> = 38)		
		<i>M (SD)</i>	<i>M (SD)</i>	<i>t</i> (82)	<i>p</i> -value
Pauses	Total Number	19.91 (4.64)	18.82 (4.27)	1.01	.313
	Mean Duration (s)	0.82 (0.31)	0.68 (0.20)	2.34	.022
Rate	Speech rate (words/s)	2.28 (0.44)	2.57 (0.37)	-3.26	.002
	Artic rate (words/s)	3.25 (0.49)	3.34 (0.39)	-0.95	.346

	Speech rate (syl/s)	2.85 (0.51)		3.19 (0.47)	-0.55	.003
	Artic rate (syl/s)	4.07 (0.62)		4.14 (0.59)	-0.53	.598
Duration	Total Duration (s)	55.40 (4.05)		56.91 (2.70)	-1.97	.052
		Male	Female	Male	Female	
		<i>n</i> = 26	<i>n</i> = 20	<i>n</i> = 20	<i>n</i> = 18	
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>		
					<i>F</i> (1, 80)	<i>p</i> -value
Pitch	Mean	123 (3.98)	170 (4.54)	110 (4.54)	182 (4.78)	8.14
	SD	20.2 (2.15)	33.6 (2.15)	23.1 (1.89)	39.2 (2.27)	4.14

Note: AUT = autistic; NA = nonautistic; stat = statistic; sig = significant; M = mean; SD = standard deviation; s = seconds; syl = syllables.

Exploring the Relationship Between Reported Prosody, Speech Acoustics, and Diagnosis

The connections between self-reported prosody challenges, acoustic speech features, and autism diagnosis were explored through three analyses: a general logistic regression model to predict diagnosis from reported challenges and speech features, a linear model to predict perceived prosody challenges from autism traits and speech features, and a data-driven cluster analysis to identify subgroupings based on these features.

Reported Prosody Challenges, Longer Pauses, and Slower Articulatory Rate Predict Autism

A logistic regression model was used to better understand how well the self-report and acoustic speech measures could predict membership in each diagnostic group. The model significantly predicted the diagnostic group ($\chi^2(6) = 67.25, p < .001$). Self-reported prosodic challenges (*VOX-Q*) was a strong predictor of AUT diagnosis, with each 1-point increase on the measure associated with 28% higher odds of an autism diagnosis (OR = 1.28, 95% CI [1.16, 1.49], $p < .001$). Total pause duration also significantly predicted an autism diagnosis (OR = 1.31, 95% CI [1.09, 1.68], $p = .011$), with every one second increase in pauses being associated with 31% higher odds of being in the AUT group. Finally, higher articulation rate was associated with lower

odds of an autism diagnosis ($OR = 0.10$, 95% CI $[0.01, 0.73]$, $p = .038$), with every increase in one word per second associated with a substantial drop in likelihood (90% reduction) of being in the AUT group. Other features, including speech rate and pitch variability, did not significantly contribute to the prediction of diagnosis.

Prediction of Perceived Prosody Challenges

A related but distinct question is how an autism diagnosis, autism traits, and/or speech features predict the degree of reported prosody challenges, as measured by the *VOX-Q*. Here, a linear model was used to predict *VOX-Q* score from diagnosis, autism traits (*CATI*), and speech features. In the full model, the *CATI* score significantly predicted reported prosody challenges ($\beta = 0.32$, $p < .001$), with diagnosis only trending towards significance ($\beta = -21.28$, $p = .052$). However, when *CATI* was removed from the model, diagnosis became a significant predictor ($\beta = -21.28$, $p < .001$), suggesting substantial shared variance between the two variables. Taken together, these results suggested that *CATI* scores may mediate or absorb much of the explanatory power of diagnosis.

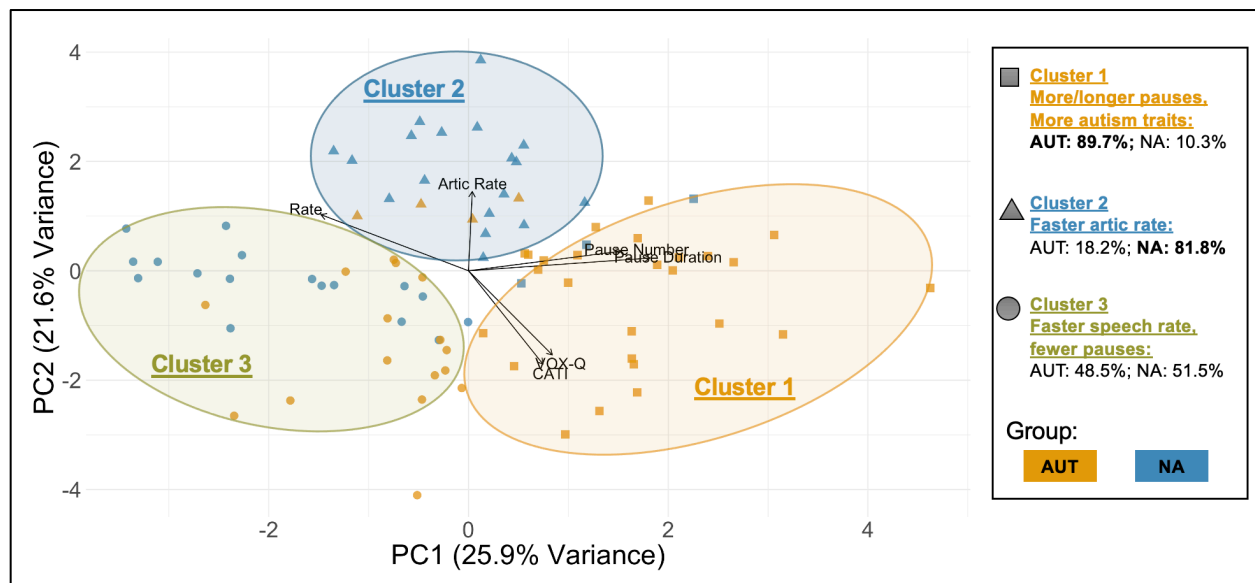
Cluster Analysis

Finally, we examined how individuals may fit into subgroups of participants based on common patterns of acoustic features, irrespective of autism diagnosis. Cluster 1 was composed almost entirely of AUT individuals (89.7% AUT; **Figure 3.3**), Cluster 2 included a majority of NA individuals (18.2% AUT), and Cluster 3 included almost equal numbers of AUT and NA individuals (48.5% AUT). The first principal component (PC1) explained 25.9% of the variance and was driven by the total pause duration, number of pauses, and rate. The second principal component (PC2) explained 21.6% of the variance and was driven most by scores on the *CATI*, *VOX-Q*, and articulation rate. Overall, individuals in Cluster 1 used longer and more pauses and

reported more autism traits and prosody challenges, individuals in Cluster 2 had a faster articulation and speaking rate, and individuals in Cluster 3 used fewer and shorter pauses.

Figure 3.3.

Visualization of the Three Clusters Identified by the K-Means Clustering Approach



Note: The direction of the arrows represent the gradient of increasing values, while the lengths represent the relative contribution to each variable in explaining the variance. Each point represents an individual participant, color-coded by diagnostic group. The shape corresponds to that individuals' assigned cluster. Note: Artic = Articulation; PC1 = Principal Component 1; PC2 = Principal Component 2.

Discussion

This is the first study to date to investigate autistic individuals' self-perceptions of prosody alongside objective acoustic measures of global prosody. Specifically, this study aimed to: (1) quantify autistic adults' subjective experiences with prosody, (2) analyze objective global prosodic features of their spontaneous speech, and (3) examine how these characteristics relate to autism traits/diagnosis, language, and nonverbal IQ. By integrating lived experiences with acoustic data - while confirming that observed prosodic differences were not due to an overt motor speech disorder (e.g., apraxia) - this work offers novel insight into spontaneous speech production in autistic adults, a topic that has received limited attention in prior research.

In relation to subjective experiences, the self-report data from the VOX-Q indicated that autistic adults experience higher levels of challenges with prosody than their nonautistic peers. These self-reported difficulties were highly correlated with autism traits across both groups (AUT and NA). In relation to objective measures of speech production, several global prosodic acoustic features related to timing and pitch showed differences between autistic and nonautistic groups. Specifically, the autistic group exhibited a similar number of pauses between words in their speech, but the pauses were longer and lead to a slower speech (but not articulatory) rate. The autistic adults also showed an attenuation of sex-effects in mean pitch and pitch variability. Finally, three complementary models showed that self-reported challenges and pause duration predicted membership (AUT vs. NA) and that self-reported autism traits (CATI scores) significantly predicted perceived prosody challenges (intriguingly for both groups). Finally, an exploratory clustering analysis identified three subgroups of participants based on both subjective and objective prosody measures: one primarily composed of AUT individuals, one largely of NA individuals, and one with roughly equal representation from both groups. Taken together, these results highlight group differences in experiences with, and production of, prosody, while also suggesting that individual variability in prosody may be better captured through data-driven subgrouping than by diagnostic labels or acoustic measures alone.

Prosody and Autism

Self-Reported Prosody Challenges

The self-reported prosody challenges captured by the VOX-Q offer novel insights into the lived experiences of autistic adults and how these experiences relate to objective acoustic measures of speech. In addition to difficulties, the VOX-Q also captured participants' self-perception of their own speech characteristics. Notably, nonautistic individuals generally reported that their speech

was similar to that of others, whereas autistic individuals were more likely to report that their speech (rate, pitch, pauses) differed in some way. While a detailed analysis of how self-perception aligns with acoustic realization is beyond the scope of the present paper to precisely examine how, one significant group difference emerged: autistic participants more frequently reported the use of *more* pauses than the “average speaker” compared to nonautistic participants. This aligned with the objective data results, showing that autistic adults had longer (though not significantly more) pauses compared to the nonautistic adults. This pair of complementary findings suggests a degree of awareness in both groups regarding the timing of their own speech, a relatively understudied area of prosody that warrants further investigation.

The degree of difficulty with prosody reported by autistic adults was strongly correlated with the extent of autism traits reported, as measured by the CATI (English et al., 2021). This relationship is notable given the lack of prosody-specific items in the CATI and the strength of correlation across both groups (AUT and NA). The finding suggests that prosody-related challenges are closely linked to the broader autism phenotype, as characterized by the CATI, and that this relationship exists on a continuum in the general population. Notably, the composite score of the CATI includes six subscales. Two of these subscales pertain to social communication (“Social Interactions,” “Communication”) and one relates to masking (“Social Camouflage”). The remaining three subscales – “Self-Regulatory Behaviors,” “Cognitive (In)Flexibility,” and “Sensory Sensitivity” – are less obviously related to prosody. Future work should aim to better characterize the relationships between the individual CATI subscales and reported prosody challenges, as well as examine how the VOX-Q scores relate to other used measures of autism trait. Among commonly used autism trait self-report measures for adults (Autism Quotient, Social Responsiveness Scale), prosody is rarely a focus. Finally, the finding that, like other autism traits,

prosody difficulties exist on a continuum in the general population, warrants further investigation to better understand the impact of these challenges, and how to best address them for both autistic and nonautistic individuals.

Acoustic characterization of autistic speech

To the author's knowledge, the acoustic characterization of global prosody presented here is the first analysis of a picture description task in autistic adults. Further, the moderate sample size of participants, and relative balance of males and females, provided the unique opportunity to examine sex-specific effects that may have been missed in prior analyses of autistic speech. In fact, in the largest systematic review of the prosody literature to date, Asghari et al. (2020) *only* included data from male participants: "in the case of studies that reported results according to gender, the results for males were only considered" (p. 3). This exclusion of females from studies is problematic due to the documented sex differences in autism (Wood-Downie et al., 2021), apparent underdiagnosis of autistic females (Cruz et al., 2023), and the particular relevance of sex to features of speech (e.g., pitch).

Timing differences and implications. The metrical pattern (i.e., rhythm) of speech is shaped in large part by the location and duration of pauses between words. Pauses serve as important acoustic cues that mark phrase boundaries, which are thought to be critical to listeners' comprehension (Carlson et al., 2005; Frazier et al., 2006). Pauses that are too long in duration, or that occur in unexpected locations within sentences, can cause speech to be perceived as less fluent, more hesitant, or less coherent (e.g., Sanderman & Collier, 1997). In studies of nonautistic individuals, pauses and rhythm of speech were found to be related to a perceived foreign accent in Foreign Accent Syndrome, highlighting the relevance of pauses in overall perceptions of speech (Katz et al., 2008). Despite the importance of speech timing for communicative efficacy,

spontaneous prosody timing - particularly with respect to pause behavior - has been relatively understudied in autistic adults. A primary finding of the present study was that autistic adults produced *longer* pauses during spontaneous speech compared to nonautistic peers, resulting in a lower overall speech rate.

The present findings are in line with Ochi et al. (2019), who found that autistic adult males spent a greater proportion of time in pauses during a semi-structured interview. Notably, the increased pause duration was correlated with a measure of social reciprocity difficulties (Ochi et al., 2019). By including both autistic males and females and employing tasks beyond social interaction contexts, the present study extends prior findings and highlights the robustness of this prosodic feature across sex and contexts. The present study also offers further context for interpreting differences in pause behavior. First, the mean pause duration was not related to non-verbal IQ or overall vocabulary size, suggesting that longer pauses were likely not related to general cognitive ability or language ability. Second, although mean pause length was positively correlated with type token ratio (linguistic complexity of narrative), the groups did not differ on type token ratio overall and there was no significant difference in the within-group correlation values. Interestingly, across both groups, longer pauses appear related to the production of more complex linguistic narratives. Taken together, these results suggest that the longer pauses observed in autistic adults may reflect differences in language planning, prosodic control, or communication style, rather than cognitive or word-finding challenges.

Additional support for pause-related differences in autism comes from studies focused on children. For example, Thurber & Tager-Flusberg (1993) found that autistic children ($n = 10$) produced fewer non-grammatical (within-phrase) pauses than nonautistic children ($n = 10$), and that the reduction in pauses was related to reduced grammatical complexity of speech. Other

research in autism has explored the use of filled pauses (e.g., “Um,” “Uh”), which serve specific pragmatic functions - such as indicating politeness or signaling new information to the listener (Fox Tree, 2001; Fruehwald, 2016). Several studies have shown a reduction in the use of the filler word “Um” during pauses (Gorman et al., 2016; Irvine et al., 2016; McGregor & Hadden, 2021). In autistic adults, research has shown differences in the prosodic realization – but not frequency – of filled pauses (Wehrle et al., 2024). Although filled pauses were not a focus in the present study, future work should examine the relationship between pause length, pause type (filled vs. silent), and syntactic and prosodic structure in spontaneous speech of autistic adults. This could offer deeper insight into the mechanisms underlying the longer pauses observed here, and their potential impact on listener perception.

The increase in pauses resulted in a slower overall speech rate (words per minute and syllables per minute), which was not accompanied by a slower articulation rate (number of words or syllables per minute after pauses were removed). Prior research examining speech rate in autism is mixed. Patel et al. (2020) reported that autistic adolescents and young adults, as well as their parents, had a slower rate of speech during a narrative production task than nonautistic peers and their parents. Similarly, Bone et al. (2016) showed that speaking rate in autistic children was related to autism traits: autistic individuals with slower speech had more autism traits as measured by the ADOS. In contrast, Nadig & Shaw (2012) did not find differences in speech rate between autistic and nonautistic children during a conversational task. Finally, the disassociation between speech and articulation rate in the present study suggests that when the autistic participants were speaking, they were articulating at a typical rate, but – as described above – they were pausing for longer between words and phrases. The typical rate of articulation rate also suggests that participants did not present with overt motor speech disorders (consistent with our exclusion

criteria), which are typically characterized by reduced rate of articulation (ASHA, 2007). Together, these findings suggest that the observed differences in the present study are unlikely due to articulatory challenges, per se, but rather from increased planning demands - whether motoric, linguistic, or both - during spontaneous speech.

Pitch differences and implications. Prior research investigating global acoustic features of prosody has often been underpowered to examine sex-based differences. In the present study, the relatively balanced sample of male and female participants allowed for the identification of interactions between sex and pitch metrics, adding important nuance to prior findings. Two meta-analyses of prosodic features in autism found that mean pitch was significantly higher in autistic individuals, both across individual studies and in aggregate (Asghari et al., 2020; Fusaroli et al., 2017). Notably, however, the meta-analysis by Asghari et al. (2020) indicated that these analyses only included male participants. Consistent with this, the present study found a higher mean pitch in autistic compared to nonautistic males. In contrast, however, autistic females in the sample had a *lower* mean pitch in autistic compared to nonautistic females. The present study also revealed a reduction in sex-based differences in pitch variability (measured by standard deviation). Specifically, the autistic females showed less variability in their pitch than the nonautistic females, while the autistic males showed more variability in their pitch than the nonautistic females. Again, this refines prior literature: Asghari et al. (2020) reported greater pitch variability in autistic (male) individuals, which is supported in the present study for autistic males, but not autistic females.

Importantly, this overall reduction in “sex-based differences” observed in the nonautistic adults does not appear to be related to increased gender diversity in autism (Bungert et al., 2024): both interactions between pitch metrics and sex remained significant when gender diverse autistic participants ($n = 6$) were excluded from the analyses. While the current data cannot definitively

explain these findings, several possible explanations are proposed. First, although the results remained significant when individuals with discordant sex and gender identity were excluded from an analysis, broader gender expression differences may not be fully captured by this operationalization of gender diversity. That is, participants classified as cisgender may still express gender in ways that diverge from conventional norms. If autistic individuals are less likely to conform to these societal norms, this could result in a reduction in sex-typical patterns. Second, autistic adults may also be less influenced by social feedback about pitch patterns, leading to pitch patterns that are less socially modulated. Third, pitch divergence may reflect conscious or unconscious masking, whereby individuals alter their speech to make it sound more “neutral” or less “atypical,” consistent with reports of camouflaging in other domains of communication (Cook et al., 2022). Future work examining mean and standard deviation of pitch across multiple contexts (i.e., social vs. nonsocial), and in relation to individual differences in masking, may help clarify between these possible interpretations. Overall, these results underscore the importance of recruiting and including autistic females in research to better understand the full autism phenotype and how it may vary by sex.

Multivariate and machine learning approaches

In addition to examining group-based differences in subjective and objective measures, the present study connected these features of prosody to autism diagnosis, autism traits, and potential subgroups through three separate analyses.

First, a logistic regression model was used to explore how well self-report and acoustic speech measures could predict group membership (AUT vs. NA). Self-reported prosody challenges (as measured by the VOX-Q) and increased pause durations were associated with increased odds of being in AUT group, while faster articulatory rates were associated with lower

odds of being in the AUT group. In a second analysis, the degree of self-reported prosody challenges (measured by the VOX-Q) was found to be significantly predicted based on self-reported autism traits (measured by the CATI), but not autism diagnosis (when CATI score was included in the model). This suggests that trait severity, rather than diagnostic status alone, may better explain individual differences in self-reported prosody challenges.

Finally, building on recent efforts in adolescents (Weed et al., 2023; Liu et al., 2024), to identify participant clusters based on acoustic features, irrespective of diagnosis, the current study applied similar methods with the addition of behavioral characterization features (e.g., autism traits, prosody challenges). Consistent with the two prior studies, three clusters were identified: one composed primarily of autistic individuals, one composed primarily of nonautistic individuals, and one with a relatively equal number of autistic and nonautistic individuals. Across all three studies, speech rate emerged as a key factor in accounting for variance between clusters. Taken together, these analyses suggest that prosody challenges and production differences are, on average, more prominent in autistic individuals, but that these differences exist along a continuum, extending into the nonautistic population as well.

Limitations

The results from the present study should be considered within the context of several limitations, specifically the online research setting, the task selection and analysis approach, and the sample of participants.

There are possible limitations of the online research setting. While online research became more common during the COVID-19 pandemic, and may allow for a wider range and larger number of participants in the research, online research presents unique challenges. In this study, participants could self-pace through the instructions and complete the task from a familiar setting,

which might have led to both advantages (e.g., self-pacing, comfort of setting, lack of social interactions with examiner) and disadvantages (e.g., potential environmental distractions, microphone differences, inability to request direct and immediate feedback on task comprehension). Explicit instructions, consistent with recommendations outlined by Newman et al. (2024), practice trials with the interface, and stringent data quality checks were used to minimize the impact of these challenges.

While the picture description task captures participants' spontaneously generated speech, the task is not fully naturalistic, and is a relatively brief speech sample. In particular, describing a picture is not comparable to social speech such as a contingent conversation, and therefore the results may not generalize to social contexts. However, focusing on a task that does not entail explicitly social context could also be viewed as a strength of the study, in that it is more focused on prosody production *without* the potentially confounding impacts of a conversation partner. Additionally, the presence of visual cues provides a more controlled framework across participants for speaking, while also reducing working memory demands necessary for other spontaneous speech tasks (e.g., story retell).

The present study evaluated global prosody (at the narrative level), rather than local prosody (at the individual word or phrase level). This study did not attempt to characterize emotional meaning from the extracted acoustic features, or to evaluate the impact of prosody-related content from the listener's perspective. Follow-up studies that specifically aim to measure these variables across different scales of granularity, or further characterize the speakers' prosody (e.g., ratings), would provide additional insight.

While this study was the first study of its kind, the findings may not generalize to adults with far below nonverbal intellectual ability or vocabulary knowledge and who are unable to

complete the study independently. Further research is needed to characterize prosody in a broader, more diverse sample of the autism spectrum, including individuals with more limited language profiles, as well as investigating the role of sex and gender. While the SPARK database has been shown to have a high degree of validity with respect to autism diagnosis (Fombonne et al., 2022), not all SPARK participants' diagnosis has been confirmed with in-person clinical assessments, and this study did not attempt to additionally confirm the presence of autism symptoms in the autism group, beyond a self-report questionnaire (CATI). Finally, this study did not factor in individual differences such as exposure to languages other than English in our analyses or ask whether participants had lived in the US during childhood and/or adulthood, or consider potential regional differences in US-English speech characteristics, all of which could impact prosody, although these are unlikely to play a role in any primary findings reported here.

