INVESTIGATING SPEECH PERCEPTION AND INDIVIDUAL VARIATION IN COGNITIVE-BEHAVIOURAL TRAITS

INVESTIGATING SPEECH PERCEPTION AND INDIVIDUAL VARIATION IN COGNITIVE-BEHAVIOURAL TRAITS

BY SIMRAN SANDAL, B.S. (Hons)

A Thesis Submitted to the School of Graduate Studies in
Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University

MASTER OF SCIENCE (2025)

Hamilton, Ontario (Cognitive Science of Language)

TITLE: Investigating Speech Perception and Individual Variation in Cognitive-Behavioural Traits

AUTHOR: Simran Sandal, B.S. (Hons) (McMaster University)

SUPERVISORS: Dr. Phoebe Gaston and Dr. Daniel Pape

NUMBER OF PAGES: x, 69

Lay Abstract

Autism Spectrum Disorder (ASD) is often linked to differences in communication and language processing. Research suggests that one contributing factor may be a tendency for people with more autistic traits to focus on fine details in sounds while using context differently than others. These traits, measured by the Autism Spectrum Quotient (AQ), affect how people hear and adapt to changes in speech. The present study examined whether autistic traits in the general population influence how listeners identify speech sounds and how their brain responds. Results showed that all participants categorized the sounds consistently, and their judgements were influenced by the surrounding vowel. However, individuals with lower AQ scores interpreted consonant sounds differently than those with higher scores. Brain activity further showed that individuals with higher AQ scores had stronger but slower response when consonants occurred with certain vowels. These findings suggest that autistic traits are linked to subtle differences in both speech perception and brain sensitivity to context.

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition associated with differences in cognition, communication and language processing. These differences may be explained by their distinct cognitive and perceptual processing styles compared to those of neurotypical individuals. Previous studies have shown that individual differences in traits linked to autism, as measured by the Autism Spectrum Quotient (AQ), influence speech perception and the ability to adapt to contextual variability in speech signals (Stewart & Ota, 2008; Yu, 2010).

The present study examined how variation in AQ traits is associated with behavioural and neural responses to ambiguous speech sounds. Fifty-two participants completed a two-alternative forced choice (2AFC) task, categorizing a continuum of fricative sounds between /s/ and /ʃ/ in different vowel contexts (/a/ and /u/). Electroencephalography (EEG) was also recorded during a passive oddball paradigm to measure mismatch negativity (MMN) brain responses to such vowel-fricative-vowel syllables.

Results from the behavioural task showed strong categorical perception across participants, with vowel context reliably influencing categorization patterns. Comparison of the highest and lowest AQ quartiles showed that Low AQ participants categorized stimuli more often as /ʃ/ overall. Analysis of the 50% cross over point (PSE) indicated that High AQ participants shifted toward /ʃ/ earlier than the Low AQ group, particularly in the /u/ context. Neural responses showed that participants exhibited reliable event-related potential (ERP) responses in the MMN time window, although polarity was positive rather than negative. Group-level differences emerged in both ERP amplitude and latency, with High AQ individuals showing stronger but slower responses to stimuli in the /u/ context. These findings highlight how individual differences can shape both perceptual and neural mechanisms of speech processing.

Acknowledgments

This project has been one of the most meaningful academic experiences of my life. It has challenged me to think critically, pushed me outside my comfort zone, and taught me what it means to be part of a research community. I hope the results of this work contribute, even in a small way, to a better understanding of how people perceive and make sense of speech. I would like to express my deepest appreciation to the following people, without whom this thesis would not have been possible:

To my supervisors, Dr. Phoebe Gaston and Dr. Daniel Pape, thank you for your guidance and support throughout this project. Phoebe, thank you for giving me the opportunity to learn EEG and for guiding me through the complexities of experimental design and analysis. Your patience with my many questions and insight have been invaluable. Daniel, thank you for introducing me to the world of phonetics and speech perception research. I have learned so much from your expertise and enthusiasm, and you have helped me discover a passion for this field that I will carry with me throughout my academic and professional journey.