Conclusions

Overall, this study suggests differences in both subjective and objective measures of prosody in autistic individuals. Additionally, this study adds the first adult data to a small but growing body of literature suggesting that subgrouping approaches may be informative for understanding autistic prosody, and that rate is a key feature in the distinctions between these subgroups. Further mixed-methods research is needed to characterize additional acoustic parameters of speech across communication contexts, particularly in adults. Examining these acoustic measures in relation to self-report of prosody and broader communication strengths and challenges (e.g. autism traits, learned strategies, masking, gender differences) will better illuminate the role of prosody in autistic social interaction.

References

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Asghari, S. Z., Farashi, S., Bashirian, S., & Jenabi, E. (2021). Distinctive prosodic features of people with autism spectrum disorder: a systematic review and meta-analysis study. *Scientific Reports*, 11(1), 23093. <https://doi.org/10.1038/s41598-021-02487-6>
- Bal, V. H., Kim, S.-H., Fok, M., & Lord, C. (2019). Autism spectrum disorder symptoms from ages 2 to 19 years: Implications for diagnosing adolescents and young adults. *Autism Research*, 12(1), 89–99.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/a:1005653411471>
- Bedford, R., Pickles, A., & Lord, C. (2016). Early gross motor skills predict the subsequent development of language in children with autism spectrum disorder. *Autism Research*, 9(9), 993–1001. <https://doi.org/10.1002/aur.1587>
- Bone, D., Bishop, S., Gupta, R., Lee, S., & Narayanan, S. S. (2016, September). Acoustic-prosodic and turn-taking features in interactions with children with neurodevelopmental disorders. *Interspeech*, 1185-1189.
- Bonneh, Y. S., Levanon, Y., Dean-Pardo, O., Lossos, L., & Adini, Y. (2011). Abnormal speech spectrum and increased pitch variability in young autistic children. *Frontiers in Human Neuroscience*, 4, 237. <https://doi.org/10.3389/fnhum.2010.00237>
- Boucher, J., Brisebois, A., Slegers, A., Courson, M., Désilets-Barnabé, M., Chouinard, A.-M., Gbeglo, V., Marcotte, K., & Brambati, S. M. (2022). Picture description of the western aphasia battery picnic scene: Reference data for the French Canadian population. *American Journal of Speech-Language Pathology*, 31(1), 257–270. https://doi.org/10.1044/2021_AJSLP-20-00388
- Bowers, L., Huisinigh, R., & LoGuidice, C. (2008). *Social Language Development Test Elementary*. East Moline, IL: LinguiSystems.
- Briend, F., David, C., Silleresi, S., Malvy, J., Ferré, S., & Latinus, M. (2023). Voice acoustics allow classifying autism spectrum disorder with high accuracy. *Translational Psychiatry*, 13(1), 250. <https://doi.org/10.1038/s41398-023-02554-8>
- Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, 30(7), 489–518. https://doi.org/10.3109/0269_9206.2016.1145740

- Bungert, C., Li, C., Cardinaux, A., O'Brien, A. M., Cannon, J., Shkolnik, V., Gabrieli, J. D. E., Strang, J., & Sinha, P. (2024). Proportional over-representation of gender diverse identities in two US-based autistic adult samples from the SPARK database. *Autism in Adulthood*. <https://doi.org/10.1089/aut.2023.0121>
- Carlson, R., Hirschberg, J., & Swerts, Marc. (2005). Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech Communication*, 46, 326-333. <https://doi.org/10.1016/j.specom.2005.02.013>
- Constantino, J. N. & Gruber C. P. (2012). *Social Responsiveness Scale–Second Edition (SRS-2)*. Torrance, CA: Western Psychological Services.
- Cook, J., Crane, L., Hull, L., Bourne, L., & Mandy, W. (2022). Self-reported camouflaging behaviours used by autistic adults during everyday social interactions. *Autism*, 26(2), 406–421. <https://doi.org/10.1177/13623613211026754>
- Craig, F., Crippa, A., Ruggiero, M., Rizzato, V., Russo, L., Fanizza, I., Trabacca, A., 2021. Characterization of autism spectrum disorder (ASD) subtypes based on the relationship between motor skills and social communication abilities. *Human Movement Science*, 77, 102802. <https://doi.org/10.1016/j.humov.2021.102802>
- Cruz, S., Zubizarreta, S. C., Costa, A. D., Araújo, R., Martinho, J., Tubío-Fungueiriño, M., Sampaio, A., Cruz, R., Carracedo, A., & Fernández-Prieto, M. (2025). Is there a bias towards males in the diagnosis of autism? A systematic review and meta-analysis. *Neuropsychology Review*, 35(1), 153–176. <https://doi.org/10.1007/s11065-023-09630-2>
- Dahl, K. L., Cádiz, M. D., Zuk, J., Guenther, F. H., & Stepp, C. E. (2024). Controlling pitch for prosody: Sensorimotor adaptation in linguistically meaningful contexts. *Journal of Speech, Language, and Hearing Research*, 67(2), 440-454. https://doi.org/10.1044/2023_JSLHR-23-00460
- de Marchena, A., & Miller, J. (2017). "Frank" presentations as a novel research construct and element of diagnostic decision-making in autism spectrum disorder. *Autism Research*, 10(4), 653–662. <https://doi.org/10.1002/aur.1706>
- Depape, A. M., Chen, A., Hall, G. B., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Frontiers in Psychology*, 3, 72. <https://doi.org/10.3389/fpsyg.2012.00072>
- Diehl, J. J., & Paul, R. (2012). Acoustic Differences In The Imitation Of Prosodic Patterns In Children With Autism Spectrum Disorders. *Research in Autism Spectrum Disorders*, 6(1), 123–134. <https://doi.org/10.1016/j.rasd.2011.03.012>
- Diehl, J.J.& Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics*, 34(1), 135-161. <https://doi.org/10.1017/S0142716411000646>

- Drimalla, H., Scheffer, T., Landwehr, N., Baskow, I., Roepke, S., Behnia, B., & Dziobek, I. (2020). Towards the automatic detection of social biomarkers in autism spectrum disorder: introducing the simulated interaction task (SIT). *NPJ digital medicine*, 3, 25. <https://doi.org/10.1038/s41746-020-0227-5>
- Eni, M., Zigel, Y., Ilan, M., Michaelovski, A., Golan, H. M., Meiri, G., Menashe, I., & Dinstein, I. (2025). Reliably quantifying the severity of social symptoms in children with autism using ASDSpeech. *Translational Psychiatry*, 15(1), 14. <https://doi.org/10.1038/s41398-025-03233-6>
- English, M. C. W., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., Enns, J. T., & Maybery, M. T. (2021). The Comprehensive Autistic Trait Inventory (CATI): development and validation of a new measure of autistic traits in the general population. *Molecular Autism*, 12(1), 37. <https://doi.org/10.1186/s13229-021-00445-7>
- Feldstein, S., Konstantareas, M., Oxman, J., & Webster, C.D. (1982). The chronography of interactions with autistic speakers: An initial report. *Journal of Communication Disorders*, 15, 451–460. [https://doi.org/10.1016/0021-9924\(82\)90018-1](https://doi.org/10.1016/0021-9924(82)90018-1)
- Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, 44(8), 1972–1981. <https://doi.org/10.1007/s10803-014-2073-2>
- Filippa, M., Lima, D., Grandjean, A., Labbé, C., Coll, S. Y., Gentaz, E., & Grandjean, D. M. (2022). Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence. *Scientific Reports*, 12(1), 17144. <https://doi.org/10.1038/s41598-022-21554-0>
- Fombonne, E., Coppola, L., Mastel, S., & O’Roak, B. J. (2022). Validation of autism diagnosis and clinical data in the SPARK cohort. *Journal of Autism and Developmental Disorders*, 52, 3383–3398. <https://doi.org/10.1007/s10803-021-05218-y>
- Fox Tree, J. E. (2001). Listeners’ uses of um and uh in speech comprehension. *Memory & Cognition*, 29(2), 320–326. <https://doi.org/10.3758/bf03194926>
- Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10(6), 244-249. <https://doi.org/10.1016/j.tics.2006.04.002>
- Fruehwald, J. (2016). Filled pause choice as a sociolinguistic variable. In H. Jeoung (Ed.), *University of Pennsylvania Working Papers in Linguistics: Selected Papers from New Ways of Analyzing Variation (NWAV)* (2 ed., Vol. 22, pp. 41-49). Penn Graduate Linguistics Society. <http://repository.upenn.edu/pwpl/vol22/iss2/6>
- Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis". *Autism*, 10(3), 384–407. <https://doi.org/10.1002/aur.1678>

Gernsbacher, M. A., Sauer, E. A., Geye, H. M., Schweigert, E. K., & Hill Goldsmith, H. (2008). Infant and toddler oral- and manual-motor skills predict later speech fluency in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(1), 43–50. <https://doi.org/10.1111/j.1469-7610.2007.01820.x>

Goldman, Eisler F. *Experiments in Spontaneous Speech*. London and New York: Academic Press; 1968. *Psycholinguistics*.

Gorman, K., Olson, L., Hill, A. P., Lunsford, R., Heeman, P. A., & van Santen, J. P. (2016). Uh and um in children with autism spectrum disorders or language impairment. *Autism Research*, 9(8), 854–865. <https://doi.org/10.1002/aur.1578>

Hilton, C. B., & Goldwater, M. B. (2021). Linguistic syncopation: Meter-syntax alignment affects sentence comprehension and sensorimotor synchronization. *Cognition*, 217, Article 104880. <https://doi.org/10.1016/j.cognition.2021.104880>

Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: Revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, 44(8), 1996–2012. <https://doi.org/10.1007/s10803-014-2080-3>

Irvine, C. A., Eigsti, I. M., & Fein, D. A. (2016). Uh, um, and autism: Filler disfluencies as pragmatic markers in adolescents with optimal outcomes from autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 46(3), 1061–1070. <https://doi.org/10.1007/s10803-015-2651-y>

Josephs, K. A., Duffy, J. R., Martin, P. R., Stephens, Y. C., Singh, N. A., Clark, H. M., Botha, H., Lowe, V. J., Whitwell, J. L., & Utianski, R. L. (2023). Acoustic analysis and neuroimaging correlates of diadochokinetic rates in mild-moderate primary progressive apraxia of speech. *Brain and Language*, 240, 105254. <https://doi.org/10.1016/j.bandl.2023.105254>

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2(3), 217–250

Katz, W. F., Garst, D. M., & Levitt, J. (2008). The role of prosody in a case of foreign accent syndrome (FAS). *Clinical Linguistics & Phonetics*, 22(7), 537–566. <https://doi.org/10.1080/02699200802106284>

Kertesz, A. (2006). *Western Aphasia Battery--Revised* (WAB-R) [Database record]. APA PsycTests. <https://doi.org/10.1037/t15168-000>

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings Biological Sciences*, 279(1738), 2698–2704. <https://doi.org/10.1098/rspb.2012.0311>

Krivokapi J. (2007). Prosodic planning: effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, 35(2), 162–179. <https://doi.org/10.1016/j.wocn.2006.04.001>

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA. MIT Press.

- Liu, T., Davison, K. E., Madsen Hardy, I., Kershenbaum, A., Weed, E., Gabrieli, J.D.E., Tager-Flusberg, H., & Zuk, J. (2024, *preprint*). Rethinking prosody in autism: examining prosody production through multiple approaches. <https://osf.io/preprints/osf/78dvn>
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P. S., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- Lyakso, E., Frolova, O., Grigorev, A. (2016). *A Comparison of Acoustic Features of Speech of Typically Developing Children and Children with Autism Spectrum Disorders*. In: Ronzhin, A., Potapova, R., Németh, G. (eds) *Speech and Computer. SPECOM 2016. Lecture Notes in Computer Science()*, vol 9811. Springer, Cham. https://doi.org/10.1007/978-3-319-43958-7_4
- Maffei, M. F., Chenausky, K. V., Gill, S. V., Tager-Flusberg, H., & Green, J. R. (2023). Oromotor skills in autism spectrum disorder: A scoping review. *Autism Research*, 16(5), 879–917. <https://doi.org/10.1002/aur.2923>
- McGregor, K. K., & Hadden, R. R. (2020). Brief report: "Um" fillers distinguish children with and without ASD. *Journal of Autism and Developmental Disorders*, 50(5), 1816–1821. <https://doi.org/10.1007/s10803-018-3736-1>
- Mody, M., Shui, A. M., Nowinski, L. A., Golas, S. B., Ferrone, C., O'Rourke, J. A., & McDougle, C. J. (2017). Communication deficits and the motor system: Exploring patterns of associations in autism spectrum disorder (ASD). *Journal of Autism and Developmental Disorders*, 47(1), 155–162. <https://doi.org/10.1007/s10803-016-2934-y>
- Morett, L. M., O'Hearn, K., Luna, B., & Ghuman, A. S. (2016). Altered gesture and speech production in ASD detract from in-person communicative quality. *Journal of Autism and Developmental Disorders*, 46, 998–1012. <https://doi.org/10.1007/s10803-015-2645-9>
- Nadig, A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders*, 42(4), 499–511. <https://doi.org/10.1007/s10803-011-1264-3>
- Nation, I.S.P. and Beglar, D. (2007) A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology: An International Review*, 70(3), 1380–1402. <https://doi.org/10.1111/apps.12302>
- O'Brien, A. M., Cardinaux, A. L., & Gabrieli, J. D. E. (2025). *The Vocal Output Experiences Questionnaire (VOX-Q): A Self-Report Measure of Prosody-Related Experiences in Adults*. <https://osf.io/gx8d7/>

- Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PloS one*, 14(12), e0225377. <https://doi.org/10.1371/journal.pone.0225377>
- Patel, S. P., Nayar, K., Martin, G. E., Franich, K., Crawford, S., Diehl, J. J., & Losh, M. (2020). An acoustic characterization of prosodic differences in autism spectrum disorder and first-degree relatives. *Journal of Autism and Developmental Disorders*, 50(8), 3032–3045. <https://doi.org/10.1007/s10803-020-04392-9>
- Peppé, S. & McCann, J. (2003). *Profiling Elements of Prosody in Speech-Communication* (PEPS-C). Retrieved from <http://www.peps-c.com/index.html>
- Posar, A., & Visconti, P. (2022). Early motor signs in autism spectrum disorder. *Children (Basel, Switzerland)*, 9(2), 294. <https://doi.org/10.3390/children9020294>
- Rybner, A., Jessen, E. T., Mortensen, M. D., Larsen, S. N., Grossman, R., Bilenberg, N., Cantio, C., Jepsen, J. R. M., Weed, E., Simonsen, A., & Fusaroli, R. (2022). Vocal markers of autism: Assessing the generalizability of machine learning models. *Autism Research*, 15(6), 1018–1030. <https://doi.org/10.1002/aur.2721>
- Sanderman, A. A., & Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech*, 40(4), 391–409. <https://doi.org/10.1177/002383099704000405>
- Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., Sen, A., Singhal, N., Erickson, D., & Singh, N. C. (2010). Sounds of melody--pitch patterns of speech in autism. *Neuroscience Letters*, 478(1), 42–45. <https://doi.org/10.1016/j.neulet.2010.04.066>
- SPARK Consortium. Electronic address: pfeliciano@simonsfoundation.org, & SPARK Consortium (2018). SPARK: A US cohort of 50,000 families to accelerate autism research. *Neuron*, 97(3), 488–493. <https://doi.org/10.1016/j.neuron.2018.01.015>
- Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, 23(2), 309–322. <https://doi.org/10.1007/BF01046222>
- Weed, E., Fusaroli, R., Simmons, E., & Eigsti, I. M. (2024). Different in different ways: A network-analysis approach to voice and prosody in autism spectrum disorder. *Language learning and development: the Official Journal of the Society for Language Development*, 20(1), 40–57. <https://doi.org/10.1080/15475441.2023.2196528>
- Wehrle, S., Grice, M., & Vogeley, K. (2024). Filled pauses produced by autistic adults differ in prosodic realisation, but not rate or lexical type. *Journal of Autism and Developmental Disorders*, 54(7), 2513–2525. <https://doi.org/10.1007/s10803-023-06000-y>
- West K. L. (2019). Infant motor development in autism spectrum disorder: A synthesis and meta-analysis. *Child Development*, 90(6), 2053–2070. <https://doi.org/10.1111/cdev.13086>

Wood-Downie, H., Wong, B., Kovshoff, H., Mandy, W., Hull, L., & Hadwin, J. A. (2021). Sex/gender differences in camouflaging in children and adolescents with autism. *Journal of Autism and Developmental Disorders*, 51(4), 1353–1364. <https://doi.org/10.1007/s10803-020-04615-z>

Zampella, C. J., Wang, L. A. L., Haley, M., Hutchinson, A. G., & de Marchena, A. (2021). Motor skill differences in autism spectrum disorder: A clinically focused review. *Current Psychiatry Reports*, 23(10), 64.

Chapter 4:

Precision fMRI Reveals Engagement of Speech Motor Regions During Prosody Perception and Differences in Autism

Author contributions: **Conceptualization & methodology** (Amanda M. O’Brien; John D. E. Gabrieli; Annie Cardinaux; Stefanie Shattuck-Hufnagel); **data acquisition and curation** (Amanda M. O’Brien; Annie Cardinaux; Cindy Li; Shruti Nishith; Luwei Quan); **statistical analysis** (Amanda M. O’Brien); **interpretation** (Amanda M. O’Brien; Annie Cardinaux; John D. E. Gabrieli); **writing—original draft** (Amanda M. O’Brien; Annie Cardinaux); **writing—revisions** (John D. E. Gabrieli; Cindy Li); **funding** (John D. E. Gabrieli); **supervision** (John D. E. Gabrieli)

Funding: Research reported in this chapter was supported by the Tan Yang Center for Autism Research at MIT. Amanda M. O’Brien was additionally supported by an NIDCD Institutional Training Grant (T32DC000038) and Fellowship (F31DC020864).

Additional acknowledgements: The authors extend their gratitude to Steven Shannon and Atsushi Takahashi for their support in the McGovern Imaging Center, Jennifer Zuk, Frank Guenther, Helen Tager-Flusberg, and Zhenghan Qi for their feedback on various aspects of the study design and interpretation; and to all of the participants who contributed their time to make this work possible. Finally, the authors thank the students who provided valuable contributions across various components of this project, including initial paradigm programming, data collection, data curation, data annotation, and scoring: Karina Artemiev, Charlie Barron, James Broderick, Isabella Chambers, Ines Choi, Diya Dronavadhya, Sabrina He, Samantha Holler, Jeffrey Kwon, Tiffani Mapp, Amber Martin, Eva Naik, Simon Radhakrishnan, Kendall Rhee, Smriti Saini, Laiba Shahid, Timothy Smith, Autumn Rambus, and Kayla Zlotnick.

Abstract

Research in nonautistic individuals, including imaging and lesion studies, has suggested the importance of key motor regions in both the production and perception of linguistic prosody. However, the specific contributions of these regions remain poorly understood in both autistic and nonautistic adults. This chapter presents novel insights from an fMRI study that precisely characterizes the speech production motor regions in autistic and non-autistic adults, and examines their contribution to prosody perception. First, each individual's speech production motor regions were localized. Then, these individualized regions were used to extract activation during a passive linguistic prosody task. This study demonstrates the novel finding that nonautistic individuals engaged many of the same precisely localized regions for speech production as prosody perception, and that autistic individuals showed a reduction in activation of these regions. Moreover, activation of these motor regions during the prosody contrast was correlated with performance on a prosody perception task across and within groups, suggesting communicative relevance of these activation differences.

Introduction

Through subtle changes in pitch, loudness, duration, and/or voice quality, a speaker can transform a single phrase into one that conveys emphasis, urgency, excitement, or sarcasm. These suprasegmental acoustic features being altered by the speaker - collectively known as *prosody* - add richness and nuance to communicative exchanges. For the listener, interpreting prosody requires the rapid integration of acoustic cues to derive meaning that extends beyond the literal content of the spoken words. In the brain, key regions of the speech production motor system have been implicated in the generation of prosody. Emerging research suggests that some of these motor regions may *also* support prosody *perception*. However, it remains unknown whether the same precise motor regions that an individual uses for speech production are also engaged during prosody perception, and how this engagement may be altered for clinical groups with prosody challenges (e.g., autistic individuals). In the present study, we investigate the role of the speech production motor system in prosody perception, by first localizing speech production regions at the individual level, and then examining activation in these regions during a passive linguistic prosody perception task in nonautistic and autistic adults.

Overview of Prosody and Its Development

Prosody serves multiple communicative purposes, which can be categorized as *affective*, conveying emotional information, and *linguistic*, conveying syntactic and/or pragmatic information (Monrad-Krohn, 1947). Here, we focus largely on linguistic prosody, which can be used to convey lexical, syntactical, and semantic information. Linguistic prosody involves the use of three key features: pitch (fundamental frequency), loudness (amplitude of the signal), and duration (length of a given speech sound or phrase). Together, these features can be used to place emphatic stress or focus on individual words in a sentence (e.g., “*Josh* brought the binoculars” vs. “Josh brought the *binoculars*”). Additionally, pauses between words can clarify syntax, and

distinguish between alternative interpretations (“I saw Josh with the binoculars” vs. “I saw Josh, [pause] with the binoculars”). Notably, these acoustic features related to *linguistic* prosody are under active control of the speaker, in contrast to some features of emotional prosody (e.g., voice quality such as breathiness).

Humans are attuned to the melody of speech from an early age. Indeed, even before birth, human fetuses show sensitivity to the melodic contour of music and language (Kisilevsky et al., 2004; DeCasper et al., 1986). Newborns show a preference for the language that they were exposed to prenatally, and can distinguish between languages with varying prosodic structures (Abboub et al., 2016). Additionally, infants show a strong preference for “*infant directed speech*” (also known as *motherese* or *parentese*), which differs from adult-directed speech in its exaggerated prosodic, lexical, and syntactic characteristics (Saint-Georges et al., 2013). A meta-analysis of prosodic aspects of infant-directed speech revealed that prototypical, exaggerated prosodic contours in adult speech were associated with better attentional and communicative outcomes in infants (Spinelli et al., 2017). Notably, the ability to perceive prosodic features of speech undergoes a protracted developmental trajectory, continuing through adolescence (Filippa et al., 2022; Kallay et al., 2022). Some clinical populations, including individuals on the autism spectrum, demonstrate differences in prosody perception (for a review of linguistic prosody, see Grice et al., 2023; for a review of affective prosody, see Zhang et al., 2022) and production (for a review, see Asghari et al., 2021).

Prosody in the Brain

Some of the earliest investigations into the neural bases of prosody began with individuals with acquired brain damage, such as stroke. These early studies focused on the right hemisphere, finding that individuals with unilateral right hemisphere lesions showed difficulties with affective

prosody perception (Ross, 1981; Gorelick & Ross, 1987). Subsequent work, however, also demonstrated the impact of left-lateralized damage on prosody perception (Pell & Baum, 1997). A meta-analysis of lesion studies further clarified that damage to *either* hemisphere may impact prosody perception, but that right hemisphere damage may have a greater impact on *affective* prosody processing (Witteman et al., 2011). Lesion studies have also shed light on the regions related to prosody generation. For example, Sakurai et al. (2015) described a case of Foreign Accent Syndrome⁴ in which an individual with focal damage to the speech production network (e.g., premotor cortex) experienced significant changes to their suprasegmental speech patterns (i.e., prosody). In other cases, individuals have lost a regional accent after damage to similar brain regions (e.g., supplementary motor area, anterior insula).

Beyond lesion-based evidence, neuroimaging research has identified key regions of the brain involved in both the perception and production of prosody, with a large focus on affective prosody. According to the asymmetric sampling in time hypothesis, the encoding of slower time-scale (150-300ms) suprasegmental parameters (such as those relevant to prosody) tend to occur in the right hemisphere, whereas more rapid time-scale changes (<50ms) associated with segmental features of speech are generally processed in the left hemisphere, though this left-ward preference has been shown to shift based experimental design (for an overview, see Oderbolz et al., 2025). Other work probing hemispheric specialization for prosody more specifically has suggested that lateralization at least partially depends on what information (linguistic vs. emotional) the prosodic acoustic signals are conveying (Wildgruber et al., 2006). Aligned with the lesion evidence

⁴ Foreign Accent Syndrome (FAS) is a rare condition first described by French neurologist Pierre Marie (1907). FAS is characterized by the sudden acquisition of a “pseudo-accent,” which resembles a foreign accent, but is characterized by differences in patterns of stress and pauses (Katz, 2008). FAS is thought to result from damage to the speech production motor network in the brain and importantly is not a variant of apraxia of speech (Bhandari, 2011).

described above, activation for emotional prosody processing has been shown to occur bilaterally, with additional activation in Brodmann's Area 47, the amygdala, and the anterior insula (Seydell-Greenwald et al., 2020).

Pertinent to the present study, activation for linguistic prosody has been shown to occur in bilateral fronto-temporal regions (Kreitewolf et al., 2014). A meta-analysis comparing the brain bases of linguistic and affective prosody highlighted common areas of activation during processing of both types of prosody, particularly in right-hemisphere auditory regions and supplementary motor area (Belyk & Brown, 2014). While both linguistic and affective prosody engaged the inferior frontal gyrus (IFG) bilaterally, these functions appeared in segregated regions of the IFG, with linguistic prosody perception occurring in the IFG pars opercularis and affective prosody perception occurring in the IFG pars orbitalis (Belyk et al., 2014). Prosody-relevant white matter pathways have also been identified: A connectivity analysis using probabilistic fiber tracking revealed dorsal and ventral temporofrontal pathways relevant to emotional prosody processing (Fruhholz et al., 2015). As a whole, these studies provide evidence for the engagement of bilateral *sensory, motor, and sensorimotor* regions during prosody perception. Indeed, a dual-stream processing model has been proposed for prosody processing, involving a ventral auditory pathway in the superior temporal lobe and a dorsal auditory-motor pathway that connects posterior temporal regions with inferior frontal and premotor regions (Sammler et al., 2015).

Some research has highlighted the overlap between regions in the brain that support prosody perception and production. Notably few imaging studies have examined the brain bases of prosody *production*, which requires the planning and generation of the prosodic cues and the execution of the voice signal. The limited research that has focused on the *planning and generation* of prosody has identified that the key brain regions involved include the inferior frontal gyrus,

ventral premotor cortex, supplementary motor area (SMA), pre-SMA, and anterior cingulate cortex (reviewed in Guenther et al., 2016). Similar to the prosody perception literature, these studies have focused largely on the production of *affective* prosody (e.g., producing “happy” speech vs. “neutral” speech). The *execution* of prosody, which involves a combination of laryngeal and respiratory movements, is thought to be controlled by a network similar to the one controlling speech articulation. In particular, dorsal and ventral regions of the precentral gyrus, as well as cerebellar lobule VI, are involved in laryngeal and respiratory control (Guenther et al., 2016). Finally, one fMRI study of adults found that prosody perception and production of meaningless phrases shared neural substrates in the premotor regions, including the inferior frontal gyrus, and that activation in these regions was correlated with a measure of prosodic production (Aziz-Zadeh, 2010).

Brain Bases of Prosody in Autism

Autism is a neurodevelopmental condition that is diagnosed based on social and communication challenges, as well as restricted and repetitive interests (APA, 2013). Given the nuanced social role of prosody in conveying communicative information, it is perhaps unsurprising that differences in prosody perception have been broadly noted in autism. The literature on these challenges, however, is mixed. Systematic reviews of the literature have suggested that autistic individuals demonstrate challenges with recognition of emotional prosody and difficulty in interpreting more “intuitive” aspects of linguistic prosody, such as use of prosody for focus in a sentence and intentions, such as conveying irony (for reviews see Fusaroli et al., 2017; Grice et al., 2023; Zhang et al., 2022). Notably, work in this area has largely been conducted with children, with limited attention given to adults.

There have been even fewer neuroimaging studies investigating the neural basis of prosody in autism (four studies in adults and two children), with a primary focus on affective, emotional prosody. The only study to date that used precision fMRI to examine prosody (*only including males*), identified individually-specific voice-specific regions (“voice patches”), and then examined activation in these regions during affective prosody perception of angry, sad, and happy voices (Hashimoto, 2024). The authors found functional differences in key auditory association cortex between groups. Rosenblau (2017) used group-analysis methods to also examine affective prosody in adults, and found activation differences in the superior temporal sulcus for processing of complex vs. basic emotions (with no group differences for the processing emotional vs. neutral prosody). In another study of affective prosody in autistic adults, Gebauer et al. (2014) did not find statistically significant group differences in between autistic ($n = 16$) and nonautistic ($n = 11$) participants for processing emotional speech. On the other hand, Hesling et al. (2010) found significantly reduced activation in the supramarginal gyrus in autistic adults ($n = 8$) compared to nonautistic adults ($n = 8$) when participants listened to a story with exaggerated prosody (affective *and* linguistic).

In work focused on children, Eigsti et al. (2011) found broader, more spatially distributed patterns of brain activation during affective (angry vs. neutral) and pragmatic (question vs. statement) prosody perception in autistic ($n = 16$) compared to nonautistic ($n = 11$) participants. Finally, Wang et al. (2006) used fMRI to examine the use of prosody for irony comprehension in autistic and nonautistic children ($n = 18$ in each group) and found that during the irony condition (compared to rest), the autistic individuals showed stronger activation of inferior frontal gyrus (IFG) and superior temporal gyrus (STS). As a whole, these studies of emotional/affective prosody in autism suggest potential differences in sensory processing regions in autism during emotional

processing. Notably, these studies did not attempt to understand to what extent these sensory regions may also play a function during speech production.

Given the overlap between prosody production and perception regions in nonautistic individuals, an outstanding question is how speech production motor differences in autism may impact prosody perception. Indeed, there have been a few studies to investigate the brain bases of the motor act of speech production in autism; these studies show brain-based differences between autistic and nonautistic individuals during basic and complex oromotor and speech production tasks. A magnetoencephalography study of the neural dynamics during phonemic production tasks revealed significant differences in magnitude and latency of activation in several motor regions between children with autism and NT (Pang et al., 2016). A diffusor tensor imaging study in adults with autism revealed a weaker connection between the ventral premotor cortex and the supplementary motor area (SMA) (Peeva et al., 2013). Other neuroimaging data show decreased engagement of the speech production network - particularly in the SMA - in the autistic compared to the nonautistic group during a speech production task that is highly correlated with language skills, nonword repetition (O'Brien et al, 2022). Taken together, this emerging literature supports a connection between motor skills and communication in autistic individuals, and necessitates a closer look at the connections between prosody and motor differences.

Gaps in the Literature

The present study attempts to address several critical gaps in the literature. First, while existing prosody research has suggested the importance of key speech production motor regions are involved in the *perception* of prosody, no study to date has examined how individually- specific motor regions are engaged during prosody perception. In other words, are the same *precise* regions

that an individual uses to produce speech also active during the perception of prosodically exaggerated speech? If so, which of these motor regions are engaged?

Second, to the authors' knowledge, there have been just three studies examining prosody perception in autistic adults, neither of which specifically investigated linguistic prosody. The present study specifically isolates the use of pitch and duration for linguistic focus (i.e., emphasis). Finally, given the relevance of the motor system in prosody perception and production, it is critical to understand participants' speech production motor functioning, though this has not been examined in prior studies. The present study characterizes speech-based motor function in order to better contextualize the results in a carefully-matched sample of adults (matched on nonverbal IQ and language).

Present study

The present study had three primary aims:

- (1) Investigate whether nonautistic adults engage individually localized speech production system during a passive prosody perception task;
- (2) Examine whether speech production engagement during prosody perception differs in IQ-, language, and gender-matched autistic individuals, and;
- (3) Explore whether this engagement is related to behavioral performance on a prosody perception task outside of the scanner.

The authors pre-registered the following hypotheses:

- (1) Nonautistic individuals will engage some motor regions more when listening to sentences with a dynamic prosody contour than a flat contour.
- (2) Autistic individuals will demonstrate reduced activation of these regions.
- (3) Activation in these regions will be related to performance on a prosody perception task outside

of the scanner. These differences will not be due to apraxia of speech.

Materials and Methods

Study Design

The overall study design included: (1) a neuroimaging session and (2) a behavioral task session. These components were completed in a single visit, separated by a short break.

Participants

A total of 69 adult participants (N = 56: 40 NA, 29 AUT) ages 18-46 years were recruited from the community. Participants were recruited via online social media advertisements, sharing study information on local email lists, and flyers posted in Boston/Cambridge neighborhoods. Autistic adults were also recruited through local organizations that serve this population, and through a database of autistic individuals maintained by the Department of Brain and Cognitive Sciences at Massachusetts Institute of Technology. Participants first completed an online screening survey to indicate interest in the study. Researchers then used survey responses to determine initial eligibility for the study, and to recruit a sample balanced on age and sex assigned at birth.

All participants were fluent speakers of English who had been exposed to English since birth, with no history of prematurity (born >37 weeks gestation), hearing impairment, head trauma (with loss of consciousness or concussion), nonfebrile seizures, or colorblindness. Participants were excluded from the nonautistic (control) group if they were currently taking psychotropic medication for a psychiatric condition (e.g., depression, anxiety), had a history of seizure disorder or epilepsy, language delay, developmental language disorder, specific language impairment, motor impairment, or dysarthria, a diagnosis of dyslexia or ADHD or OCD, or family history of (1st degree relative) schizophrenia/schizoaffective disorder, ASD, PDD-NOS, Asperger's syndrome, or ADHD. **Table 4.1** summarizes participant characteristics on behavioral measures.

After applying exclusionary criteria (described below), 56 participants (34 NA, 22 AUT) were included in the present study (Table 4.1). Seven participants who were recruited in the AUT were excluded for not meeting diagnostic criteria for autism based on the ADOS ($n = 4$), excessive motion in the scanner ($n = 2$), or incomplete data ($n = 1$). Six nonautistic participants were excluded from due to technical difficulties during the scan ($n = 3$) or incomplete data ($n = 3$). A Fisher's exact test was used to confirm no differences in sex ratio between groups ($p = .59$).

Participant Characteristics

Table 4.1.

Group Averages and Standard Deviations of Participant Demographic and Assessment Characteristics

	Autistic (AUT)	Nonautistic (NA)	t(53)	p-value
	n=22: 13 M; 9 F	n=33: 15 M; 18 F		
Measure	Mean (SD)	Mean (SD)		
Age	27.7 (5.0)	29.4 (8.3)	-0.82	.42
KBIT (Standard score)	111.27 (15.4)	113.37 (14.1)	-0.52	.61
CATI	156.00 (21.9)	94.90 (25.3)	9.09	<.001
Language Measures				
CELF Formulated Sentences	47.19 (1.25)	47.65 (0.71)	-1.67	.10
CELF Semantic Relationships	14.90 (2.69)	15.10 (2.45)	-0.27	.79
Prosody Perception Task	11.76 (2.05)	11.67 (1.77)	0.18	.86
VOX-Q	50.41 (9.48)	34.76 (8.95)	6.03	<.001

Note. Demographic, behavioral, and self-reported characterization of participants by group. Group comparisons were conducted using a *t*-test. M = Male; F = Female; NR = Prefer not to report; SD = Standard Deviation; KBIT = Kaufman Brief Intelligence Test (Second Edition); CATI = Comprehensive Autistic Trait Inventory; CELF = Clinical Evaluation of Language Fundamentals; VOX-Q = Vocal Output Experiences Questionnaire.

Diagnostic Confirmation: ADOS Administration

All participants in the AUT group had received a clinical diagnosis of autism prior to participating in the study. Additionally, autistic participants met criteria for autism or autism

spectrum on the Autism Diagnostic Observation Schedule-Revised (ADOS-2) Module 4, on the original and/or the revised algorithm (Lord et al. 2000; Lord et al. 2006; Hus and Lord 2014). The ADOS-2 was administered by a research-reliable member of the research team, typically during a separate visit prior to the fMRI and behavioral study session. ADOS scores using the revised algorithm for the included participants ranged from 8 to 19 ($M = 12.8$, Median = 13, $SD = 3.09$).

Behavioral Measures

Participants completed a comprehensive battery of assessments and questionnaires to characterize their autism traits, cognition, language (including prosody), and speech motor skills.

Nonverbal Intelligence

The Kaufman Brief Intelligence Test, Second Edition (KBIT-2) Matrices Nonverbal subtest assesses fluid nonverbal reasoning. Participants were presented with a series of matrices of pictures or designs, each of which was missing an element from the matrix. Participants were then asked to identify the missing element from a list of options based on the relationship between the extant elements. The relationship between elements within each matrix gradually increases in abstraction and complexity. The KBIT-2 is standardized by age for individuals aged 4 to 90 years, with a mean of 100 and standard deviation of 15. Participants' standard scores were generated.

Autism Traits

The Comprehensive Autistic Trait Inventory (CATI) is a self-report survey used to quantify autistic traits and behaviors in the general population (English et al., 2021; English et al., 2024). Participants were asked to agree or disagree with available response options on questions probing Social Interactions, Communication, Social Camouflage, Self-Regulatory Behaviours, Cognitive (In)Flexibility, and Sensory Sensitivity. Scores on the CATI can range from 48 to 240, with higher scores indicating more reported autism traits.

Language

Two subtests of the Clinical Evaluation of Language Fundamentals, Fifth Edition (CELF-5) were administered to characterize participants' expressive and receptive language abilities: *Formulated Sentences* and *Semantic Relationships*. The *Formulated Sentences* subtest measures participants' ability to produce complete sentences that are semantically and grammatically correct. For each item on this subtest, the experimenter provided participants with one or two words and instructed them to use this word/word pair in a sentence to describe an illustration. Participants' sentences were scored on completeness, semantic relatedness to the target picture, and grammatical correctness. Participants also completed the *Semantic Relationships* subtest, a measure of receptive language that evaluates concepts such as time, location, and serial order. Experimenters presented participants with a statement and asked participants to select two correct answers from a list of four written choices. The CELF-5 is standardized for individuals from 5 to 21 years old; because the participants exceeded this age range, raw scores are reported.

Prosody

Experimental Prosody Perception Task. A custom prosody task was used to assess participants' interpretation of prosodic cues signaling linguistic focus. Participants listened to recorded sentences in which pitch and duration contours were manipulated to place focus on a single word (e.g., "She did not want the SHORT blue cup"). After hearing each sentence, participants were asked to answer a multiple-choice question that required them to use the prosodic cues to distinguish between alternative interpretations (e.g., short blue cup, tall blue cup, short green cup). Two practice trials with feedback were provided before the test items. Response accuracy was calculated across the 15 test items (maximum possible score of 15).

Self-Reported Prosody Experiences. To characterize participants' lived experiences with

prosody, The Vocal Output eXperiences Questionnaire (VOX-Q) was administered via Qualtrics. The VOX-Q is a 42-item Speech and Listening self-response measure developed by members of this research team (O'Brien et al., 2025) through two focus group sessions with autistic consultants. Respondents indicate degree of agreement with items pertaining to the perception and production of prosody on a 5-point Likert scale. For more information about this open-access questionnaire, including to see the full measure, additional subscales, and scoring, visit: <https://osf.io/gx8d7/>. Complete data from this survey including additional subscores will be published separately. In the present paper, we report the “*Composite*” score (16 items used as an overall index of challenges with prosody perception and production).

Motor Speech Challenges

Two measures were used to characterize participants’ motor speech abilities and screen for signs of a primary motor speech disorder (i.e., apraxia of speech). These included DDK rate (number of correctly produced syllables per second) and effortful speech production/listening during a narrative production task.

Oral Diadochokinetic (DDK) Tasks. Alternating Motion Rate (SMR) and Sequential Motion Rate (AMR) tasks were used to examine basic motor abilities of participants’ oral articulators (Kent et al. 1987, Duffy et al., 2005). DDK tasks are commonly used clinically as part of a comprehensive motor speech evaluation, and have been shown to be useful in identifying childhood apraxia of speech (Heller et al., 2015). The AMR tasks require participants to repeatedly produce the same syllable (e.g., “puh”) as quickly and accurately as possible during a short time interval. The SMR tasks require participants to sequence multiple syllables (“puh-tuh-kuh”) as quickly and accurately as possible during a short time interval. In the present study, participants completed three AMR trials (one trial each with “puh,” “tuh,” and “kuh”) and two SMR trials (one

trial each with “puh-tuh-kuh” and “buttercup”). Normative data for these tasks have been well researched: adults between the ages of 18 and 40 produce approximately 6 syllables per second (95% CI: 4.6–7.85; Icht & Ben-David, 2014). The number of *correctly* produced syllables per second was calculated for each trial and averaged across trial type (AMR and SMR) for each participant. No participants from either group were excluded from analyses based on DDK performance.

Picture Description. To obtain a more holistic picture of participants’ spontaneous motor speech skills, a one minute picture description speech sample was elicited. Participants described a scene-based line drawing (“The Picnic Scene”) from the Western Aphasia Battery-Revised (WAB-R; Kertesz, 2006). Participants were instructed to describe as much as they could about the picture for a full minute. If participants stopped describing the picture before 30 seconds had passed, they were prompted to keep going. A trained speech-language pathologist listened to all speech samples from the picture description task and assessed the speech production in these files to ensure the following features were “Within Normal Limits”: Rate, volume, breath support, intelligibility. No participants from either group were excluded from analyses based on these factors.

Neuroimaging Data Collection and Analysis

fMRI Experimental Session

This study was part of a larger imaging project that included two additional fMRI tasks (not reported here). Before entering the scanner, participants were provided with an overview of the tasks that they would complete in the scanner and engaged in a brief practice session including each task. Instructions were reviewed again immediately preceding each task in the scanner. Auditory stimuli were presented binaurally using Sensimetrics S14 earphones (www.sens.com).

Participants' speech was recorded during the speech production localizer task via an MRI-compatible microphone. Stimulus presentation was controlled by the software PsychoPy (Peirce et al., 2019).

Experimental Scanning Tasks

Speech Production Localizer Task. The speech production localizer was used to identify specific speech production regions in individual brains (Fedorenko et al., 2010 for an example). The task was designed to identify regions that are more engaged during speech production (i.e., naming an image aloud) compared to other related, but non-production, perceptual tasks (i.e., hearing a picture being image, looking at an image). A picture naming task was chosen instead of a nonword reading task in order to avoid any potential activation differences associated with reading nonwords in the autism group (Vale et al., 2022).

Conditions. During the speech production localizer task, participants viewed photographs and (1) named the item aloud (Naming), (2) listened to the item being named, or looked at the item in silence, including (3) looking at real, nameable images (Looking) and (4) looking at AI-generated un-nameable images). The critical contrast for the present study was the Naming > Looking at a nameable image contrast, as this maximally activates motor regions, including those that may be active during perception of naming (e.g., areas involved in monitoring auditory feedback during speech production). Participants were randomly assigned to one of 24 possible combinations of stimuli presentation, randomized across order of presented conditions. All participants viewed the set of real pictures in the same order, with the AI-generated images interspersed according to each participant's assigned task condition.

Task Design. Participants saw the task instructions (instructing them whether to name, listen, look at the upcoming image (1s), then two subsequent images (2s each), during which they

completed the directed task condition (named the picture, listened to the picture being named, looked at the picture in silence), and then had a scan (2s). A sparse scanning sequence was used such that scans occurred every 5 seconds and participants viewed and named pictures in the absence of scanning noises (Hall et al., 1999). After completing two blocks of each condition, there was a longer rest block (14s) and scan, during which time the participant was instructed to look at a fixation cross in silence. This sequence occurred eight times, such that there were 32 scans per condition per run. Participants completed 2 runs of this task, and each run was 9 minutes 20 seconds in length (**Figure 4.1A**).

Accuracy. Research staff directly listened to the participants' naming in real-time during the task to ensure participants were completing the task as expected. Clarity on task instructions was provided between runs if needed. Following task completion, audio recordings from the scanning session were reviewed by researchers after the scan to assess accuracy of (1) task completion and (2) naming.