To the amazing group of research assistants in the EEG lab, thank you for your dedication and hard work during the data collection process. Your effort and teamwork made this project possible, and I am so grateful to have had the chance to work alongside all of you.

To my parents, Prabhjit and Gurpal Sandal, thank you for your endless love and encouragement. Your belief has carried me through every stage of this journey, and this thesis is as much of a reflection of your support as it is my work.

Finally, thank you to all the participants, my friends and colleagues, and everyone at ARiEAL Research Centre for their support. This journey has been an incredible opportunity, and I am profoundly grateful to everyone who supported me through it.

Table of Contents

Lay Abstract ... iii
Abstract ... iv
Acknowledgments ... v
List of Figures and Tables ... viii
List of all Abbreviations and Symbols ... x
1. Introduction ... 1
1.1 Categorical Perception ... 1
1.2 Perceptual Compensation for Coarticulation ... 2
1.3 Coarticulatory Vowel Context Effects ... 3
1.4 Autism Spectrum Disorder (ASD) ... 4
1.4.1 Weak Central Coherence (WCC) ... 4
1.4.2 Enhanced Perceptual Functioning (EPF) ... 6
1.4.3 Neural Correlates of Atypical Speech Perception ... 7
1.5 Mismatch Negativity (MMN) ... 7
1.6 The Present Study ... 10
2. Methods ... 11
2.1 Participants ... 11
2.2 Autism Spectrum Quotient (AQ) ... 12
2.3 Stimuli ... 14
2.4 Behavioural Task ... 14
2.5 EEG Task ... 15
2.5.1 EEG Recordings ... 15
2.5.2 EEG Pre-Processing and Analysis ... 16
3. Results ... 17
3.1 AQ score ... 17
3.2 Two-Alternative Forced Choice (2AFC) Identification Task ... 18
3.2.1 Continuous AQ ... 19
3.2.2 Group-Based AQ ... 21
3.2.3 Confidence Rating ... 23
3.2.4 Point of Subjective Equality (PSE) ... 24
3.3 MMN Amplitudes ... 28
3.3.1 Continuous AQ ... 28
3.3.2 Group-Based AQ ... 30
3.4 MMN Latency ... 33
3.4.2 Group-Based AQ ... 35
3.5 Exploratory Results ... 36
3.5.1 Gender Differences in Behavioural Categorization ... 36
3.5.2 P200 Response ... 39
3.5.3 P200 Amplitude ... 39
3.5.4 P200 Latency ... 42
3.5.5 Late MMN ... 45
3.5.6 Late MMN Amplitude ... 45
3.5.7 Late MMN Latency ... 47
4. Discussion ... 50

4.1 Behavioural Findings 50
4.2 Neural Findings 53
4.2.1 ERP Response Amplitude.. 53
4.2.2. ERP Latency 54
4.3.3 P200 Overlap and Late MMN 56
4.2.4 Summary of Neural Patterns.. 57
4.3 Limitations... 58
4.4 Conclusion.. 59
References 61