Prosody Perception Task. Participants engaged in a linguistic prosody perception task, during which they passively listened to sentences with flat and dynamic pitch contours. Participants were instructed to press a button when they noticed a break in the sentences (at the start of each rest block), as an attention check. Stimuli were generated to include 160 sentences in the present tense, with an average of 5.99 words and a mean duration of 1.90s for accented speech and 1.82s for flat speech. The sentences were first recorded by a female speaker who was instructed to read the sentence aloud in a natural way, while limiting variation of pitch and duration as much as possible. Each original sentence was then manipulated in two ways using the Praat software in order to create a version of the sentence with a “dynamic” contour and another with a “flat”

contour. A visualization of the sentence manipulations and description of the procedures used are included in **Appendix B**.

Conditions. Participants passively listened to sentences that had been manipulated to have dynamic or flat pitch contours. The critical contrast for the present study was the Dynamic > Flat pitch contour, as this condition isolates the use of pitch and duration for focus. Participants were randomly assigned to one of two possible combinations of stimuli presentation (hearing dynamic sentences first, hearing flat sentence first) and one of two possible orders of stimuli (Set A and Set B), for a total of four possible presentations of the stimuli. The sentences in Set A and Set B were always presented in the same order, regardless of the participant's assigned condition. Sentences presented in each task condition (Set A and Set B) did not differ significantly in the total number of words ($t(158) = 1.225, p = .222$), or duration within accented ($t(158) = 0.825, p = .410$) or flat speech ($t(178) = 0.929, p = .354$).

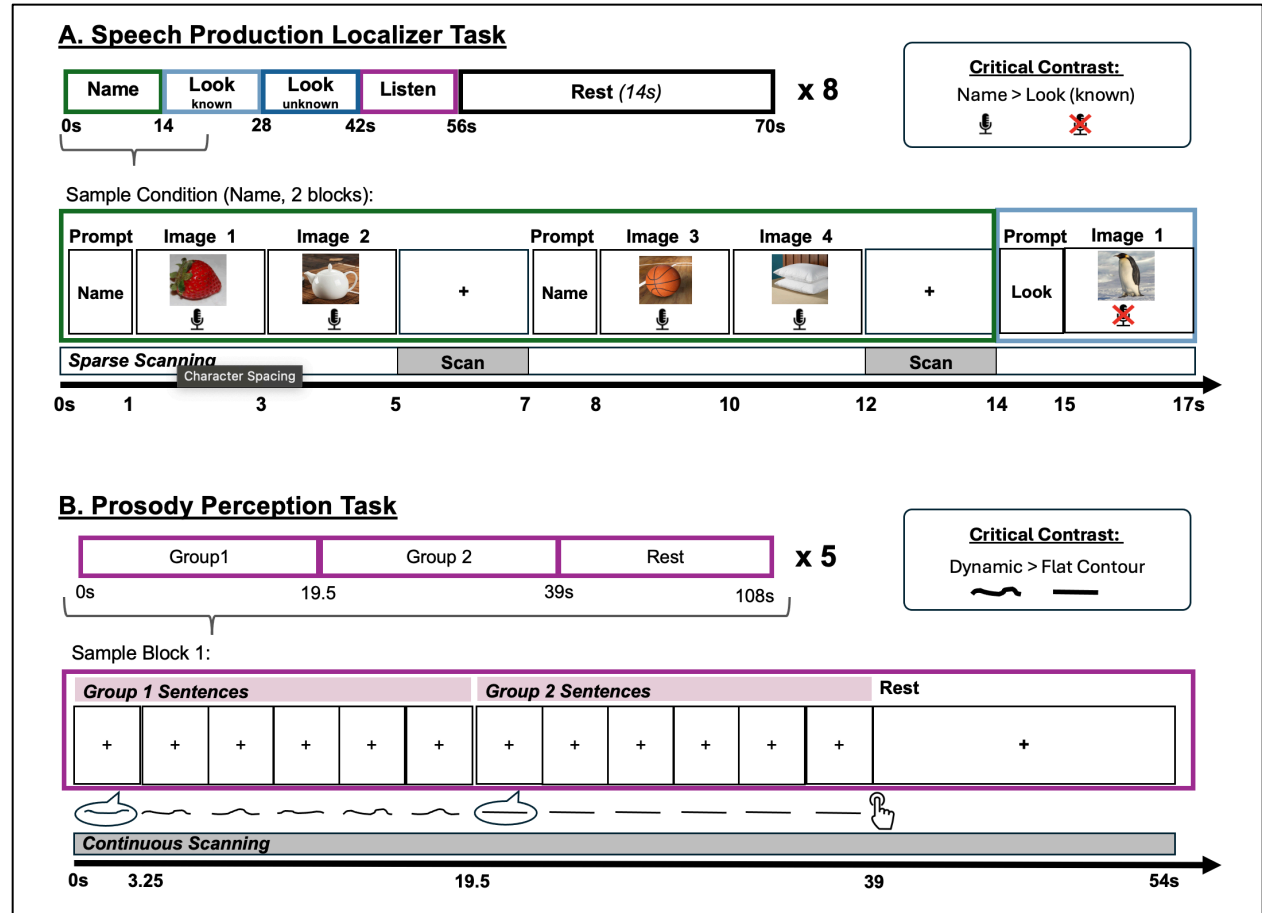
Task Design. Participants heard blocks of sentences that included a group of 6 sentences with dynamic pitch contours and a group of 6 sentences with flat pitch contours. The audio files containing each sentence were 2.5 seconds in length (including 0.2 seconds of initial silent padding, and subsequent silent padding following), and a new sentence began every 3.25 seconds, providing a short buffer between sentences. Each block was followed by a 15s period of rest. This sequence occurred 5 times, such that there were 30 stimuli per condition per run. Participants saw a fixation cross on the screen throughout the study. Participants completed 2 runs of this task and each run was 4.5 minutes in length (**Figure 4.1B**).

Task Performance. A button press task was used as an attention check after each block of sentences, for a total of 5 attention checks. The consistency of button presses during attention

checks was assessed for each participant and compared between groups. Groups did not differ in button presses during the attention checks ($t(81)=1.225, p = .222$).

Figure 4.1.

Task Design and Timing For Speech Production Localizer Task and Passive Prosody Perception Task



Note: Condition order was randomized across participants for both tasks.

Imaging Acquisition

Data were acquired on a Siemen's PRISMA 3T scanner (Siemens AG, Berlin and Munich, Germany) with a 32-channel head coil. A T1-weighted (T1w) volume was acquired at the start of each scanning session (Parameters: TR = 2530 ms; TE 1/2/3/4 = 1.69/3.55/5.41/7.27 ms; voxel size = 1 mm³; slices = 176; FOV = 256 mm³).

T2*-weighted EPI functional scans were collected with a set of parameters specific to each functional task. For the speech production localizer task, the following multi-echo parameters were used: TR = 700 ms, TE 1/2 = 3.96/6.42 ms, voxel size = 2.5 mm³, slices = 55, FOV = 210 mm³. There were 73 volumes per run and 146 volumes per task. A sparse scanning protocol was used to improve dynamic range in auditory brain responses and reduce spurious activation in brain regions associated with the cognitive demands of noise exclusion (Gaab et al., 2007; Gaab et al., 2008; Peelle, 2014). A sparse-sampling rate of TR = 7 s with a 5 s delay was selected to allow participants sufficient time to name the two pictures while still capturing the peak hemodynamic response for each trial (Perrachione & Ghosh, 2013).

For the prosody perception task, the following parameters were used for image acquisition: TR = 2500 ms, TE = 30 ms, voxel size = 2.5 mm³, slices = 56, FOV = 210 mm³. There were 108 volumes per run and 216 volumes per task.

Imaging Data Pre-processing

First, the fMRI DICOM data were converted to NIFTI files following Brain Imaging Data Structure (BIDS) formatting using HeuDiConv (<https://github.com/nipy/heudiconv>; Gorgolewski et al., 2017; Yarkoni et al., 2019). Pre-processing was completed via fMRIPrep (Esteban et al., 2019), including T1 correction, skull stripping, brain extraction, and normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c. Anatomical and functional images were preprocessed using fMRIPrep (v.1.4.1), including T1 correction, brain extraction, normalization to the fsaverage5 template, tissue segmentation, and motion correction procedures [(Esteban et al., 2019); fMRIPrep Available from: <https://doi.org/10.5281/zenodo.852659>, RRID:SCR_016216], a Nipype-based tool (Gorgolewski et al., 2017); Nipype available from: <https://doi.org/10.5281/zenodo.581704>, RRID:SCR_002502).

Framewise displacement (FD; Power et al., 2014), together with other commonly used confounds, such as motion parameters, was calculated for each functional run using the implementation in Nipype. FD and spatial s.d. of the data after temporal differencing (known as ‘DVARs’) are calculated for each functional run, both with their implementations in Nipype (Esteban et al., 2019). Volumes with an $FD \geq 1.0$ mm were regressed out of the analysis (Siegel et al., 2014). Participants with more than 40% of excluded volumes were not included in the analyses (ASD: $n = 2$).

FMRI Analyses

Hypotheses and related analyses for this project were pre-registered on Open Science Framework: https://osf.io/nwrq9/?view_only=1cf4d74a61324069b2ec526d45ac8064

First Level Modeling. For both imaging tasks, all first level modeling was done in SPM12 (<https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>). Normalized and smoothed images were included in a first-level GLM, with task conditions entered as regressors and outliers as nuisance regressors. For the speech production task, the critical contrast was Naming > Looking (Nameable). For the prosody perception task, the critical contrast was Dynamic > Flat pitch contour.

Speech Production Localizer. The 13 parcels generated a priori from the speech production localizer task (**Appendix B**) were used as the regions of interest. For each individual, the top 10% of voxels within each parcel were identified as their functional region of interest (fROI) for that parcel. We confirmed the appropriateness of using these regions across groups by extracting the parameter estimates in these fROIs for each individual for the critical speech production contrast. Parameter estimates were extracted using the *numpy* and *nilearn* packages in

Python. Activation between groups was compared using a linear mixed-effects model predicting activation from group and region of interest, with participant ID included as a random effect.

Prosody Perception Task. Two primary analyses were performed: within participants' speech production functional regions of interest (within and between groups) and at the whole brain level (within group).

Functional Regions of Interest. To investigate whether participants engage speech production motor regions when passively perceiving prosody, the activation from each participants' speech production functional Regions of Interest (fROI) was extracted from the prosody perception contrast (Dynamic > Flat). Parameter estimates were extracted using the *numpy* and *nilearn* packages in Python. For each parcel, a two-sided t-test was used to assess whether the NA activation was significantly different than 0, and *p*-values were adjusted for multiple comparisons using FDR correction. Similarly, a series of two-sided t-tests were used to assess whether the AUT activation in these parcels was different than 0.

To foreshadow results, seven parcels were identified in the NA group as being actively engaged during prosody perception. The activation in these regions was compared between NA and AUT individuals using a linear mixed-effects model predicting activation from group and parcel, with participant ID included as a random effect, with follow-up pairwise comparisons between groups in each parcel.

Whole Brain Analysis. To better understand the landscape of activation within each group during prosody perception beyond that measured in the parcels, and identify any potential compensatory activation in the AUT group that was not captured by the primary analysis, whole-brain group-level modeling was performed in SPM12. Three whole brain analyses were

completed: NA (Dynamic vs. Flat); AUT (Dynamic vs. Flat); NA (Dynamic vs. Flat) vs. AUT (Dynamic vs. Flat).

Brain-behavior comparison. The final goal of the present study was to explore the relationship between activation in motor regions during passive prosody perception with behavioral performance on a prosody perception task outside of the scanner. A Pearson's product-moment correlation was used to compare motor parcels that had significant between group differences. Correlations were calculated across all participants, as well as within each group.

Results

The present study investigated whether nonautistic individuals' speech-production motor brain regions are more engaged when listening to sentences with dynamic - compared to flat - prosodic contours, and whether this activation differs in autism. Each individual's specific speech production motor regions were first localized via a speech production localizer task, and then the activation in these regions during prosody perception was examined.

No group difference in activation in speech production functional regions of interest

Each participants' most active voxels (top 10%) from each of the 13 speech production parcels (SI-1) during the speech production task were identified as their functional regions of interest (fROIs). The magnitude of activation during the critical speech production task contrast (Naming pictures > Looking at pictures) was extracted from these fROIs. A linear mixed effects model revealed that the groups did not differ in their engagement of these regions during the speech production task ($F(1, 134.8)=0.39, p = .53$), suggesting suitability of these regions across groups for subsequent analyses (SI-2). One subtle difference, which did not remain significant after correcting for multiple corrections, was in the left premotor cortex: the AUT group had significantly higher activation in this region than the NA group ($t(33)=2.13, p(unadjusted) = .04$).

Nonautistic adults engage speech production fROIs during prosody perception

After identifying each participants' speech production motor fROIs, a critical and unanswered question is whether nonautistic individuals indeed activated these regions when passively perceiving prosody, consistent with hypotheses. Each participants' activation magnitude from the critical prosody perception contrast (Listening to sentences with dynamic pitch contours > Listening to sentences with flat pitch contours) was extracted from their fROIs. For each parcel, a two-tailed *t*-test was performed to identify parcels that had activation significantly greater than zero (i.e., greater response to dynamic than flat pitch), but also identify any regions with the opposite pattern of activation (i.e., greater response to flat than dynamic pitch). Seven speech production parcels showed significantly positive activation (all corrected *p*'s < .01; **Table 4.2**; **Figure 2A**), suggesting that nonautistic individuals were engaging these specific motor regions when passively perceiving dynamic vs. flat prosodic contours.

Table 4.2.

Activation Difference for the Critical Prosody Contrast (Dynamic > Flat) in Motor Parcels

ROI	NA			AUT		
	<i>t</i>	<i>p-value</i>	<i>p-value (adj)</i>	<i>t-value</i>	<i>p-value</i>	<i>p-value (adj)</i>
ACC	3.88	<.001	.001	0.98	.34	.58
Bilat SMA	4.04	<.001	<.001	0.99	.33	.58
Cingulate	4.16	<.001	<.001	2.09	.05	.32
LH Ant Insula	5.46	<.001	<.001	2.33	.03	.32
LH Cerebellum	1.46	.15	.25	1.21	.24	.58
LH Premotor	-0.57	.57	.58	-0.94	.36	.58
LH STG	-1.21	.24	.34	0.70	.49	.66
LH vPremotor	-0.60	.55	.58	-0.94	.36	.58

RH PT	3.13	.004	.007	0.45	.66	.71
RH Ant Insula	5.55	<.001	<.001	1.74	.10	.42
RH Cerebellum	1.03	.31	.40	0.28	.79	.79
RH Prefrontal	4.20	<.001	<.001	-0.68	.51	.66
RH STG	0.56	.58	.58	0.48	.64	.71

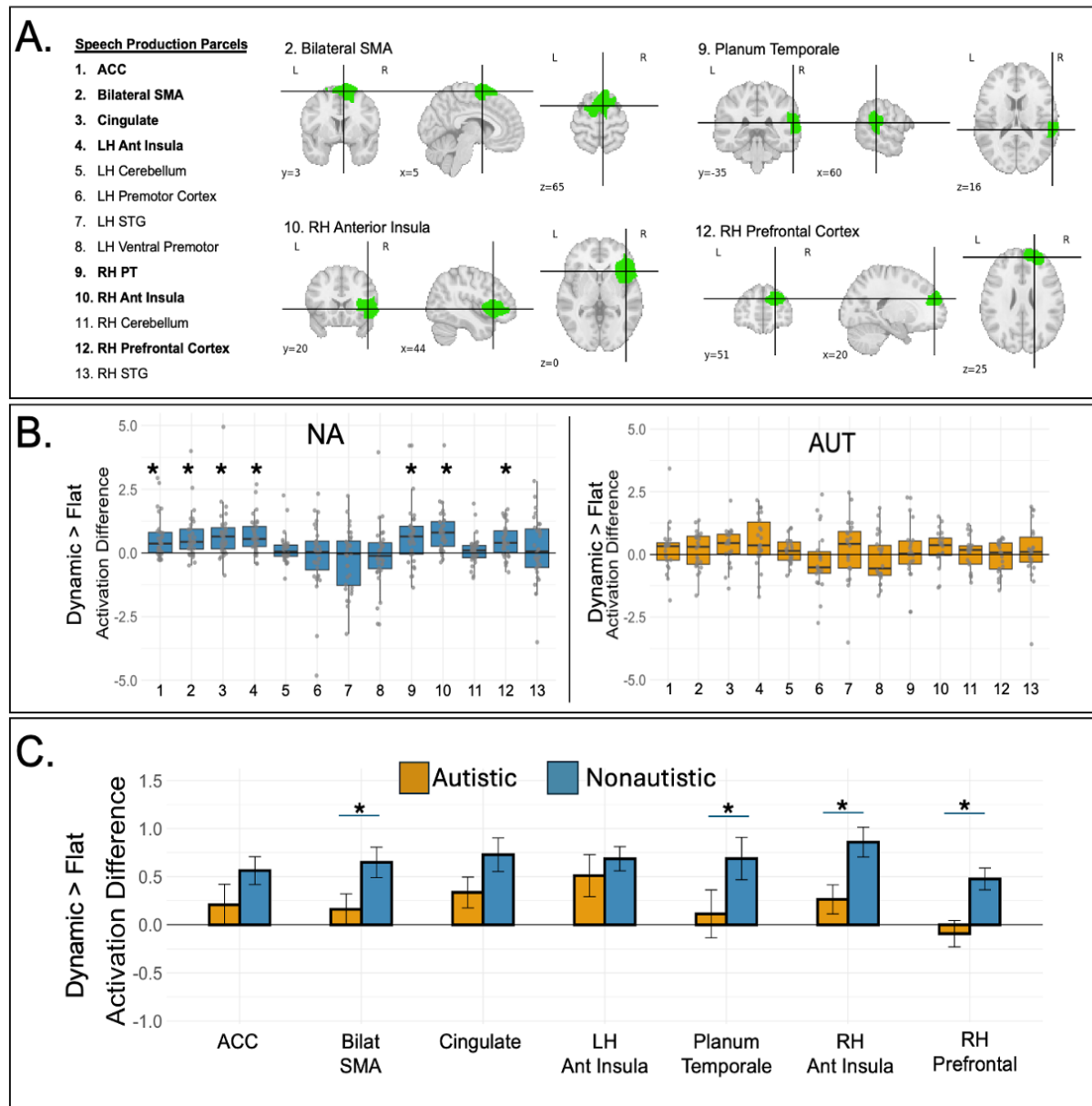
Note: Nonautistic adults showed significantly positive activation in seven individually localized speech production regions of interest during prosody perception, whereas the autistic adults did not show significantly positive activation in any of these regions that survived adjustments for multiple comparisons. NA = Nonautistic; AUT = Autistic; ROI = Region of Interest; LH = Left hemisphere; RH = Right hemisphere; Bilat SMA = Bilateral Supplementary Motor Area; ACC = Anterior Cingulate Cortex; STG = Superior temporal gyrus; PT = Planum Temporale; Ant Insula = Anterior Insula; vPremotor = Ventral Premotor

Autistic adults do not show significant engagement of speech production fROIs during prosody perception

To examine how autistic individuals engage their speech production fROIs during prosody perception, each participants' activation magnitude from the prosody perception task contrast was extracted from their functionally-defined regions. As with the nonautistic adults, two-tailed t-tests were used to examine whether the activation in these regions was significantly different than 0. In contrast to the nonautistic group, only one of the speech production parcels, the LH Anterior Insula, had significantly positive activation, but it did not remain significant after correcting for multiple comparisons (**Table 4.2; Figure 4.2**).

Figure 4.2.

Activation in precisely localized speech production parcels during passive prosody perception.



Note: A. Visualization of key parcels used in analyses. B. NA autistic adults engage seven of the thirteen speech production parcels more during passive prosody perception (listening to sentences with dynamic pitch contours > Listening to sentences with flat pitch contours). In contrast, the NA group did not engage any of these regions more during passive prosody perception. C. The nonautistic adults engage speech production regions more than the autistic adults, with specific group differences in bilateral supplementary motor area (SMA), planum temporale, RH anterior insula, and RH prefrontal regions. ACC = anterior cingulate cortex; Bilat SMA = bilateral supplementary motor area; LH ant insula = left hemisphere anterior insula; RH Ant Insula = right hemisphere anterior insula; RH prefrontal = right hemisphere prefrontal cortex.

Group differences in engagement of motor functional regions

Upon identifying the speech production motor parcels, the nonautistic individuals engaged during prosody perception, a linear mixed effects model was used to predict activation from group and

location. Results indicated that the nonautistic individuals activated these regions significantly more than the autistic individuals during prosody perception (Main effect of group; $F(1, 52) = 5.30, p = .025$; **Figure 4.2**). Pairwise comparison revealed specific group differences for four of the parcels: bilateral SMA, planum temporale, RH anterior insula, RH prefrontal cortex (all p 's < .05).

Whole brain analyses reveal additional group differences in prosody processing

To better understand the full landscape of activation during prosody perception in each group, beyond the speech production motor system, and identify potential compensatory activation in the autistic individuals, a whole brain group analysis was performed for each group individually for the critical prosody contrast, and between groups. The NA group demonstrated significant clusters of activation in expected motor regions, largely occurring in the RH, including the RH premotor cortex and RH anterior insula (**Table 4.3, Figure 4.3**). The AUT group had fewer and smaller regions of significantly positive activation for the critical prosody contrast, including in the RH anterior insula, LH IFG, and RH Cerebellum. A direct comparison between the NA and AUT groups revealed one significant cluster of activation in the middle frontal gyrus that was more active in the NA group than the AUT group for the critical prosody contrast which survived cluster-level corrections; no regions were identified as being more active in the AUT group than the NA group for the critical prosody contrast (**Table 4.3, Figure 4.3**).

Table 4.3.

Regions of Significant Activation for the Critical Prosody Perception Contrast in Whole Brain Analysis

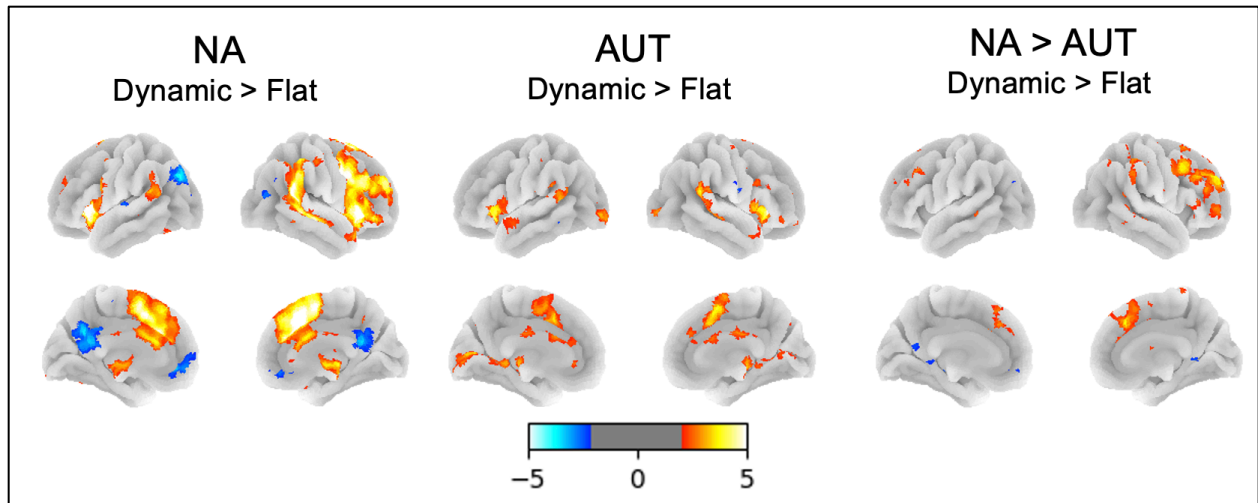
Group	Volume	Peak (t -value)	Cluster-level p (FWE-corr)	Peak (XYZ)	Peak Structure
-------	--------	--------------------	---------------------------------	------------	----------------

NA	4635	7.55	<0.001	42×2×40	RH Premotor Cortex
	1836	6.87	<0.001	36×26×-2	RH Anterior Insula
	3601	6.77	<0.001	12×20×34	Anterior Cingulate
	1695	6.62	<0.001	62×-44×24	RH Temporoparietal Junction
	484	5.28	<0.001	26×56×30	RH Middle Frontal Gyrus
	217	5.26	0.007	40×-54×48	RH Parietal Cortex
	346	4.95	<0.001	8×-24×-12	Substantia Nigra
	252	4.87	0.003	-20×-68×-32	LH Cerebellum (Crus I)
AUT	162	5.73	0.013	36×30×0	RH Anterior Insula
	131	5.68	0.035	-36×30×4	LH Inferior Frontal Gyrus
	121	5.08	0.049	14×-56×-36	RH Cerebellum
NA > AUT	211	4.74	0.009	25×55×30	RH Middle Frontal Gyrus

Note: There were no significant clusters of activation for the AUT > NA comparison.

Figure 4.3.

Whole Brain Analyses for the Critical Prosody Contrast.



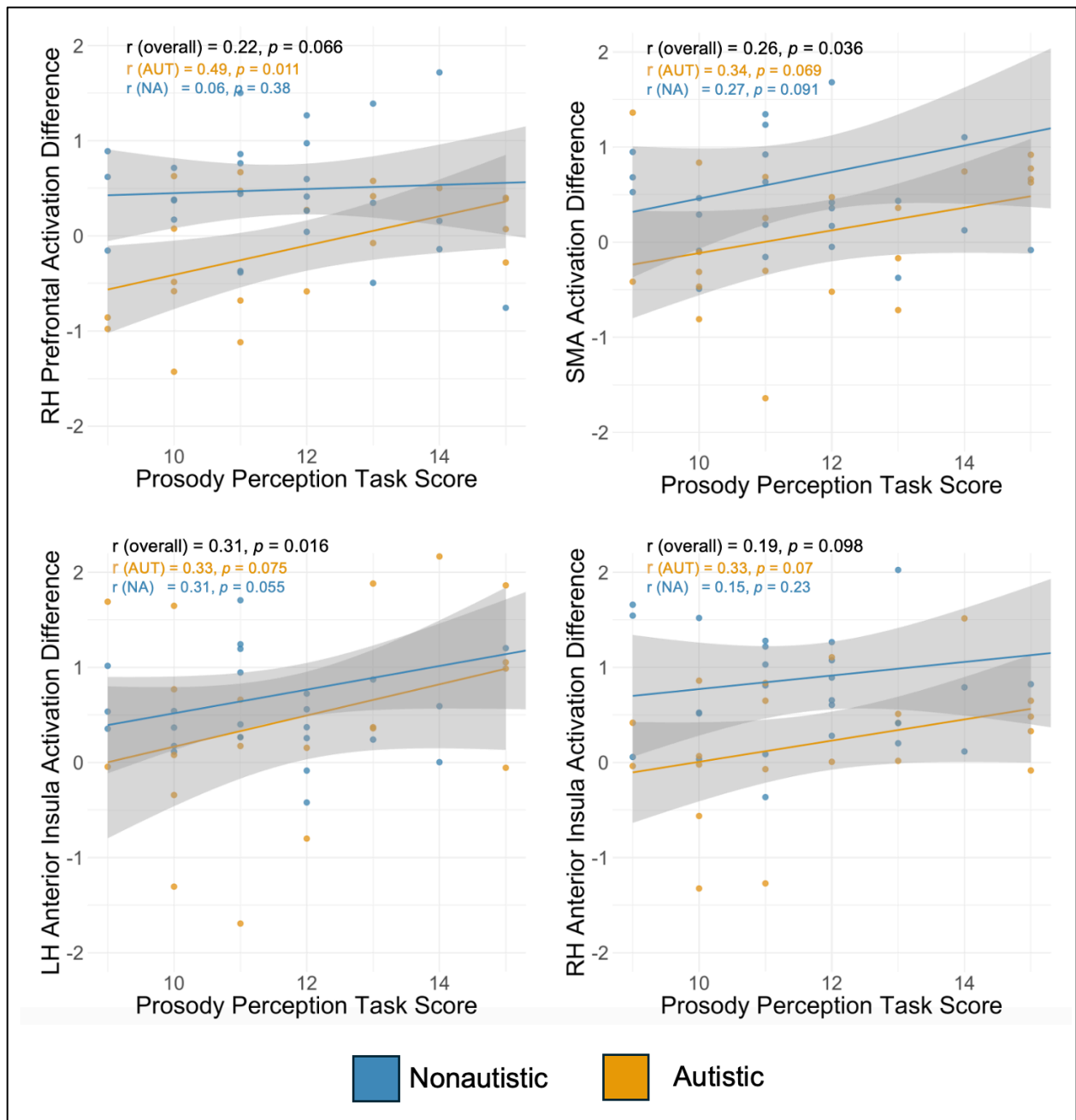
Note: NA = nonautistic; AUT = Autistic

Increased speech production motor activation during prosody perception is associated with better performance on prosody task

A Pearson's product-moment correlation was used to compare motor parcels that had significant between group differences: the right prefrontal, bilateral SMA, and the LH and RH anterior insula. A significant or marginally significant group-based correlation was identified between behavioral performance and activation in all four regions (**Figure 4.4**, p 's < .10). Further evaluation of these correlations by diagnosis revealed significant or trending correlations in each group for two of the parcels: bilateral SMA and LH Anterior Insula (p 's < .10). The other two parcels showed significant brain-behavior correlations for the AUT group but not the NA group.

Figure 4.4.

Correlations Between Brain Activation and Out-of-Scanner Prosody Task



Note: Activation represents the critical prosody contrast (Dynamic > Flat). AUT = Autistic; NA = Nonautistic; RH = Right hemisphere; LH = Left hemisphere; SMA = Supplementary Motor Area

Discussion

To the authors' knowledge, the present study is the the first study to examine the brain basis of altered prosody in autism, and the first to use precision fMRI to examine linguistic prosody perception among nonautistic and autistic individuals. The study addressed three primary aims:

(1) whether nonautistic adults engage individually localized speech production motor regions during a passive prosody perception task, (2) whether this engagement differed in a group of age-, IQ-, language, and sex-matched autistic adults, and (3) to what extent motor engagement related to behavioral performance on an out-of-scanner prosody perception task. The results revealed that nonautistic adults engaged precisely localized motor regions during prosody perception, whereas the autistic group showed reduced engagement of multiple motor regions. Furthermore, motor-region activation during prosody perception was positively associated with performance on a prosody task across and within groups, supporting the idea that motor regions are engaged in prosody perception.

All three primary pre-registered hypotheses were supported. First, nonautistic individuals showed significant activation in over half of their speech production motor regions for the critical prosody contrast (dynamic contour > flat contour). Second, in contrast, the autistic group did not show increased activation for dynamic prosody relative to flat prosody in any of these regions. Third, across both groups – and within each group - activation in motor regions was related to performance on the behavioral prosody perception task. Importantly, these differences are unlikely to reflect a co-occurring speech motor disorder (e.g., apraxia of speech) among the autistic participants, as confirmed by DDK measures and clinical evaluation of speech samples. Together, these findings offer novel insight into the role of the speech production motor network in prosodic language perception in nonautistic adults and contribute to growing evidence linking the motor system to language challenges in autism (O’Brien et al., 2022; Mody et al., 2019).

Prosody in the Brain

While prior *group-level* imaging research on linguistic prosody perception and production has implicated speech production motor regions – such as the left dorsal premotor cortex and IFG

-- in both processes (e.g., Aziz-Zadeh et al., 2010), the present study is the first to directly examine whether the *same*, precisely localized regions are engaged during both prosody perception and speech production at the *individual level*. Indeed, in nonautistic individuals, many of the same speech production regions were more strongly activated during the perception of dynamic prosodic contours than flat ones (i.e., the critical prosody contrast). These regions were distributed across the brain, including the left anterior insula, bilateral SMA, anterior cingulate cortex, cingulate, and several right hemisphere structures such as the right planum temporale, anterior insula, prefrontal cortex.

Many of the regions engaged during both prosody perception and speech production align with areas previously implicated in group-level studies of prosody perception, including both linguistic and affective. For example, Seydell-Greenwald et al. (2020) found bilateral activation of the anterior insula during emotional prosody processing, which was not observed during general sentence processing. Of note, the anterior insula was the only region in the present study to show significant clusters of activation in both the right and left hemisphere in nonautistic adults. Moreover, the right anterior insula was identified in the whole brain analyses as being engaged in both the autistic and nonautistic individuals during prosody perception, suggesting a particular relevance of this structure bilaterally that merits further investigation. Prior work has also identified the importance of bilateral SMA for both affective and linguistic prosody (Belyk & Brown, 2014), and the anterior cingulate cortex for pragmatic and emotional prosody (Doherty et al., 2004; Fruhholz et al., 2012), all regions supported in the present study. Related to prosody production, damage to the SMA has been observed in cases of Foreign Accent Syndrome, believed to be the result of atypical prosody production (Berthier et al., 2015). Together, these findings point to potential common neural substrates supporting both the perception and production of

prosody, across both linguistic and affective domains – highlighting the need for future work that explores how these regions contribute to shared versus distinct mechanisms underlying different aspects of prosodic communication.

Importantly, not all speech production motor regions were significantly engaged during prosody perception. For example, regions that are traditionally associated with earlier auditory processing (e.g., left and right superior temporal gyri [STG]) did not show significantly greater activation for the dynamic vs. flat prosodic contours in the present task. This is in contrast, however, to work suggesting the importance of the STG in processing of pitch accents during connected speech (Gnanateja et al., 2025). Additionally, although whole-brain analyses identified IFG regions as being more active during the dynamic than flat contrast, this was not observed for the individually defined speech production regions. Both of these contradictions suggest that functional regions supporting prosody perception may lie adjacent to - but not overlapping with – those recruited during speech production. This finding adds nuance to prior group-level findings of reported IFG overlap between prosody perception and production (Aziz-Zadeh et al., 2010). The apparent co-localization of production and perception-related activity in areas such as the IFG and STG remains an open question for future investigation.

Brain Bases of Prosody in Autism

While the NA group engaged precisely localized motor regions during prosody perception, the AUT group did not show this same increase in engagement in any speech production motor region of interest. This group difference existed despite similar levels of activation in these regions during the simple speech production naming task (i.e., the speech production localizer), suggesting that observed differences were specific to prosody perception rather than general speech motor functioning. Direct group comparisons revealed significant activation differences across four

regions: Bilateral SMA, right planum temporale, right anterior insula, and right prefrontal cortex. Moreover, across all participants, lower activation in these four regions was associated with poorer performance on a prosody perception task outside of the scanner.

This study contributes to a small but growing body of work investigating prosody processing in autism ($N = 6$ studies) and is the first to explicitly examine linguistic prosody in autism. Prior studies, largely focused emotional prosody, have largely identified regions of the posterior auditory cortex (e.g., Rosenblau et al., 2017) and auditory association cortex (e.g., Hashimoto et al., 2024) as being related to prosody processing challenges in autism. Consistent with these studies, the present study also identified differences in the right planum temporale - a posterior auditory region. Additionally, though falling short of statistical significance, the NA group engaged the ACC during prosody perception while the NA group did not. This aligns with prior work showing that a stronger functional connection between the right superior temporal sulcus (STS) and the ACC predicted prosody recognition in nonautistic, but not autistic, adults (Rosenblau et al., 2017), suggesting a broader functional role for the ACC in prosody perception across domains.

Notably, the present study differed from prior work in several critical ways, offering important context for interpreting cross-study comparisons. First, it employed precision fMRI techniques, allowing for individually localized regions of interest. Of the prior literature, only Hashimoto et al. (2024) also used precision fMRI; however, their analyses were limited to voice-specific processing regions, restricting their ability to examine motor-specific contributions. Other studies of prosody in autism have utilized group-averaging based methods, which may obscure meaningful individual variability (Fedorenko et al., 2010) – particularly in clinical populations where functional organization may be more variable (O’Brien et al., *In Prep*). Finally, the present

study is the first to focus on linguistic prosody in autism. While prior work in nonautistic individuals has shown similarities between common regions of activation for both affective and linguistic prosody (e.g., Belyk & Brown, 2014), this overlap has not yet been examined in autism. By focusing on linguistic rather than affective prosody, the present study minimizes potential confounds related to differences in social-emotional processing in autism (Lartseva et al., 2015) that may influence performance on affective prosody tasks. Taken together, the present study provides unique insight into the motor system's involvement in prosody perception at the individual level.

Among the regions implicated in the present study, the supplementary motor area (SMA) – including both pre-SMA and SMA proper – may be particularly informative target for future investigation in both speech production and perception in autism. The SMA is known for its role in initiating, planning, and controlling motor and speech actions (Guenther, 2016), and an intracranial EEG study highlighted the SMA's importance in speech initiation (Bullock et al., 2024). Extending this line of evidence to speech production in autism, an investigation of the relationship between white matter and speech timing in autistic children found that shorter pauses in autistic children was associated with higher fractional anisotropy in the frontal aslant tract, connecting the SMA and IFG (Davison et al., *in press*). This finding relates prosody production differences in autism to brain differences in motor control. Beyond production and speech timing – and of particular relevance here – the SMA has also been shown to play an important role in non-production language tasks in nonautistic individuals, particularly as task demands become more challenging (reviewed in Hertrich et al., 2016). The SMA has been implicated in many aspects of language processing in nonautistic individuals, including consolidation of lexical knowledge (Lesage et al., 2015) and semantic judgements (Moore-Parks et al., 2010). In the field

of autism research, emerging evidence also suggests that atypical SMA activation may contribute to language differences; for example, O’Brien et al., 2023 found reduced SMA activation during nonword repetition in autistic compared to nonautistic children. Altogether, these findings point to the SMA as a core region at the intersection of motor and language systems, particularly as related to differences in autism.

Functional Impact of Prosody Differences in the Brain

Consistent with the group-level differences in motor activation during prosody perception, autistic and nonautistic adults also differed in their reported prosodic challenges. On the VOX-Q, autistic adults reported greater difficulties with both prosody perception and production in everyday communication than nonautistic adults. To further assess the real-world relevance of the brain findings, participants completed a prosody perception task outside the scanner, in which they listened to sentences with manipulated pitch contours that disambiguated meaning and selected the appropriate interpretation. Although the groups did not differ significantly in their performance on this task, activation in motor regions during the in-scanner prosody task was strongly correlated with accuracy on the out-of-scanner task within and across both groups. This suggests that motor engagement during prosodic processing is meaningfully linked to perceptual accuracy, highlighting a functional impact of these brain differences. Moreover, the presence of individual variation in motor activation and prosody performance across both groups supports a dimensional – rather than strictly categorical – account of prosody processing differences.

Motor System and Language

There is a long history of research attempting to characterize the relationship between the speech motor system and language processing. The Motor Theory of Speech Perception (Lieberman et al., 1967) posited that speech perception *depends* upon motor representations of

articulation. This theory has been disproven in its purest form: lesions to frontal speech production areas (e.g., as in the case of Broca’s aphasia) do not entirely impair speech perception, meaning that the motor system is not *critically necessary* for speech perception (Hickock et al., 2011). “Weaker” or more nuanced versions of the motor speech theories have been put forth and gained traction. One such model, analysis-by-synthesis, aligned with contemporary Bayesian frameworks, proposes that the motor system supports speech processing by creating production-based constraints on the perception system (Poeppel & Monahan, 2010; Skipper et al., 2005).

As is demonstrated in the present study, some cortical areas associated with speech production are also active during speech perception (see also: Fedorenko, Ivanova & Regev, 2024; Skipper et al., 2015). This effect is particularly strong in naturalistic, audio-visual speech contexts (e.g., Skipper et al., 2007) but has also been shown in audio-only contexts, particularly during non-optimal listening conditions (e.g., Park, 2020; Du et al., 2014), during conversation (e.g., Scott et al., 2009), and when processing abstract language, particularly in a second language (Tian et al., 2020). A meta-analysis using activation likelihood estimation of overlap between production and perception identified both context-independent regions (e.g., those that are consistently engaged during linguistic processing) such as the left planum temporale, that are activated across a full range of language perception and production and also context-dependent regions such as the pre-SMA and posterior lateral IFG, that are recruited in more specific situations (Perron et al., 2024).

Data from both the individual- and group-level data in the present study underscore the importance of distinguishing between motor regions that are specifically engaged at the individual level for both production and perception (e.g., the SMA) and those that are engaged at the group-level, but may not represent the same neural substrates at the individual level (e.g., IFG), to further refine these motor-based theories. Notably, the role that these motor regions play during language

processing in autism has received limited attention. Given the current findings – and the behavioral link between motor skills and language abilities in autism (e.g., Bal et al., 2020) – further research in this area is warranted.

Limitations

The results of the present study should be interpreted in the context of several limitations, primarily those related to the in-scanner prosody stimuli and the generalizability of findings across the autism spectrum.

First, the in-scanner prosody task was developed specifically for this study, with the critical prosody contrast designed to capture the processing associated with a *linguistic* pitch accents (i.e., dynamic prosodic contours). To generate these stimuli, the speaker was instructed to speak naturally, but with minimal pitch variation. These recordings were then acoustically modified to create the dynamic and flat versions of the stimuli. However, in attempting to suppress pitch variation, the speaker may have unconsciously increased their use of other prosodic cues, like pauses or volume had prosodic – a phenomenon known as “cue-trading” (Lieberman, 1960). These timing or intensity features were not modified in subsequent manipulation, so may be present in the “flat” condition. That said, these features would also remain in the “dynamic” condition and thus should be accounted for in the contrast. Additionally, while the focus of the present study was on linguistic prosody, the manipulated pitch contours may have inadvertently conveyed emotional tone (e.g., the flat contour may be interpreted as sounding sadder than the dynamic one). In this way, it is difficult to fully isolate linguistic prosody from affective prosody.

Second, the results should be interpreted in relation to the phenotypic profiles of the included sample. The autistic and nonautistic participants were matched on measures of language ability and nonverbal IQ. As such, these results may not generalize to individuals with different

language or cognitive profiles. It is notable, however, that even among these matched participants, robust differences in brain activation were observed.

Conclusions

Overall, the present study found evidence that individually localized speech motor regions are engaged during prosody perception in nonautistic adults, but not in autistic adults who are matched on language and nonverbal IQ. Notably, activation in these regions was correlated with performance accuracy on an out-of-scanner prosody task. The results add to an emerging body of literature indicating differences in motor system engagement in autism. Future work should examine the specificity of these differences to autism, whether similar motor differences underlie other prosody challenges (e.g., with emotional prosody), as well as the developmental emergence of these differences in autism.