List of Figures and Tables

Figure 1.Example trial sequence for the two-alternative forced-choice (2AFC) behavioral task implemented in Gorilla 15
Figure 2. Autism-Spectrum Quotient (AQ) subscale scores by participant........ 18
Figure 3. Proportion of /ʃ/ responses across z-scored continuum steps, plotted by vowel context 20
Figure 4. Categorization curves for Low and High AQ groups, plotted by vowel context........ 22
Figure 5. Mean confidence ratings for each audio stimulus in the 2AFC task, separated by vowel context........ 23
Figure 6. Mean confidence ratings across continuum steps by AQ group and vowel context........ 24
Figure 7. Relationship between z-scored AQ scores and point of subjective equality (PSE) values 25
Figure 8. Distribution of phoneme categorization boundaries (PSEs) by AQ group (Low vs. High) and vowel context (/a/ vs. /u/) 27
Figure 9. Event-related potentials (ERPs) at electrode FCz for the /a/ and /u/ vowel contexts in Low AQ and High AQ groups........ 31
Figure 10. Scalp topographies of the MMN difference wave (deviant minus standard) in the 100–250 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups........ 32
Figure 11. Mean ERP response amplitudes (µV) in the 100–250 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups. 32
Figure 12. Relationship between z-scored AQ scores and ERP latency (ms) for the /a/ (left) and /u/ (right) vowel contexts........ 34
Figure 13. Mean ERP latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups 36
Figure 14. Proportion of /ʃ/ responses across the fricative continuum by vowel context and gender........ 38
Figure 15. Mean P200 response amplitudes (µV) in the 180–250 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups. 41
Figure 16. Relationship between z-scored AQ scores and P200 latency (ms) for the /a/ and /u/ vowel contexts 43
Figure 17. Mean P200 latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups 44
Figure 18. Mean Late MMN response amplitudes (µV) in the 300–500 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups........ 47
Figure 19. Relationship between z-scored AQ scores and Late MMN latency (ms) for the /a/ and /u/ vowel contexts 48
Figure 20. Mean Late MMN latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ and High AQ groups........ 49

Table 1. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with AQ as a continuous predictor........ 21
Table 2. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with AQ grouped........ 23
Table 3. Effects of vowel context and continuous AQ (AQ_z) on PSE........ 26
Table 4. Effects of vowel context and AQ group (Low vs. High) on PSE........ 27
Table 5. Effects of vowel context and continuous AQ (AQ_z) on categorization slope........ 27
Table 6. Effects of vowel context and AQ group (Low vs. High) on categorization slope........ 28
Table 7. Linear mixed-effects model estimates for ERP response amplitude as a function of AQ score, vowel context and scalp region........ 29
Table 8. Linear mixed-effects model estimates for ERP response amplitude as a function of AQ group, vowel context and scalp region........ 33
Table 9. Linear mixed-effects model estimates for ERP Response latency as a function of AQ score, vowel context and scalp region........ 35
Table 10. Linear mixed-effects model estimates for ERP response latency as a function of AQ group, vowel context and scalp region........ 36
Table 11. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with gender as a predictor........ 38
Table 12. Linear mixed-effects model estimates for P200 response amplitude as a function of AQ score, vowel context and scalp region........ 40
Table 13. Linear mixed-effects model estimates for P200 response amplitude as a function of AQ group, vowel context and scalp region........ 41
Table 14. Linear mixed-effects model estimates for P200 response latency as a function of AQ score, vowel context and scalp region........ 43
Table 15. Linear mixed-effects model estimates for P200 response latency as a function of AQ group, vowel context and scalp region........ 44
Table 16. Linear mixed-effects model estimates for Late MMN response amplitude as a function of AQ score, vowel context and scalp region........ 45
Table 17. Linear mixed-effects model estimates for Late MMN response amplitude as a function of AQ group, vowel context and scalp region........ 47
Table 18. Linear mixed-effects model estimates for Late MMN response latency as a function of AQ score, vowel context and scalp region........ 48
Table 19. Linear mixed-effects model estimates for Late MMN response latency as a function of AQ group, vowel context and scalp region........ 50