References

- Aziz-Zadeh, L., Sheng, T., & Gheytaichi, A. (2010). Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS ONE*, 5(1). <https://doi.org/10.1371/journal.pone.0008759>
- Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, 162, 46–59. <https://doi.org/10.1016/j.bandl.2016.08.002>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Asghari, S. Z., Farashi, S., Bashirian, S., & Jenabi, E. (2021). Distinctive prosodic features of people with autism spectrum disorder: a systematic review and meta-analysis study. *Scientific Reports*, 11(1), 23093. <https://doi.org/10.1038/s41598-021-02487-6>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/a:1005653411471>
- Belyk, M., & Brown, S. (2014). Perception of affective and linguistic prosody: An ALE meta-analysis of neuroimaging studies. *Social Cognitive and Affective Neuroscience*, 9(9), 1395–1403. <https://doi.org/10.1093/scan/nst124>
- Berthier, M. L., Dávila, G., Moreno-Torres, I., Beltrán-Corbellini, Á., Santana-Moreno, D., Roé-Vellvé, N., Thurnhofer-Hemsi, K., Torres-Prioris, M. J., Massone, M. I., & Ruiz-Cruces, R. (2015). Loss of regional accent after damage to the speech production network. *Frontiers in Human Neuroscience*, 9, 610. <https://doi.org/10.3389/fnhum.2015.00610>
- Bhandari, H. S. (2011). Transient foreign accent syndrome. *BMJ Case Reports*. <https://doi.org/10.1136/bcr.07.2011.4466>
- Bullock, L., Forseth, K. J., Woolnough, O., Rollo, P. S., & Tandon, N. (2024). Supplementary motor area in speech initiation: A large-scale intracranial EEG evaluation of stereotyped word articulation. *iScience*, 28(1), 111531. <https://doi.org/10.1016/j.isci.2024.111531>
- Carlson, Rolf & Hirschberg, Julia & Swerts, Marc. (2005). Cues to upcoming Swedish prosodic boundaries: Subjective judgment studies and acoustic correlates. *Speech Communication*. 46. 326-333. <https://doi.org/10.1016/j.specom.2005.02.013>
- Craig, F., Crippa, A., Ruggiero, M., Rizzato, V., Russo, L., Fanizza, I., Trabacca, A., 2021. Characterization of Autism Spectrum Disorder (ASD) subtypes based on the relationship between motor skills and social communication abilities. *Human Movement Science*, 77, 102802. <https://doi.org/10.1016/j.humov.2021.102802>

- DeCasper, A.J., and Spence, M.J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavioral Development*, 9, 133–150. [https://doi.org/10.1016/0163-6383\(86\)90025-1](https://doi.org/10.1016/0163-6383(86)90025-1)
- Depape, A. M., Chen, A., Hall, G. B., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Frontiers in Psychology*, 3, 72. <https://doi.org/10.3389/fpsyg.2012.00072>
- English, M. C. W., Gignac, G. E., Visser, T. A. W., Whitehouse, A. J. O., Enns, J. T., & Maybery, M. T. (2021). The Comprehensive Autistic Trait Inventory (CATI): development and validation of a new measure of autistic traits in the general population. *Molecular Autism*, 12(1), 37. <https://doi.org/10.1186/s13229-021-00445-7>
- Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194. <https://doi.org/10.1152/jn.00032.2010>
- Fedorenko, E., Ivanova, A. A., & Regev, T. I. (2024). The language network as a natural kind within the broader landscape of the human brain. *Nature Reviews: Neuroscience*, 25(5), 289–312. <https://doi.org/10.1038/s41583-024-00802-4>
- Filippa, M., Lima, D., Grandjean, A., Labbé, C., Coll, S. Y., Gentaz, E., & Grandjean, D. M. (2022). Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence. *Scientific Reports*, 12(1). <https://doi.org/10.1038/s41598-022-21554-0>
- Frühholz, S., Ceravolo, L., & Grandjean, D. (2012). Specific brain networks during explicit and implicit decoding of emotional prosody. *Cerebral cortex*, 22(5), 1107–1117. <https://doi.org/10.1093/cercor/bhr184>
- Frühholz, S., Gschwind, M., & Grandjean, D. (2015). Bilateral dorsal and ventral fiber pathways for the processing of affective prosody identified by probabilistic fiber tracking. *NeuroImage*, 109, 27–34. <https://doi.org/10.1016/j.neuroimage.2015.01.016>
- Fruehwald, J. (2016). Filled pause choice as a sociolinguistic variable. In H. Jeoung (Ed.), *University of Pennsylvania Working Papers in Linguistics: Selected Papers from New Ways of Analyzing Variation (NWAY)* (2 ed., Vol. 22, pp. 41–49). Penn Graduate Linguistics Society. <http://repository.upenn.edu/pwpl/vol22/iss2/6>
- Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis. *Autism Research*, 10(3), 384–407. <https://doi.org/10.1002/aur.1678>
- Gaab, N., Gabrieli, J. D., & Glover, G. H. (2007). Assessing the influence of scanner background noise on auditory processing. I. An fMRI study comparing three experimental designs with varying degrees of scanner noise. *Human Brain Mapping*, 28(8), 703–720. <https://doi.org/10.1002/hbm.20298>

Goldman, Eisler F. (1968) *Experiments in Spontaneous Speech*. London and New York: Academic Press; 1968.

Gorelick, P. B., & Ross, E. D. (1987). The aprosodias: further functional-anatomical evidence for the organisation of affective language in the right hemisphere. *Journal of Neurology, neurosurgery, and psychiatry*, 50(5), 553–560. <https://doi.org/10.1136/jnnp.50.5.553>

Grice, M., Wehrle, S., Krüger, M., Spaniol, M., Cangemi, F., & Vogeley, K. (2023). Linguistic prosody in autism spectrum disorder—an overview. *Language and Linguistics Compass*, 17(5). <https://doi.org/10.1111/lnc3.12498>

Guenther, F. H. (2016). *Neural Control of Speech*. The MIT Press.

Hashimoto, R. I., Okada, R., Aoki, R., Nakamura, M., Ohta, H., & Itahashi, T. (2024). Functional alterations of lateral temporal cortex for processing voice prosody in adults with autism spectrum disorder. *Cerebral Cortex*, 34(9), bhac363. <https://doi.org/10.1093/cercor/bhae363>

Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011). The role of Broca's area in speech perception: evidence from aphasia revisited. *Brain and Language*, 119(3), 214–220. <https://doi.org/10.1016/j.bandl.2011.08.001>

Hertrich, I., Dietrich, S., & Ackermann, H. (2016). The role of the supplementary motor area for speech and language processing. *Neuroscience and Biobehavioral Reviews*, 68, 602–610. <https://doi.org/10.1016/j.neubiorev.2016.06.030>

Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: Revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, 44(8), 1996–2012. <https://doi.org/10.1007/s10803-014-2080-3>

Icht, M., & Ben-David, B. M. (2014). Oral-diadochokinesis rates across languages: English and Hebrew norms. *Journal of Communication Disorders*, 48(Apr), 27–37. <https://doi.org/10.1016/j.jcomdis.2014.02.002>

Kallay, J. E., Dilley, L., & Redford, M. A. (2022). Prosodic Development During the Early School-Age Years. *Journal of speech, language, and hearing research: JSLHR*, 65(11), 4025–4046. https://doi.org/10.1044/2022_JSLHR-22-00124

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2(3), 217–250

Katz, W. F., Garst, D. M., & Levitt, J. (2008). The role of prosody in a case of foreign accent syndrome (FAS). *Clinical Linguistics & Phonetics*, 22(7), 537–566. <https://doi.org/10.1080/02699200802106284>

Kisilevsky, S., Hains, S.M., Jacquet, A.-Y., Granier-Deferre, C., and Lecanuet, J.P. (2004). Maturation of fetal responses to music. *Developmental Science*. 7, 550–559. <https://doi.org/10.1111/j.1467-7687.2004.00379.x>

- Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage*, 102, 332–344. <https://doi.org/10.1016/j.neuroimage.2014.07.038>
- Krivokapi J. (2007). Prosodic Planning: Effects of Phrasal Length and Complexity on Pause Duration. *Journal of Phonetics*, 35(2), 162–179. <https://doi.org/10.1016/j.wocn.2006.04.001>
- Lartseva, A., Dijkstra, T., & Buitelaar, J. K. (2015). Emotional language processing in autism spectrum disorders: a systematic review. *Frontiers in Human Neuroscience*, 8, 991. <https://doi.org/10.3389/fnhum.2014.00991>
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America*, 32, 451–454. <https://doi.org/10.1121/1.1908095>
- Leiberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461. <https://doi.org/10.1037/h0020279>
- Lesage, E., Nailer, E. L., & Miall, R. C. (2016). Cerebellar BOLD signal during the acquisition of a new lexicon predicts its early consolidation. *Brain and Language*, 161, 33–44. <https://doi.org/10.1016/j.bandl.2015.07.005>
- Llanos, F., German, J. S., Gnanateja, G. N., & Chandrasekaran, B. (2021). The neural processing of pitch accents in continuous speech. *Neuropsychologia*, 158, 107883–107883. <https://doi.org/10.1016/j.neuropsychologia.2021.107883>
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, second edition (ADOS-2) manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.
- Lord, C., Rutter, M., DiLavore, P. S., & Risi, S. (1999). *Autism Diagnostic Observation Schedule (ADOS)*. Los Angeles, CA: Western Psychological Services.
- Marie, P. (1907). Presentation de malades atteints d'anarthrie par lesion de l'hémisphère gauche du cerveau. *Bulletins et Memoires Societe Medicale des Hopitaux de Paris*, 1: 158–160.
- Moore-Parks, E. N., Burns, E. L., Bazzill, R., Levy, S., Posada, V., & Müller, R. A. (2010). An fMRI study of sentence-embedded lexical-semantic decision in children and adults. *Brain and Language*, 114(2), 90–100. <https://doi.org/10.1016/j.bandl.2010.03.009>
- Monrad-Krohn, G. H. (1947). The prosodic quality of speech and its disorders: (A brief survey from a neurologist's point of view). *Acta Psychiatrica Scandinavica*, 22(3-4), 255–269.
- Murray, E., McCabe, P., Heard, R., & Ballard, K. J. (2015). Differential Diagnosis of Children with Suspected Childhood Apraxia of Speech. *Journal of Speech, Language, and Hearing Research*, 58(1), 43–60. https://doi.org/10.1044/2014_JSLHR-S-12-0358

Nadig, A., & Shaw, H. (2012). Acoustic and perceptual measurement of expressive prosody in high-functioning autism: increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders*, 42(4), 499–511. <https://doi.org/10.1007/s10803-011-1264-3>

O'Brien, A. M., Cardinaux, A. L., & Gabrieli, J. D. E. (2025). *The Vocal Output Experiences Questionnaire (VOX-Q): A Self-Report Measure of Prosody-Related Experiences in Adults*. <https://osf.io/gx8d7/>

O'Brien, A. M., Perrachione, T. K., Wisman Weil, L., Sanchez Araujo, Y., Halverson, K., Harris, A., Ostrovskaya, I., Kjelgaard, M., Kenneth Wexler, Tager-Flusberg, H., Gabrieli, J. D. E., & Qi, Z. (2023). Altered engagement of the speech motor network is associated with reduced phonological working memory in autism. *NeuroImage. Clinical*, 37, 103299. <https://doi.org/10.1016/j.nicl.2022.103299>

O'Brien, A. M., Perrachione, T. K., Gabrieli, J. D. E., D'Mello, A. M. (In Prep). The organization of functional regions is atypical in autism.

Ochi, K., Ono, N., Owada, K., Kojima, M., Kuroda, M., Sagayama, S., & Yamasue, H. (2019). Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PloS one*, 14(12), e0225377. <https://doi.org/10.1371/journal.pone.0225377>

Oderbolz, C., Poeppel, D., & Meyer, M. (2025). Asymmetric Sampling in Time: Evidence and perspectives. *Neuroscience and Biobehavioral Reviews*, 171, 106082. <https://doi.org/10.1016/j.neubiorev.2025.106082>

Park, H. (2020). The role of frontal speech motor system in speech perception: An fMRI study. *PLoS*, 45(1), 145-165.

Pell, M. D., & Baum, S. R. (1997). Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. *Brain and Language*, 57(2), 195-214. <https://doi.org/10.1006/brln.1997.1736>

Perron, M., Vuong, V., Grassi, M. W., Imran, A., & Alain, C. (2024). Engagement of the speech motor system in challenging speech perception: Activation likelihood estimation meta-analyses. *Human Brain Mapping*, 45(13), e70023. <https://doi.org/10.1002/hbm.70023>

Poeppel, D., & Monahan, P. J. (2010). Feedforward and feedback in speech perception: Revisiting analysis by synthesis. *Language and Cognitive Processes*, 26(7), 935–951. <https://doi.org/10.1080/01690965.2010.493301>

Rosenblau, G., Kliemann, D., Dziobek, I., & Heekeren, H. R. (2017). Emotional prosody processing in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*, 12(2), 224–239. <https://doi.org/10.1093/scan/nsw118>

- Ross E. D. (1981). The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Archives of Neurology*, 38(9), 561–569. <https://doi.org/10.1001/archneur.1981.00510090055006>
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M. C., & Cohen, D. (2013). Motherese in interaction: at the cross-road of emotion and cognition? (A systematic review). *PloS one*, 8(10), e78103. <https://doi.org/10.1371/journal.pone.0078103>
- Sakurai, Y., Itoh, K., Sai, K., Lee, S., Abe, S., Terao, Y., & Mannen, T. (2015). Impaired laryngeal voice production in a patient with foreign accent syndrome. *Neurocase*, 21(3), 289–298. <https://doi.org/10.1080/13554794.2014.892622>
- Schwartz, M., Tavano, A., Schröger, E., & Kotz, S. A. (2012). Temporal aspects of prediction in audition: cortical and subcortical neural mechanisms. *International Journal of Psychophysiology*, 83(2), 200–207.
- Scott, S. K., McGettigan, C., & Eisner, F. (2009). A little more conversation, a little less action--candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, 10(4), 295–302. <https://doi.org/10.1038/nrn2603>
- Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain and Language*, 164, 77–105. <https://doi.org/10.1016/j.bandl.2016.10.004>
- Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: motor cortical activation during speech perception. *NeuroImage*, 25(1), 76–89. <https://doi.org/10.1016/j.neuroimage.2004.11.006>
- Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, 44, 1–18. <https://doi.org/10.1016/j.dr.2016.12.001>
- Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, 23(2), 309–322. <https://doi.org/10.1007/BF01046222>
- Tian, L., Chen, H., Zhao, W., Wu, J., Zhang, Q., De, A., Leppänen, P., Cong, F., & Parviainen, T. (2020). The role of motor system in action-related language comprehension in L1 and L2: An fMRI study. *Brain and Language*, 201, 104714. <https://doi.org/10.1016/j.bandl.2019.104714>
- Vale, A. P., Fernandes, C., & Cardoso, S. (2022). Word reading skills in autism spectrum disorder: A systematic review. *Frontiers in Psychology*, 13, 930275. <https://doi.org/10.3389/fpsyg.2022.930275>

Wehrle, S., Grice, M., & Vogeley, K. (2024). Filled pauses produced by autistic adults differ in prosodic realisation, but not rate or lexical type. *Journal of Autism and Developmental Disorders*, 54(7), 2513–2525. <https://doi.org/10.1007/s10803-023-06000-y>

Witteman, J., Van Ijzendoorn, M. H., Van de Velde, D., Van Heuven, V. J., & Schiller, N. O. (2011). The nature of hemispheric specialization for linguistic and emotional prosodic perception: a meta-analysis of the lesion literature. *Neuropsychologia*, 49(13), 3722-3738.

Wu, Z. M., Chen, M. L., Wu, X. H., & Li, L. (2014). Interaction between auditory and motor systems in speech perception. *Neuroscience Bulletin*, 30, 490-496.

Zhang, M., Xu, S., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2022). Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism*, 26(4), 798–813. <https://doi.org/10.1177/1362361321995725>

Chapter 5:

Discussion & Conclusions

The present dissertation presents evidence from three studies that the speech production motor system plays an important role in supporting language and communication in autism. The key findings from each study, and their alignment with the corresponding aims and hypotheses are presented in Table 5.1. The remainder of the dissertation aims to synthesize themes that emerge across the studies, discuss key considerations for interpreting these findings, and outlines directions for future research.

Table 5.1.

Research Questions, Aims, Hypotheses, and Key Findings for Each Study Comprising the Dissertation

Chapter 2	<i>Title:</i> Atypical Motor System Engagement in Autistic Children During Nonword Repetition <i>Research Question:</i> What are the brain bases of phonological working memory impairments in autism? <i>Study Aims:</i> (3) Identify functional brain differences during nonword repetition in autistic vs. neurotypical children in the three key brain networks involved in nonword repetition: speech perception, working memory, and speech production. (4) Investigate the specificity of these differences to autism by comparing neural responses to the same task in children with a reading disability. <i>Study Hypotheses:</i> (3) Multivariate pattern analysis (decoding accuracy) and univariate analysis will reveal the brain regions that are most consistent or distinctive in response to more challenging phonological tasks in the autism group, and whether specific brain regions are being recruited to compensate during challenging phonological processing. (4) These differences will be specific to autism (i.e., not appearing in the reading disability group), and will be related to performance on language tasks and communication measures. <i>Primary Findings:</i> (1) Multivariate and univariate analyses identified key differences in the engagement of speech production motor regions – but not other networks of interest – in the autistic participants.
------------------	---

	(2) The specific speech motor brain differences found in the autistic participants were not observed in the reading disability group. (3) The engagement in the supplementary motor area for the 5 syllables > 2 syllable contrast, was correlated to performance on the nonword repetition task for the nonautistic but not autistic group. For the autistic participants, lower decoding accuracy (from the multivariate analysis) in the speech production network was associated with higher levels of communication challenges and autism traits.
Chapter 3	<i>Title:</i> Acoustic and Self-Reported Prosodic Differences in Autistic Adults <i>Research Question:</i> How do autistic adults experience prosody production and perception, what global prosodic features distinguish between autistic and nonautistic adults in a spontaneous speech task, and how do these subjective and objective features relate to autism traits and diagnosis? <i>Study Aims:</i> (4) Characterize the self-reported experiences of prosody in autistic and nonautistic adults. (5) Quantify and compare acoustic features (pitch, rate, pauses) of speech during a spontaneous picture description task between autistic and nonautistic adults. (6) Explore the relationship between self-reported prosody experiences, prosodic acoustic features, autism traits, nonverbal IQ, and language (measured via vocabulary size). <i>Study Hypotheses:</i> (3) Autistic adults will report greater difficulty with prosody perception and production. (4) There will be a combination of acoustic features of global prosody that distinguish between the autistic and nonautistic groups. <i>Primary Findings:</i> (1) Autistic adults reported more difficulty with prosody perception and production in their daily lives. (2) During a spontaneous speech sample, the autistic adults had a longer mean pause duration and a lower speech rate than the nonautistic adults. (3) Subgroups were identified based on acoustic and behavioral measures, including one group largely comprised of autistic individuals, one group largely comprised of nonautistic individuals, and one comprised of relative equal numbers. Key features in determining these groups included pause duration and articulation rate.
Chapter 4	<i>Title:</i> Precision fMRI Reveals Engagement of Speech Motor Regions During

	Prosody Perception and Differences in Autism <i>Research Question:</i> Are speech-production motor brain regions engaged differently during language tasks in autistic compared to neurotypical adults? <i>Study Aims:</i> (4) Investigate how nonautistic individuals engage their individually localized speech production system during a passive prosody perception task. (5) Examine if this engagement differs from that of autistic individuals. (6) Explore whether this engagement is related to behavioral performance on a prosody perception task outside of the scanner. <i>Study Hypotheses:</i> (4) Nonautistic individuals will engage some motor regions more when listening to sentences with a dynamic prosody contour than a flat contour. (5) Autistic individuals will demonstrate reduced activation of these regions. (6) Activation in these regions will be related to performance on a prosody perception task outside of the scanner. <i>Primary Findings:</i> (1) Nonautistic participants engaged over half of their precisely localized motor regions during the critical prosody perception contrast. (2) The autistic group had less activation in these motor regions during prosody perception. (3) Activation in these motor regions was associated with performance on a prosody task outside of the scanner within and across groups.
--	---

The Motor System, Language System, and Autism

The studies comprising the current dissertation provide evidence for (1) a role of the speech production motor system in supporting language processes in nonautistic children and adults in context-specific ways (i.e., related to nonword repetition and prosody perception) and (2) differences in the engagement of the motor system for autistic individuals across these contexts. Additionally, Chapter 3 provides insight into the specific behavioral impact that motor-related differences may have related to prosody, both when considering subjective lived experiences and objective speech production. The research in this dissertation adds to a growing and nuanced body

of work initially stemming from the motor theory of speech perception, which was proposed by Lieberman (1967) and refined in the subsequent decades (e.g., Wu et al., 2014). Indeed, while there has been significant consideration of the language and motor networks in isolation, the dynamic and interrelated nature of these regions, particularly at the individual level, is relatively understudied. Work that continues to uncover these motor-language links will be informative for general understanding of how language is processed, and may shed new light onto communication challenges in autism.

The work presented in this dissertation should *categorically not* be interpreted as support of Facilitated Communication (“FC”) or any related method (e.g., Rapid Prompting Method), all of which have been widely discredited as pseudoscientific (Hemsley et al., 2018; ISAAC, 2014; Schlosser et al., 2019). Proponents of these methods misleadingly argue that due to motor challenges, a “facilitator” must guide or stabilize an autistic individuals’ hand in order to enable expression (typically via typing). Empirical investigation of this method has undeniably identified the facilitator as being the source of the messages (Shane & Kearns, 1994). The motor-related brain differences reported in the present dissertation do not reflect such profound “limb apraxia” or even apraxia of speech. Instead, the findings here suggest that motor regions are being engaged less during language processing, and that this is related to communication challenges, but in no way suggest that external physical assistance is necessary to communicate.

Heterogeneity of Speech and Language Profiles in Autism

As has been highlighted throughout the present dissertation, autism is a highly heterogeneous condition. The participants included in the studies herein included individuals who use speech to communicate, and who were matched to the nonautistic participants on nonverbal IQ. Consequently, while the findings offer valuable insight into this subset of the autism spectrum,

they may not generalize to individuals with different language or cognitive profiles. Importantly, up to one-third of autistic individuals have limited or no speech (Rose et al., 2016), highlighting the need to consider a wider range of communication abilities in future work. Modeling work has suggested that a subgroup of minimally speaking individuals may present with motor speech disorder (Chenausky et al., 2019), and that motor speech ability is predictive of expressive language output (i.e., number of words) in this population. These findings underscore the importance of understanding the role of motor function in shaping language outcomes across the spectrum. Prosody may be a particularly informative component of language in this context. Unlike lexical content, prosody can be present even when word use is limited, making it a potentially useful window into communicative intent and the relationship between motor and linguistic processes across the spectrum. Indeed, preliminary efforts have been made to understand the prosody of children who are minimally verbal (Thorson et al., 2016; Johnson et al., 2022), offering insights into how speech motor control and expressive communication may interact in these individuals. The current findings, focused on individuals who primarily use speech to communicate, provide a foundation for extending this line of inquiry. Future research should continue to explore the intersections of motor ability, speech, and language outcomes across the full spectrum of autistic communication profiles.

Motor Prediction and Autism

There have been many “unifying” theories that attempt to account for the range of traits associated with autism, aiming to identify core mechanisms that might underlie the diverse manifestations of autism. One such theory suggests that difficulty with prediction (i.e., the brain’s ability to anticipate, sense, and respond to incoming information) may provide a common explanation for many autism-related traits (Sinha et al., 2014; for a review, see Cannon, O’Brien

et al., 2020). Related to the present dissertation, contemporary models of speech production propose that the motor cortex generates predictions, processes prediction errors, and uses these prediction errors to update speech gestures in real time (Guenther et al., 2016). Other work has suggested a predictive role of the speech motor cortex, including in supporting auditory-phonetic processing (Perron et al., 2024). There is some support for more general motor-based prediction differences in autism (Landa et al., 2017; Northrup et al., 2017; Brisson et al., 2012), although other work has shown intact low-level motor predictions (Finnemann et al., 2021). The connections between the speech motor system, language, and prediction – in relation to autism – are intriguing and merit further work. Indeed, it remains a possibility that differences in motor-based predictions related to speech production and perception may be contributing to the observed communication challenges in autism. Future investigation utilizing temporally sensitive methods (e.g., electroencephalography) may help clarify the timing and dynamics of predictive processes during speech and language processing in autism, shedding light on how these mechanisms contribute to individual differences in communication and motor function.

Considerations for Future Work

The present dissertation establishes the foundation for future research aimed at supporting autistic communication through the lens of the interconnected language and motor systems. A range of methods were employed to explore how lived experiences, speech and language patterns, and motor-based brain differences contribute to communication in autism. These approaches included consultation with a focus group of autistic consults, co-development of a custom self-report questionnaire, online behavioral experiments, behavioral characterization measures, and in-person behavioral and neuroimaging tasks. Together, these methods reflect an interdisciplinary

approach aimed at integrating perspectives from cognitive neuroscience, speech-language pathology, and autistic lived experiences to better understand speech and language in autism.

Future work should continue along this interdisciplinary path, not only to deepen scientific understanding of human communication more broadly, but also to better contextualize experimental findings with the real-world experiences of autistic individuals. Emphasizing meaningful collaboration with autistic individuals, their families, and other individuals who support them is essential for producing research that is both scientifically rigorous and also socially relevant (Keating, 2021). Future research should aim for a *deeper* understanding of the mechanisms driving motor system differences, including those related to prosody, and a *broad*er understanding of how motor system differences may impact other aspects of language in autism. Ultimately, such work has the potential to inform the development of more targeted and evidence-based strategies for supporting autistic individuals' language and communication.

References

- Brisson, J., Warreyn, P., Serres, J., Foussier, S., & Adrien-Louis, J. (2012). Motor anticipation failure in infants with autism: a retrospective analysis of feeding situations. *Autism, 16*(4), 420–429. <https://doi.org/10.1177/1362361311423385>
- Cannon, J., O'Brien, A. M., Bungert, L., & Sinha, P. (2021). Prediction in autism spectrum disorder: A systematic review of empirical evidence. *Autism Research, 14*(4), 604–630. <https://doi.org/10.1002/aur.2482>
- Chenausky, K. V., Brignell, A., Morgan, A. T., Norton, A. C., Tager-Flusberg, H. B., Schlaug, G., & Guenther, F. H. (2021). A Modeling-Guided Case Study of Disordered Speech in Minimally Verbal Children With Autism Spectrum Disorder. *American Journal of Speech-Language Pathology, 30*(3S), 1542–1557. https://doi.org/10.1044/2021_AJSLP-20-00121
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the United States of America, 111*(19), 7126–7131. <https://doi.org/10.1073/pnas.1318738111>
- Finnemann, J. J. S., Plaisted-Grant, K., Moore, J., Teufel, C., & Fletcher, P. C. (2021). Low-level, prediction-based sensory and motor processes are unimpaired in Autism. *Neuropsychologia, 156*, 107835–107835. <https://doi.org/10.1016/j.neuropsychologia.2021.107835>
- Hemsley, B., Bryant, L., Schlosser, R. W., Shane, H. C., Lang, R., Paul, D., Banajee, M., & Ireland, M. (2018). Systematic review of facilitated communication 2014–2018 finds no new evidence that messages delivered using facilitated communication are authored by the person with disability. *Autism & Developmental Language Impairments, 3*. SAGE Publications. <https://doi.org/10.1177/2396941518821570>
- ISAAC Position Statement on Facilitated Communication. (2014). *Augmentative and Alternative Communication, 30*(4), 357–358. <https://doi.org/10.3109/07434618.2014.971492>
- Johnson, K., O'Brien, A., Kershenbaum, A., Narain, J., Radhakrishnan, S., & Picard, R. (2022). Affective ratings of nonverbal vocalizations produced by minimally-speaking individuals: What do naive listeners perceive? *Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 1–8. <https://doi.org/10.1109/ACII55700.2022.9953820>
- Keating, C. T. (2021). Participatory autism research: How consultation benefits everyone. *Frontiers in Psychology, 12*, 713982. <https://doi.org/10.3389/fpsyg.2021.713982>
- Landa, R. J., Haworth, J. L., & Nebel, M. B. (2016). Ready, set, go! Low anticipatory response during a dyadic task in infants at high familial risk for autism. *Frontiers in Psychology, 7*, 721. <https://doi.org/10.3389/fpsyg.2016.00721>

Northrup, J. B., Libertus, K., & Iverson, J. M. (2017). Response to changing contingencies in infants at high and low risk for autism spectrum disorder. *Autism Research*, 10(7), 1239–1248. <https://doi.org/10.1002/aur.1770>

Rose, V., Trembath, D., Keen, D., & Paynter, J. (2016). The proportion of minimally verbal children with autism spectrum disorder in a community-based early intervention programme. *Journal of Intellectual Disability Research: JIDR*, 60(5), 464–477. <https://doi.org/10.1111/jir.12284>

Schlosser, R. W., Hemsley, B., Shane, H., Todd, J., Lang, R., Lilienfeld, S. O., Trembath, D., Mostert, M., Fong, S., & Odom, S. (2019). Rapid prompting method and autism spectrum disorder: Systematic review exposes lack of evidence. *Review Journal of Autism and Developmental Disorders*, 6(4), 403–412. <https://doi.org/10.1007/s40489-019-00175-w>

Shane, H. C., & Kearns, K. (1994). An examination of the role of the facilitator in “facilitated communication.” *American Journal of Speech-Language Pathology*, 3(3), 48–54.

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., ... & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences*, 111(42), 15220–15225. <https://doi.org/10.1073/pnas.141679711>

Appendix A: Supplementary Information for Chapter 2

Stimuli Used

Two-, three-, four, and five-syllable nonword stimuli. Nonword pronunciation is shown in parentheses using International Phonetic Alphabet transcription.)

2-Syllable Nonwords

plever / pever (ˈplevə / ˈpevə), pencid / pecid (ˈpensɪd / ˈpesɪd), tector / tetor (ˈtektə / ˈtətə), mubler / muber (ˈmublə / ˈmubə), sufting / sufing (ˈsɛftɪŋ / ˈsɛfɪŋ), bicket / bippet (ˈbɪkət / ˈbɪpət), blimpet / blimset (ˈblɪmpət / ˈblɪmsət), nadle / naidle (ˈneɪdəl / ˈnaɪdəl), shoken / shopen (ˈʃoʊkən / ˈʃoʊpən), sploiter / spleiter (ˈsplɔɪtə / ˈsplɛɪtə), cranscow / cranksow (ˈkrænskəʊ / ˈkrænskəʊ), troglen / trolgem (ˈtrægləm / ˈtrælgəm), kolite / kilote (ˈkoʊləɪt / ˈkaɪləʊt), preport / preprot (ˈpri:pərt / ˈpri:pɒt), riftoon / roofin (ˈrɪftʊn / ˈruftɪn), priote (ˈpraɪəʊt), kabant (ˈkæbənt), ballom (ˈbæləm), kather (ˈkæðə), gammert (ˈgæmət), shable (ˈʃeɪbəl), morant (ˈmɔrənt), neaving (ˈnivɪŋ), klamic (ˈklæmɪk), ablit (ˈæblɪt), blekking (ˈblekɪŋ), dorson (ˈdɔrsən), quipid (ˈkwɪpɪd), yantel (ˈyəntəl), promin (ˈproʊmən)

3-Syllable Nonwords

esconter (ˌɛsˈkɑntə) / thiprisant (ˈθɪpˌrɪsənt) / femotal (ˌfəˈmɒtəl) / breckerment (ˈbriˌkæmənt) / bindacy (ˈbɪnˌdæsi) / candronish (ˌkænˌdrɒnɪʃ) / seblantic (ˌsəˈblæntɪk) / kalandly (ˈkæˌlændli) / saplitize (ˈsæˌplɪtaɪz) / mesedine (ˌmesəˌdaɪn) / rekipance (ˌriˌkɪpəns) / siprisom (ˈsɪˌprɪsəm) / prothader (ˈproˌðeðə) / gronify (ˈgroˌnɪfaɪ) / minerack (ˈmɪnəˌræk) / accrevate or accrebat (ˈækriˌvet or ˈækriˌbet) / taggrement (ˈtægəˌmənt) / premonish (ˌpraɪməˌnɪʃ) / guppering (ˈgʌˌpəɪŋ) / restandent (ˌrɪˈstændənt) / misinter (ˌmɪzɪnˌtə) / jandersing (ˈdʒændəˌsɪŋ) / vagalish (ˈvegəˌlɪʃ) / trabalize (ˈtreɪˌlaɪz)

4-Syllable Nonwords

icnidator / icidator (ˈɪkniˌdeɪtə / ˈɪkiˌdeɪtə), metretry / meretry (ˈmetrəˌtɔri / ˈmerəˌtɔri), astragular / atragular (ˌæˈstræɡjʊlə / ˌæˈtræɡjʊlə), hibernatist / hiberatist (haɪˈbɜːnəˌtɪst / haɪˈbɜːtəˌtɪst), gasprodoxy / gasrodoxy (ˈɡæsprəsˌdaksɪ / ˈɡæsrəˌdaksɪ), decepoment / decegoment (ˌdiˌsɪpəmənt / ˌdiˌsɪɡəmənt), exvomition / exvotition (ˌɛksvəʊˈmɪʃən / ˌɛksvəʊˈtɪʃən), challopism / chollopism (ˈʃæləˌpɪzəm / ˈʃʊələˌpɪzəm), trallocistic / trallopistic (ˌtræləˌsɪstɪk / ˌtræləˌpɪstɪk), canastocize / canistocize (kæˈnæstoˌsaɪz / kænɪstoˌsaɪz), crimipism / cripimism (ˈkrɪmɪˌpɪzəm / ˈkrɪpɪˌmɪzəm), matastrophy / mastatrocy (məˈtæstroˌsɪ / məˈstætroˌsɪ), candalopy / canladopy (kænˈdæloˌpɪ / kænˈlædoˌpɪ), besepalment / bepesalment (biˌsɪpəlˌmənt / biˌpɪsəlˌmənt), reaquishment / reasquiment (ˌriˌækwɪzmənt / ˌriˌæzkwɪmənt), hemostify (ˌhiˌmoʊstəfaɪ), allotastry (ˌæləʊˌtæstri), shagonazle (ˈʃæɡəˌnæzəl), kamasticize (kəˈmæstoˌsaɪz), rendoristat (rənˈdɔriˌstæt), posidriate (ˌpoˌsɪdriət), benopify (ˌbiˌnɒpɪfaɪ), athandanate (ˌæˈθændənɛɪt), nuplarative (ˌnuˌplərətɪv), fandosity (ˌfænˈdoʊsɪti), masadolyte (məˈsædoˌlaɪt), kinimerate (ˌkɪˌnɪməˌɛɪt), cavanator (ˈkævəˌneɪtə), reostify (ˌriˌoʊstɪfaɪ), illostratic (ˌɪləˌstrætɪk)

5-Syllable Nonwords

ionificate (aɪəˌnɪfɪˌket) / uninagable (ˌʌnɪnˌægəbl) / mallerogasty (ˌmələrəˌgæsti) / osacreatic (oˈzakriˌætɪk) / androchiable (ænˈdrokiˌəbl) / veronicity (ˌvɜrəˌnɪsɪti) / kadastaline (kəˈdæstˌlaɪn) / kilometric (ˌkɪliəˌmɛtrɪk) / apristiation (əˌprɪstiˌeɪʃən) / kredomiotic (ˌkredəməˌatɪk) / cyoblisity

(,sɔɪə'blɪsəti) / aglomantacy (,ɛɡlo'mæntəsi) / aniobity (,æni'obɪti or ,æni'obɪdɪ) / reancidity (,riæn'sɪdɪti) / hieromaly (,hɪə-rə'məli) / amerondable (,æmə-'andəbl̩) / hereomatic (,hɪrɪo'mætɪk) / lamonicify ('læ,mənɪsɪfaɪ) / thidiometry (,θɪdɪ'amətri or ,vɪdɪ'amətri) / cabernacity (,kæbə-'næsɪti) / trandoriable (,træn'doriəbl̩) / remastography (rɪ,mæ'stagrəfi) / etocrastical (ɪto'kræ,stɪkl̩) / antriovacy ('æn,triəvəsi)

S1. Supplementary Methods

S1.1 Participant Information

S1.1.1 Diagnostic Confirmation. All children participants were confirmed by trained research staff using the Autism Diagnostic Observation Schedule (ADOS/ADOS-2) Module 3 or 4 (Lord et al., 2000; 2012). Where children with existing ASD diagnoses did not meet ASD criteria on the ADOS (n = 3), the ADOS videos were reviewed by a licensed neuropsychiatrist for final clinical judgment and group assignment. In each case the original diagnostic classification was confirmed.

S1.1.2 Language-Matched Groups

We compared two smaller, language-matched groups of NT (n = 18) and autism (n = 18) children who were matched on the core language score of the CELF, as well as age, sex, and nonverbal IQ. To generate these language-matched groups, we included an additional matching criteria of core language score. This effectively removed the lower performing autistic participants and the higher performing neurotypical participants, while adding two new neurotypical participants. Thus, the language-matched samples and the full samples are neither strict subsets nor entirely independent from each other, but these samples allowed us to understand if the differences between the autism and NT groups were being driven by the autistic children with poorer language skills and/or the NT children with superior language skills.

S1.2 fMRI Acquisition Methods

S1.2.1 Practice Scan. Before the experimental fMRI scan, all participants completed a practice session in a mock scanner. During the practice session, participants listened to and repeated eight nonwords (2 nonwords per syllable condition) that were not included in the experimental scan and practiced two rest trials during which they were prompted to be silent and still.

S1.2.2 In-Scanner Methods. When participants entered the scanner, the experimenter asked them to name the most important thing for the scanner. The experimenter prompted them to say “Staying still” if they did not independently provide this answer. A scan “buddy” remained in the MRI room with the participant and provided reminders to reduce motion during the scan.

S1.3 Analysis Methods

S1.3.1 Additional fMRIPrep Details. Each T1w (T1-weighted) volume was corrected for INU (intensity non-uniformity) using N4BiasFieldCorrection v2.1.0 (Tustison et al., 2010) and skull-stripped using antsBrainExtraction.sh v2.1.0 (using the OASIS template). Spatial normalization to the fsaverage5 template [(Fonov et al., 2009), RRID:SCR_008796] was performed through nonlinear registration with the antsRegistration tool of ANTs v2.1.0 [8, RRID:SCR_004757], using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast [(Zhang et al., 2001)] (FSL v5.0.9, RRID:SCR_002823).

Functional data was motion corrected using mcflirt (FSL v5.0.9 [9]). This was followed by co-registration to the corresponding T1w using boundary-based registration (Greve & Fischl, 2009) with six degrees of freedom, using flirt (FSL). Motion correcting transformations, BOLD-to-T1w transformation and T1w-to-template (MNI) warp were concatenated and applied in a single step using antsApplyTransforms (ANTs v2.1.0) using Lanczos interpolation. Physiological noise regressors were extracted using CompCor (Behzadi et al., 2007). Principal components were estimated for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). A mask to exclude signals with cortical origin was obtained by eroding the brain mask, ensuring it only contained subcortical structures. Six tCompCor components were then calculated including only the top 5% variable voxels within that subcortical mask. For aCompCor, six components were calculated within the intersection of the subcortical mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run.

S1.3.2 Probabilistic Maps. To visualize the degree of heterogeneity of activation patterns from the univariate analyses in each group, we generated probabilistic maps using voxels that met a threshold of $p < 0.01$ (cf. Fedorenko et al., 2010; Scott & Perrachione, 2019; Lipkin, B. et al., *in prep*). To verify that the NOI-specific brain regions identified through Neurosynth.org were actively engaged during our specific task, we overlaid the NOIs over parcellation maps derived from the thresholded probabilistic maps, and calculated percent overlap (**Supplementary Figure 2**).

S1.3.3 Anatomical Identification. Significant anatomical brain regions reported in this manuscript were identified using the Atlas of Intrinsic Connectivity of Homotopic Areas (AICHA) atlas, with cross-reference to Neurosynth.org.

S3.3 Supplementary Tables

Supplementary Table 1. ADOS calibrated severity scores by group

	Module 3	Module 4	CSS
NT Mean	1.21	1.17	1.19
NT SD	1.58	0.98	1.28
ASD Mean	10.73	9.63	10.18
ASD SD	3.83	5.13	4.48
RD Mean	2.15	2.00	2.08
RD SD	1.82	3.46	2.64
LM-NT Mean	1.33	1.75	1.54
LM-NT SD	1.72	1.26	1.49
LM-ASD Mean	10.64	9.86	10.25
LM-ASD SD	4.32	5.49	4.90

ADOS = Autism Diagnostic Observation Schedule; CSS = Calibrated Severity Score (averaged across modules 3 and 4); ASD = Autism Spectrum Disorder; NT = Neurotypical; LM = Language-matched; SD = standard deviation

Supplementary Table 2. Summary of racial and ethnic makeup of each group.

	Race					Ethnicity			
	White	Asian	More than one (White, American Indian/Alaska Native)	More than one (White, Asian)	Black or African American	No Response	Hispanic / Latino	Non-Hispanic	No Response
NT	20	5	0	3	0	0	1	26	1
ASD	19	0	1	3	0	0	0	23	0
LM-NT	12	1	3	2	0	0	0	16	0
LM-ASD	16	0	0	2	0	0	1	17	0
RD	13	0	1	0	1	1	0	18	0

LM = Language-matched; ASD = Autism Spectrum Disorder; NT = Neurotypical

Supplementary Table 3. Demographic information for language-matched groups.

		LM-NT n = 18 (4F)		LM-ASD n = 18 (5F)		LM-ASD < LM-NT p-value
		Mean	SD	Mean	SD	
Demographics	Age	12.1	2.9	12.9	3.1	.212
IQ	KBIT	112.3	11.9	109.6	15.7	.277
Language	CELF (CLS)	112.5	11.1	105.9	20.6	.131
NWR	In-scanner NWR	82.7	13.8	79.1	14.4	.037
	CTOPP (NWR)	9.7	3.1	8.1	2.2	.040
	CNRep (Raw)	33.8	4.6	32.1	4.6	.126
Reading	TOWRE (SWE)	107.6	8.9	97.7	11.2	.003
	TOWRE (PDE)	108.6	11.2	102.3	13.2	.067
	WRMT (WI)	110.9	10.4	103.2	15.8	.048
	WRMT (WA)	106.3	11.6	102.4	15.8	.203
	WJ	114.5	12.3	106.6	17.4	.062

LM = Language-matched; ASD = Autism Spectrum Disorder; NT = Neurotypical; CELF (CLS) = Core Language Score (Clinical Evaluation of Language Fundamentals, Fourth edition); NWR = nonword repetition; CNRep = Children's Test of Nonword Repetition; CTOPP = Comprehensive Test of Phonological Processing 2; KBIT = Kaufman Brief Intelligence Test 2 (Matrix subtest); WRMT = Woodcock Reading Mastery Tests; WI = Word Identification; WA = Word Attack; TOWRE = Test of Word Reading Efficiency; SWE = Sight Word Efficiency; PDE = Phonemic Decoding Efficiency. WJ = Woodcock-Johnson III Tests of Achievement (Sentence Reading Fluency subtest)

Supplementary Table 4. Mean number of motion outliers excluded (out of 96 volumes) and average framewise displacement (FD) for included trials across the three runs. Trials were excluded if the participant had a framewise displacement > 2.00mm. There were no group differences in the number of excluded trials or in average FD for included trials.

		NT		Autism		NT vs Autism	RD		NT vs RD	Autism vs RD
		Mean	SD	Mean	SD	p-value	Mean	SD	p-value	p-value
Full Sample	# of Outliers	4.52	4.54	6.00	4.99	.27	3.94	4.49	.92	.36
	Motion ¹	0.59	0.28	0.71	0.31	.15	0.65	0.24	.90	.16
Language-Matched Sample	# of Outliers	4.00	4.53	5.95	5.04	.22				
	Motion ¹	0.60	0.28	0.71	0.32	.31				

NT = Neurotypical; RD = Reading disability; SD = Standard deviation

¹ for each participant, motion was measured by mean framewise displacement.

Supplementary Table 5. MNI coordinates and anatomical labeling (based on the Automated Anatomical Labeling Atlas) of the regions included within each of the networks of interest used in the univariate and multivariate analyses. Coordinates represent anatomic centers of voxel clusters from meta-analysis derived network maps from Neurosynth. The numbers in parentheses indicate the percent of each cluster located in each structure.