List of all Abbreviations and Symbols

ADHD – Attention Deficit Hyperactivity Disorder

ARiEAL – Advanced Research in Experimental and Applied Linguistics

ASD – Autism Spectrum Disorder

AQ – Autism Spectrum Quotient

CMS – Common Mode Sense

CI – Confidence Interval

DRL – Driven Right Leg

EEG – Electroencephalography

EPF – Enhanced Perceptual Function

ERP – Event-Related Potential

ICA – Independent Component Analysis

MEG – Magnetoencephalography

MMN – Mismatch Negativity

OR – Odds Ratio

REST – Reference Electrode Standardization Technique

TD – Typically Developing

VFV – Vowel-Fricative-Vowel

WCC – Weak Central Coherence

2AFC – 2-Alternative Forced Choice Task

1. Introduction

Speech perception plays an important role in language processing as it allows individuals to interpret speech sounds as meaningful words and sentences. However, speech in naturalistic environments is highly variable due to factors such as speaker differences, contextual influences and articulation patterns. Despite this variability, most individuals will accurately perceive speech through mechanisms such as categorical perception and perceptual compensation, which help listeners to resolve ambiguity and be consistent in their interpretations of speech sounds. However, these mechanisms are not uniform across all listeners. Cognitive-behavioural traits, such as those associated with Autism Spectrum Disorder (ASD), have been argued to shape the way listeners process and compensate for variability in the speech signal. This thesis investigates how individual differences in cognitive-behavioural traits, specifically those related to the Autism Spectrum Quotient (AQ), influence both behavioural and neural response to a continuum of fricative[1]-vowel syllables. By using a combination of behavioural (speech perception) tasks and event-related potential (ERP) measures, specifically the Mismatch Negativity (MMN) response, this work aims to explore how fine-grained acoustic variation is processed across individuals, and how this relates to adapting speech perception to context.

1.1 Categorical Perception

A characteristic of speech processing is categorical perception, where acoustically varying speech sounds are heard as belonging to distinct categories (Fuhrmeister & Myers, 2021). For example, in English, /s/ as in "*sip*" and /ʃ/ as in "*ship*" are both fricatives (friction sounds, and particularly sibilants (hissing sounds), however, they differ in their spectral properties. Specifically, /ʃ/ has a lower center of gravity (the frequency at which energy in the

[1] Fricatives are consonant sounds, like /s/ and /ʃ/, produced by pushing air through a narrow constriction in the mouth.

sound is concentrated) than /s/. The difference in spectral properties allows listeners to perceive these fricatives as distinct phonetic categories, creating a perceptual boundary between them (Fuhrmeister & Myers, 2021). The organization of speech sounds into phonological categories can be influenced by both language experience and exposure, meaning that two sounds that are easily distinguishable for one listener may be difficult to differentiate for another (Fuhrmeister & Myers, 2021). An early study by MacKain et al. (1981) investigated the categorical perception of a /r/-/l/ continuum in English speakers and Japanese bilinguals with varying levels of English experience. English and Japanese speakers differ in how they categorize /r/ and /l/, in English, they are distinct phonemes that separate words, but in Japanese, they are not. The results showed that Japanese listeners with English experience could categorically perceive the contrasts, similar to native English speakers, but their identification functions were less sharply defined (MacKain et al., 1981). In contrast, those with little English experience performed poorly across identification tasks when shown different exemplars (MacKain et al., 1981). Although cross-linguistic research has demonstrated that language experience plays a role in categorical perception, accumulating evidence suggests that even speakers of the same language vary in how they perceive speech sounds. For instance, Kong and Edwards (2016) found that some English listeners made more gradient responses to a stop voicing continuum, such as /k/ versus /g/. Individual response patterns varied from highly categorical to more continuous. Together, these findings highlight that speech perception is shaped by not only language background but also by individual differences.

1.2 Perceptual Compensation for Coarticulation

Speech sounds are not produced in isolation, phonetic context plays a crucial role because the way sounds are articulated changes depending on surrounding consonants, vowels, as well as speaking conditions (Campbell et al., 2018). As a result of this contextual variation, sounds are not always perceived as fixed categories (Campbell et al., 2018). Coarticulation refers to the natural overlap and alteration of speech sounds based on neighbouring sounds that occur during fluent speech (Yu & Lee, 2014; Daniloff & Hammarberg, 1973). To accommodate this variability in speech production, the perceptual system engages in perceptual compensation for the acoustic effects of coarticulation (Yu & Lee, 2014). Perceptual compensation describes the brain's ability to help listeners adjust their interpretation of speech sounds, allowing them to accurately perceive speech despite the acoustic fluctuations introduced by coarticulation (Yu & Lee, 2014). In addition to the acoustic influences, perceptual compensation has also been shown to be shaped by social and interactional factors, including the identity of the speaker, expectations of the listener and broader sociolinguistic background (Yu, 2010). Despite significant research in this area, further investigation is needed to understand how individual-level and contextual factors affect speech perception and the ability to adapt to variability.

1.3 Coarticulatory Vowel Context Effects

Among the many sources of contextual variation that influence speech perception, vowel context is a common contributor of coarticulatory effects. In natural speech, vowels can significantly affect the articulation and acoustics of adjacent consonants (Yu, 2010). This is especially noticeable with rounded vowels like /u/ (as in *suit*), which involve lip rounding and a retracted tongue position that shortens the front cavity of the vocal tract (Recasens, 2018). These articulatory adjustments can cause a spectral frequency lowering effect on the preceding fricatives (Yu, 2010; Yu & Lee, 2014). For example, when /s/ occurs before rounded vowels, its spectral peak shifts downward, making it acoustically more similar to /ʃ/ (Mann & Repp, 1980; Yu, 2010, 2019; Yu & Lee, 2014). As a result, the perceptual boundary between /s/ and /ʃ/ may become blurred, increasing ambiguity in fricative identification. Several studies have reported findings where listeners are more likely to categorize ambiguous speech sounds along a /s/-/ʃ/ continuum as /s/ when they occur before /u/ than before /a/ (Mann & Repp, 1980; Yu & Lee, 2014; Yu, 2019). However, the strength of this context effect may vary across individuals, as some may be more sensitive to coarticulatory cues or differ in how strongly they compensate for acoustic variations.

1.4 Autism Spectrum Disorder (ASD)

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition characterized by a wide range of abilities and challenges that vary across individuals (Stewart & Ota, 2008). One area where differences are commonly observed is in speech perception and language processing. Individuals with ASD often show atypical patterns of auditory perception and reduced sensitivity to phonological context (Yu & To, 2020). These differences may contribute to broader difficulties in language comprehension and social communication.

Many of the traits associated with ASD, such as a heightened focus on detail or reduced integration of contextual cues, are not exclusive to clinical populations. Instead, they are thought to exist along a continuum that extends into the general population (Baron-Cohen et al., 2001b; Stewart & Ota 2008). The Autism Spectrum Quotient (AQ) is a widely used self-report measure of these traits, allowing researchers to study how differences in cognitive-behavioural traits may influence speech perception among the broader population (Baron-Cohen et al., 2001b).

1.4.1 Weak Central Coherence (WCC)

To better understand how ASD and related traits influence perception and language processing, researchers have proposed cognitive models that characterize differences in information processing style. One influential framework proposed is the Weak Central

Coherence (WCC) theory. According to this model, individuals with ASD exhibit a cognitive style characterized by an enhanced focus on local-level details and a reduced tendency to integrate information into a broader context (Happé & Frith, 2006; Stewart & Ota, 2008). Importantly, WCC is not considered a deficit but a difference in how perceptual and cognitive information is prioritized. Happé and Frith (2006) argue that this style of processing is domain general and can influence a range of behaviours, from visual pattern recognition to language comprehension.

In the context of speech perception, this processing bias may lead to a reduced reliance on top-down cues (information derived from prior knowledge, experience and expectations) such as sentence context or lexical expectations, and a stronger attention to bottom-up acoustic features (based on the incoming acoustic signal itself) (Happé & Frith, 2006). In other words, individuals with ASD may focus on the specific sounds they hear, rather than using broader contextual information. For example, in a homograph (words that are spelled the same but have different meanings) disambiguation task, individuals with ASD relied less on sentence-level information to resolve ambiguous words compared to the typically developing (TD) group (Happé, 1997). Additionally, Booth and Happé (2010) used a Sentence Completion Task, in which participants were asked to finish sentence stems that could be completed in either a contextually appropriate way (global completion) or literal/detail-focused way (local completion). For example, given the prompt "You can go hunting with a knife and…," a global completion would be "catch a bear" and a local completion would be "fork" (Booth & Happé, 2010). They found that individuals with ASD produced significantly more local completions compared to both ability-matched Attention Deficit Hyperactivity Disorder (ADHD) and TD groups (Booth & Happé, 2010). Similarly, Yu and To (2020) showed that individuals with ASD exhibited reduced sensitivity to coarticulatory vowel context effects during fricative perception.

Similar variation in speech perception patterns has been observed in the broader population using the AQ. Stewart and Ota (2008) found that neurotypical listeners with higher AQ scores showed a weaker Ganong effect (tendency to bias phonetic categorization toward real words over non-words in ambiguous contexts), indicating reduced influence of lexical context on categorization. Neurophysiological evidence from Ahtam et al. (2020) further supports these findings by demonstrating that individuals with ASD show atypical brain responses during semantic processing, including reduced early sensitivity to contextual information and increased right-hemisphere lateralization. These neural differences are associated with altered integration of sentence context and word meanings.

1.4.2 Enhanced Perceptual Functioning (EPF)

While WCC emphasizes reduced contextual integration, another framework, the Enhanced Perceptual Functioning (EPF) model, proposes that individuals with ASD have superior low-level perceptual processing that may contribute to their detailed focus on sensory input (Mottron et al., 2006). For example, Bonnel et al. (2003) reported that individuals with ASD demonstrated heightened pitch sensitivity, performing significantly better than typically developing controls in tasks requiring pitch discrimination and categorization. In a similar study, Järvinen-Pasley et al. (2008) observed that children with autism outperformed controls on perceptual discrimination tasks involving pitch contour and temporal patterns. However, these same participants showed poorer sentence comprehension, suggesting that enhanced perceptual abilities do not necessarily translate to efficient linguistic integration (Järvinen-Pasley et al., 2008).

1.4.3 Neural Correlates of Atypical Speech Perception

The pattern of perceptual enhancement, along with challenges in discrimination, may be linked to differences in brain organization in ASD. In particular, increased right-hemisphere lateralization observed in ASD reflects a deviation from the typical left-dominant pattern seen in neurotypical individuals during language processing. Ahtam et al. (2020) found that individuals with ASD showed significantly greater right-hemisphere activity, especially over the occipito-temporal regions, during both early and later stages of semantic processing. Additional support for non-left dominant hemispheric organization in autism comes from positron emission tomography (PET) studies. Boddaert et al. (2003) found that both the TD and autism group showed bilateral activation of the superior temporal gyrus (a brain region important for processing sounds and language) when listening to speech-like sounds. However, the group with autism exhibited significantly larger activation in the right hemisphere compared to the left, which was reversed in the TD group (Boddaert et al., 2003). The lateralization pattern observed in individuals with ASD may reflect a shift toward right-hemisphere dominance, which is more specialized for spectral processing (Haesen et al., 2011). The right-hemisphere dominance may explain enhanced pitch and vowel perception in ASD, while differences in left hemisphere function (critical for temporal processing) could contribute to challenges seen in syllable discrimination and overall speech processing (Haesen et al., 2011).

1.5 Mismatch Negativity (MMN)

Mismatch Negativity (MMN) is an Event-Related Potential (ERP) component that reflects the brain's automatic ability to detect changes in a train of auditory stimuli (Garrido et al., 2009). In experimental conditions, MMN is typically elicited using an oddball paradigm where a frequently repeated "standard" sound is occasionally replaced by a "deviant" sound that differs from the standard in characteristics such as frequency, intensity or phonological features (Garrido et al., 2009). MMN can be understood with the predictive coding framework, which posits that perception arises from comparing sensory input with top-down predictions (Garrido et al., 2009). According to this view, MMN reflects a prediction error signal generated when a deviant sound violates the brain's expectation based on the previously established auditory pattern (Garrido et al., 2009). The process involves hierarchical interactions between auditory cortical areas, combining local adaptation and synaptic plasticity (Garrido et al., 2009). Thus, MMN reflects the brain's mechanism for making perceptual inferences and learning, as it continuously updates its internal model to minimize prediction errors and optimize sensory processing (Garrido et al., 2009). The MMN is measured by subtracting the ERP response to the standard sounds from the response to the deviant sounds and, thus, creating a difference wave (Garrido et al., 2009). The MMN commonly appears as a negative deflection in the difference waveform and is typically maximal across the fronto-central electrodes, while showing a polarity inversion at posterior electrode sites (Garrido et al., 2009). MMN can be recorded using methods such as electroencephalography (EEG) and magnetoencephalography (MEG). Importantly, MMN occurs even when attention is directed away from the sounds and no active behavioural response is required, making it a reliable index of automatic, pre-attentive auditory processing (Garrido et al., 2009). Due to its ability to objectively measure sound discrimination without requiring conscious effort, MMN is a valuable neurophysiological marker for studying the neural mechanisms underlying speech perception.

Research using MMN paradigms has shown altered auditory processing in individuals with ASD. Several studies have reported reduced MMN amplitudes in response to speech and non-speech stimuli. Ludlow et al. (2014) found that children with ASD showed significantly reduced MMN amplitudes in response to both words and pseudowords, compared to TD children. The effect was especially pronounced for meaningful speech stimuli, with reduced activation observed in the frontal and central parietal regions. Vlaskamp et al. (2017) found that compared to TD controls, children with ASD exhibited reduced MMN amplitudes in response to duration and frequency-duration deviants. The researchers also observed a relationship between reduced MMN amplitudes and greater social and communication difficulties (Vlaskamp et al., 2017). Although, a MEG study by Kasai et al. (2005) found no differences in mismatch field (MEG analog of MMN) amplitude, it found significantly prolonged mismatch field latencies in the left hemisphere during across-category vowel changes. These longer latencies were also associated with greater clinical severity of autism symptoms (Kasai et al., 2005). The observed reductions and delays in MMN responses likely reflect atypical neural mechanisms involved in early, automatic detection of auditory changes, including altered cortical connectivity and reduced efficiency in encoding sensory inputs (Ludlow et al., 2014). Reduced MMN amplitudes may indicate weaker sensory discrimination, while longer latencies suggest slower processing of speech cues critical for language (Ludlow et al., 2014).

Together, these studies suggest that MMN may index individual differences not only between diagnostic groups but also along a continuum of related traits. These findings raise the possibility that similar patterns of auditory processing may be observed in the broader population, where subclinical traits can be captured using dimensional measures such as the AQ. Therefore, investigating MMN responses in relation to AQ scores may offer insights into how auditory processing varies with broader cognitive-behavioural traits.

1.6 The Present Study

The present study investigated how individual differences in cognitive-behavioural traits, specifically those measured by AQ, influenced speech perception under conditions of phonetic ambiguity. Using a two-part design, the study examined both behavioural and neural responses to vowel-fricative-vowel (VFV) syllables that vary in their acoustic and contextual properties. Task 1 employed a two-alternative forced choice (2AFC) identification task to assess how participants categorized a /s/-/ʃ/ continuum, with fricative onsets cross-spliced with either /a/ or /u/ vowels. Task 2 assessed neural responses recorded using EEG during a passive oddball paradigm to measure MMN.

The study addressed the following research questions: (1) Do listeners exhibit categorical perception of fricatives, and does this change based on vowel context (/a/ vs. /u/)? (2) Does the influence of vowel context on fricative perception differ depending on individuals' AQ scores? (3) Does neural discrimination of speech sounds (as indexed by MMN) vary as a function of vowel context and AQ?

Based on prior work (Yu, 2019; Yu & Lee, 2014), it was expected that participants would exhibit categorical perception of the acoustic fricative continuum between /s/ and /ʃ/, showing a category shift in the identification response near the midpoint. It was also expected that if vowel context influenced categorization patterns, where fricatives preceding the rounded vowel /u/ would be more likely to be identified as /ʃ/ than those preceding /a/, this would indicate that identification was driven primarily by low-level acoustic cues. The prediction is consistent with spectral frequency lowering effect caused by articulatory properties of rounded vowels, which alter the acoustic structure of preceding fricatives (Mann & Repp, 1980; Yu, 2019; Yu & Lee, 2014). However, it was predicted that this context effect would be impacted by individual cognitive-behavioural traits, where participants with higher AQ scores would show either reduced or enhanced sensitivity to the acoustic vowel context, reflected in an earlier or later shift between /s/ and /ʃ/ in the perceptual boundary. Lastly, at the neural level, it was anticipated that individuals with lower and higher AQ scores would differ in MMN amplitudes and/or latencies to deviant syllables, correlated with their behavioural responses.

By examining these effects in a neurotypical population, this study aims to contribute to the growing body of research demonstrating that AQ traits influence perceptual and cognitive processes along a continuum. Investigating these traits outside of a clinical ASD diagnosis avoids potential diagnostic biases and allows for a more in-depth understanding of individual variability in auditory processing. These findings may have broader implications for models of speech perception and may also inform the design of auditory technologies that accommodate diverse perceptual strategies.

2. Methods

2.1 Participants

In this study, 52 undergraduate/graduate students from McMaster University were recruited (45 females, mean age = 20.2 years, SD = 1.6 years) (7 males, mean age = 20.6 years, SD = 2.7 years). Participants were recruited using three main methods: McMaster's Linguistics Research Participation System software (SONA), re-contacted from a previous study that had collected AQ scores as part of a Master's thesis project in the Phonetics Lab at the Centre of Advanced Research in Experimental and Applied Linguistics (ARiEAL), and responses to recruitment posters or emails circulated to faculty and students in the Linguistics and Languages or Psychology, Neuroscience, and Behaviour departments.

Participants underwent a screening process to ensure their eligibility prior to taking part in the experiment. All included participants were native English speakers or had started learning English by the age of 10, right-handed, had normal or corrected to normal vision and no history of neurological problems or hearing impairments. Meeting these criteria was based on participants' self-reports, and no formal screening was conducted.

Eligible participants completed a demographic questionnaire which collected general background information such as age, gender, current year of study at McMaster and their Faculty. Participants also completed a 50-question AQ questionnaire. To reduce any potential bias, the AQ questionnaire was referred to as a "cognitive-behavioural trait questionnaire" prior to participation. Participants were debriefed at the end of the experiment and informed about the true nature of the questionnaire.

All questionnaires were administered online, primarily in Microsoft Forms. The only exception was the screening questionnaire where SONA-recruited participants completed screening within the SONA system, while externally recruited participants completed the identical screening in Microsoft forms.

Written informed consent was obtained from all participants prior to data collection. Participants received either course credit (through SONA) or monetary compensation for their involvement. The study was approved by the McMaster Research Ethics Board (MREB).

2.2 Autism Spectrum Quotient (AQ)

To measure individual differences in traits commonly associated with autism, participants completed the AQ, a 50-item self-report questionnaire developed by Baron-Cohen et al. (2001b). The AQ is designed to measure the degree to which an individual exhibits traits associated with the autism spectrum and has been widely adopted in research to examine how such traits vary across the general population (Baron-Cohen et al., 2001b). In addition to the AQ, multiple assessment tools exist to capture cognitive and pragmatic processing styles, such as the Broad Autism Phenotype Questionnaire (Hurley et al., 2007) and the