Network of Interest	Volume	XYZ	Structure(s)
Perception			Temporal_Sup_R(64)Temporal_Mid_R(25)
			Temporal_Pole_Sup_R(3) -(3) Heschl_R(2)
	15621	57.7, -20.1, 1.2	Rolandic_Oper_R(1) Temporal_Pole_Mid_R(1)
	509	51.8, -6.0, 41.1	Precentral_R(67) Postcentral_R(33)
	101	43.0, 12.4, 14.5	Frontal_Inf_Oper_R(86) -(10) Rolandic_Oper_R(4)
			Cerebelum_7b_R(54) Cerebelum_8_R(32) -(11)
	148	38.5, -50.3, -43.7	Cerebelum_Crus2_R(3)
			Temporal_Mid_L(43) Temporal_Sup_L(41) -(8)
			Rolandic_Oper_L(3) Heschl_L(2)
	17669	-55.1, -24.5, 4.2	Temporal_Pole_Sup_L(1) Postcentral_L(1)
	795	-54.4, -2.4, 45.5	Precentral_L(53) Postcentral_L(45) -(2)
	154	-49.2, 15.3, -17.2	Temporal_Pole_Sup_L(100)
			Precentral_L(35) Postcentral_L(34)
	807	-50.7, -4.6, 19.7	Rolandic_Oper_L(28) -(3)
	101	-47.0, -42.9, -43.7	Cerebelum_7b_L(54) Cerebelum_Crus2_L(46)
Working Memory	241	-48.5, 16.1, 18.9	Frontal_Inf_Oper_L(96) Frontal_Inf_Tri_L(4)
	100	-19.0, -72.4, -31.9	Cerebelum_Crus1_L(94) Cerebelum_Crus2_L(6)
	292	-3.5, 5.8, 63.9	Supp_Motor_Area_L(100)
	54	-10.9, 63.3, 27.8	Frontal_Sup_Medial_L(79) Frontal_Sup_L(21)
	2236	31.9, -62.8, -31.2	Cerebelum_Crus1_R(65) Cerebelum_6_R(33) -(2)
	1736	34.1, 22.0, 0.5	Insula_R(71) -(22) Putamen_R(4) Frontal_Inf_Orb_R(3)

803	-34.5, -62.1, -29.7	Cerebelum_Crus1_L(74) Cerebelum_6_L(25) Cerebelum_Crus2_L(1) Frontal_Mid_L(96) Frontal_Mid_Orb_L(2)
1924	-35.2, 51.5, 9.4	Frontal_Inf_Tri_L(1) Frontal_Sup_L(1)
1143	-30.0, 21.2, 0.5	Insula_L(73) -(25) Putamen_L(2) Cerebelum_7b_R(70) Cerebelum_8_R(21)
340	30.4, -73.2, -51.1	Cerebelum_Crus2_R(5) -(4) Cerebelum_Crus2_L(48) Cerebelum_Crus1_L(39)
249	-4.2, -79.8, -22.4	Vermis_7(13) Cerebelum_Crus1_R(37) Cerebelum_6_R(31)
422	7.6, -75.4, -23.1	Vermis_7(28) Vermis_6(3) Cerebelum_Crus2_R(1) Frontal_Mid_R(64) Frontal_Inf_Tri_R(30) -(5)
11114	40.8, 36.7, 24.1	Frontal_Sup_R(1) Frontal_Inf_Oper_R(1) Parietal_Inf_L(43) Parietal_Sup_L(25) Precuneus_L(8) - (7) Occipital_Mid_L(7) Angular_L(5)
16192	-30.8, -58.4, 45.5	Occipital_Sup_L(4) Frontal_Inf_Oper_R(57) Precentral_R(33)
1319	49.6, 8.0, 24.8	Rolandic_Oper_R(10) Frontal_Mid_R(25) Supp_Motor_Area_L(19) Frontal_Sup_R(15) -(11) Cingulum_Mid_R(8) Supp_Motor_Area_R(8) Precentral_R(6) Cingulum_Mid_L(4) Frontal_Sup_Medial_L(3)
15623	17.2, 9.4, 50.7	Frontal_Sup_Medial_R(2) Parietal_Inf_R(28) Angular_R(20) Precuneus_R(15) - (11) Parietal_Sup_R(8) SupraMarginal_R(8)
13827	31.9, -56.2, 45.5	Occipital_Sup_R(7) Occipital_Mid_R(3) Precentral_L(29) Frontal_Mid_L(22) Frontal_Inf_Tri_L(22) Frontal_Inf_Oper_L(11) -(9)
20730	-38.9, 11.7, 37.4	Frontal_Sup_L(6) Postcentral_L(1)
52	-24.1, 49.3, 2.0	Frontal_Sup_L(39)
355	-16.0, -0.9, 13.8	Caudate_L(11) Thalamus_L(4) Putamen_L(2)
107	-29.3, -73.2, -56.3	Cerebelum_7b_L(1)
82	-10.9, -61.4, -57.0	Cerebelum_8_L(4) Cerebelum_7b_L(65) Cerebelum_Crus2_L(18)
403	-33.7, -70.2, -50.4	Cerebelum_8_L(9) -(8)
Production 1310	-19.7, -13.4, -1.0	Thalamus_L(35) -(34) Pallidum_L(19) Putamen_L(11) Cerebelum_6_L(94) -(2) Cerebelum_Crus1_L(2)
2501	-15.3, -63.6, -20.9	Lingual_L(1) Frontal_Inf_Oper_R(80) Rolandic_Oper_R(11)
350	56.2, 16.8, 2.7	Frontal_Inf_Tri_R(9) -(1)
477	-21.2, 3.5, 4.9	Putamen_L(57) Pallidum_L(41) -(1)
3057	17.9, -64.3, -19.4	Cerebelum_6_R(96) Lingual_R(2) -(1) Fusiform_R(1) Temporal_Sup_R(40) Postcentral_R(23) Temporal_Mid_R(10) -(9) Rolandic_Oper_R(8) Precentral_R(7) Heschl_R(2) SupraMarginal_R(1)
23557	56.2, -14.9, 14.5	Temporal_Pole_Sup_R(1)

		Postcentral_L(27) Temporal_Sup_L(23)
		Temporal_Mid_L(15) -(10) Rolandic_Oper_L(9)
		Precentral_L(8) Heschl_L(3) Frontal_Inf_Oper_L(3)
32262	-52.9, -15.6, 16.7	Temporal_Pole_Sup_L(1) Insula_L(1)
		Supp_Motor_Area_R(38) Supp_Motor_Area_L(35)
3937	0.2, 3.5, 55.8	Cingulum_Mid_L(19) Cingulum_Mid_R(8) -(1)
56	1.7, -38.5, -34.9	-(96) Vermis_10 (4)
336	-16.8, -29.6, 55.8	Postcentral_L(6) Cingulum_Mid_L(3)
52	-66.2, -38.5, -19.4	Temporal_Inf_L(92) -(8)
107	-58.1, 5.0, -20.9	Temporal_Mid_L(100)
187	26.7, -66.5, -51.1	Cerebelum_8_R(100)
50	6.8, -72.4, -46.7	-(66) Cerebelum_7b_R(25) Cerebelum_8_R(9)
427	14.9, -41.4, -38.6	-(75) Cerebelum_9_R(24) Cerebelum_10_R(1)
111	-5.7, -40.0, -32.7	-(94) Cerebelum_9_L(4) Vermis_10 (3)
52	48.9, -44.4, -34.9	Cerebelum_Crus1_R(92) Cerebelum_Crus2_R(8)
76	21.6, -1.6, 5.7	Pallidum_R(87) -(7) Putamen_R(6)
33	-38.9, 10.9, 4.9	Insula_L(100)
51	-27.1, 24.2, 8.6	Insula_L(100)
		Frontal_Inf_Oper_R(95) Insula_R(5)
142	43.7, 13.1, 9.4	Rolandic_Oper_R(1)
181	-63.2, -34.1, -12.8	Temporal_Mid_L(83) Temporal_Inf_L(17)
132	-16.8, 28.6, -9.8	Frontal_Sup_Orb_L(12) Rectus_L(9)
51	-58.1, -56.2, 22.6	Temporal_Mid_L(94) Angular_L(6)
302	-38.9, -85.7, -3.9	Occipital_Mid_L(62) Occipital_Inf_L(34) -(4)

Supplementary Table 6. Peak coordinates and anatomical locations of significant clusters for within and between group analyses for all syllable (task) > rest contrast.

Cluster-wise									
Comparison	<i>p</i> (FWE-corr)	<i>p</i> (uncorr)	k	T	Voxel-level <i>p</i> (uncorr)	x	y	z	Anatomical location
NT (within group)	.0008	.0002	890	14.56	.0002	48	-21	8	R STG
	.04	.02	47	10.56	.0002	-15	-21	-2	L STG
	.02	.006	153	9.69	.0002	0	0	62	SMA
ASD (within group)	.0002	.0001	1279	14.27	.0002	-42	-33	16	L STG
	.002	.0006	911	13.84	.0002	42	-9	30	R STG
	.02	.01	128	9.92	.0002	0	3	59	SMA

.03	.02	61	7.68	.0002	-12	-60	-	L Cerebellum
							20	
.03	.01	89	6.67	.0002	12	-60	-	R Cerebellum
							20	

Supplementary Table 7. MVPA classification accuracy for decoding two syllables vs. four syllable trials within each canonical NOI for the ASD participants in the full ASD group, compared to decoding accuracy for 5 syllables vs. 2 syllables. There are no significant differences between the classification accuracy in these two contrasts.

	Speech perception	WM	Speech production	Union	Intersection	Complement
Mean (4-syl v 2-syl)	52.3	52.9	51.7	53.1	49.3	53.8
Mean (5-syl v 2-syl)	52.5	52.0	52.7	52.2	51.6	53.3
p-value	.93	.65	.64	.67	.48	.79

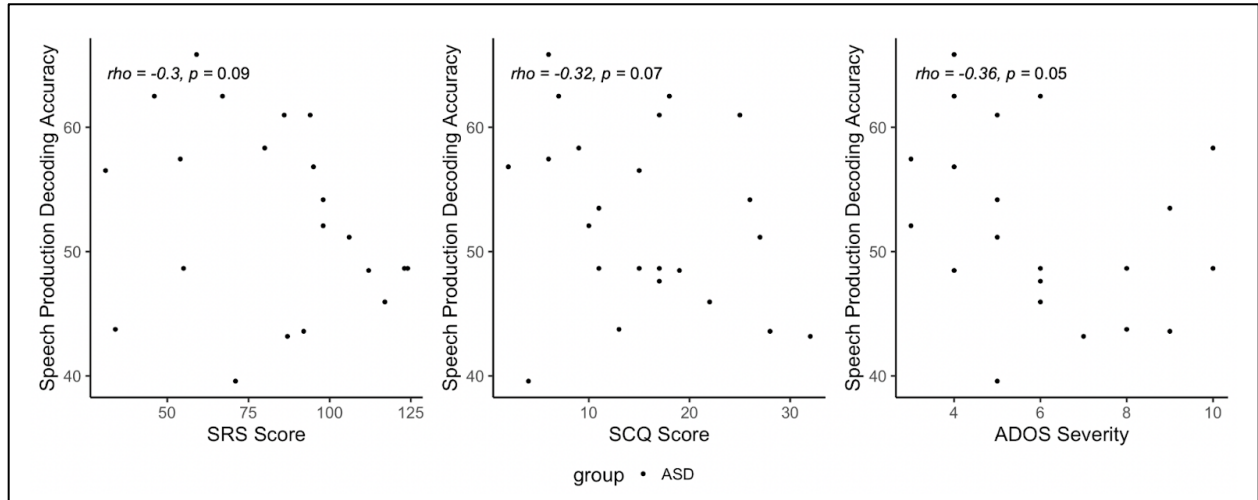
Supplementary Table 8. MVPA classification accuracy for decoding two syllables vs. five syllables in the RD and LM groups within each NOI.

Reading Disability (RD)			Language-matched NT			Language-matched ASD			
Network of Interest	Mean	SD	Against Chance FDR corrected p-value	Mean	SD	Against Chance FDR corrected p-value	Mean	SD	Against Chance FDR corrected p-value
Speech Perception Working Memory	52.0	12.1	.26	56.3	7.8	.0026	52.1	9.3	.18
Memory Speech Production	52.4	10.0	.26	54.5	9.7	.042	52.1	7.0	.17
	54.3	10.0	.16	57.6	8.8	.0026	52.6	7.0	.17

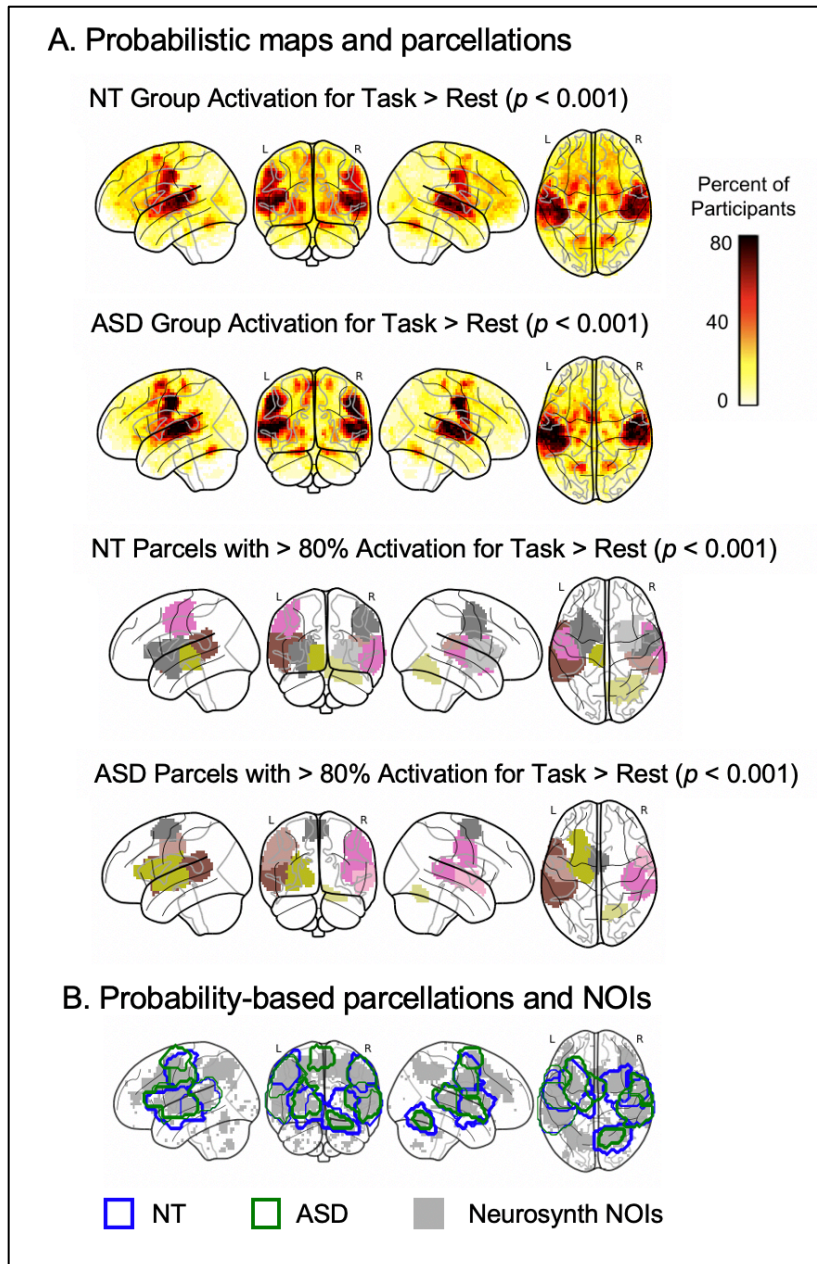
Supplementary Table 9. Two-sided Spearman rank correlations between age and the two primary brain measures (SMA activation and MVPA decoding accuracy) within the speech production region, within each group.

		NT	ASD
SMA Activation	Rho	0.17	0.087
	p-value	.38	.69
MVPA Decoding Accuracy	Rho	-0.16	0.27
	p-value	.42	.21

S4. Supplementary Figures



Supplementary Figure 1. Spearman's correlations between speech production decoding accuracy and scores on the Social Responsiveness Scale (SRS), Social Communication Questionnaire (SCQ), and ADOS Severity Score show moderate negative correlations between decoding accuracy in the autism group and autism-based symptoms (with increased severity of symptoms correlating with lower decoding accuracy). p 's are reported as one-tailed values.



Supplementary Figure 2. Probabilistic and parcellation maps. **A.** Probabilistic and parcellation maps of NT and ASD groups in all syllables > rest contrast ($p < .001$) showing the similar overall distribution and of common activation between groups. **B.** Notably, the NT and ASD parcels identified for the nonword repetition > rest contrast were similar in size and location to each other, and generally inclusive of the phonological working memory NOIs derived a priori from Neurosynth.org. The volumes of the speech perception, working memory, and speech production NOIs are: 29,516 mm³, 95,358 mm³, and 77,436 mm³, respectively. The overlap between the regions including significant ($p < .001$) activation for all syllables > rest contrast (in at least 80% of participants from either the NT or ASD group) and those from the speech perception, working memory, and speech production NOIs are: 29,516 mm³, 2,948 mm³, and 48,567 mm³, respectively.

Appendix A Supplementary References

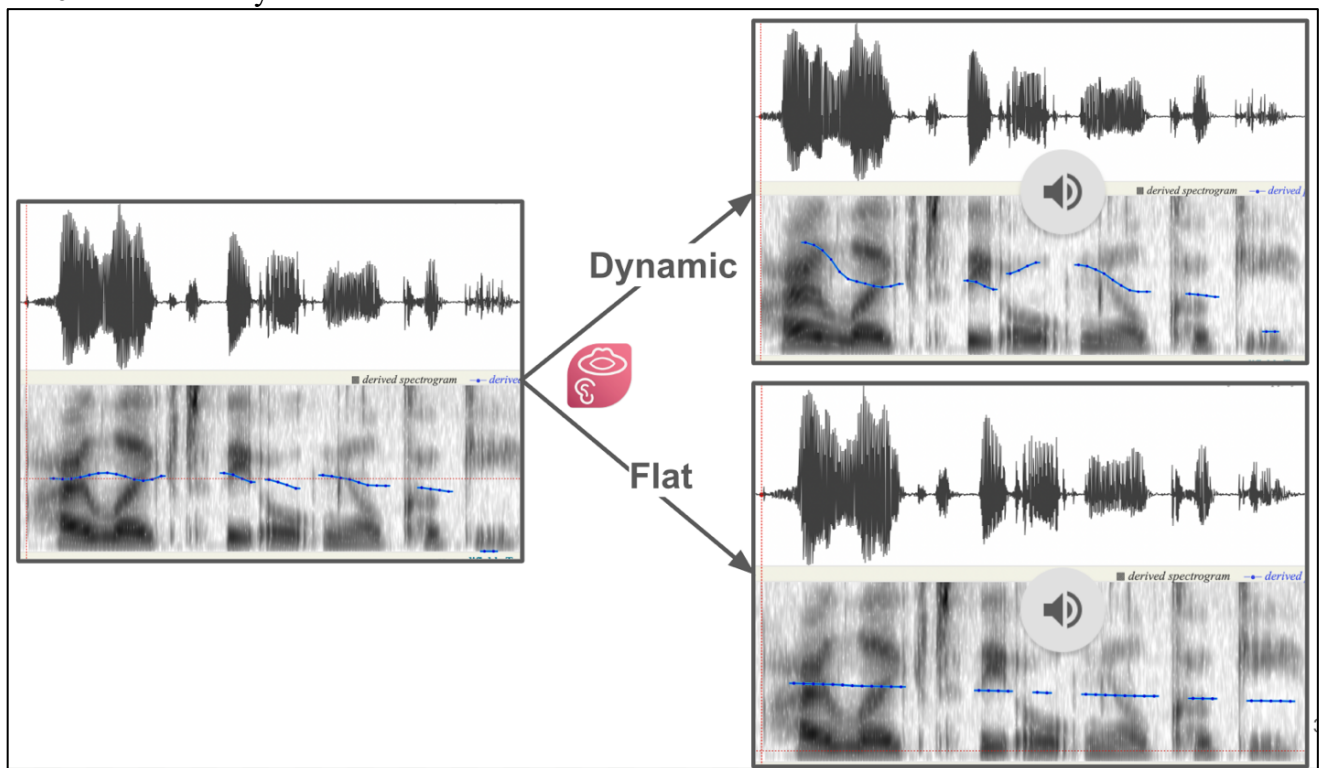
- Behzadi, Y., Restom, K., Liao, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37(1), 90–101. <https://doi.org/10.1016/j.neuroimage.2007.04.042>
- Fedorenko, E., Hsieh, P. J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, 104(2), 1177–1194. <https://doi.org/10.1152/jn.00032.2010>
- Fonov, V. S., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage (Orlando, Fla.)*, 47, S102–S102. [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5)
- Greve, & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage (Orlando, Fla.)*, 48(1), 63–72. <https://doi.org/10.1016/j.neuroimage.2009.06.060>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr, Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223.
- Lord, C., DiLavore, P. C., Gotham, K., Guthrie, W., Luyster, R., Risi, S., & Rutter, M. (2012). *Autism diagnostic observation schedule: ADOS-2* (2nd Edition). Western Psychological Services. <http://www.worldcat.org/title/autism-diagnostic-observation-schedule-ados-2/oclc/851410387>
- Scott, T. L., & Perrachione, T. K. (2019). Common cortical architectures for phonological working memory identified in individual brains. *NeuroImage*, 202, 116096.
- Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. *IEEE Transactions on Medical Imaging*, 29(6), 1310–1320. <https://doi.org/10.1109/TMI.2010.2046908>
- Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20(1), 45–57. <https://doi.org/10.1109/42.906424>

Appendix B: Supplementary Information for Chapter 4

Prosody Stimuli Manipulations

Instructions for creating flat sentences:

- For **FLAT** sentences:
 - Open file in Praat > To manipulate *[use standards]*
 - Remove pitch tiers between first and final pitch of sentence
 - Option Command T
 - Listen subjectively, make small adjustments for naturalness
 - Publish Resynthesis
- For **DYNAMIC** sentences:
 - Open file in Praat > To manipulate *[use standards]*
 - Remove pitch tiers between first and final pitch of sentence
 - Option Command T
 - Record the pitch of the midpoint, identify the new pitch (1.5x higher)
 - Add a new pitch point (Command T) at the same point where you measured, and drag it to match the new pitch
 - Add duration marker at beginning, end, and middle of the target word (Command D)
 - Move midpoint marker to 1.5
 - Listen subjectively, make small adjustments for naturalness
 - Publish Resynthesis



Note. The example sentence is: "I liked the oatmeal cookie"

Localizer Task and Parcels

The speech production localizer task successfully activated motor-specific brain regions in individual NA subjects ($N=33$; Figure SI-1) and resulted in sufficient group-level overlap to extract 13 speech production parcels that met threshold criteria (individual map threshold of $p<.0001$; region of interest and voxel overlap $\geq 10\%$; ROI size of $\geq 1,000$ voxels). These group-constrained parcels comprise regions in the left hemisphere ($n=5$), right hemisphere ($n=5$), and bilateral ($n=3$; Figure SI-1).

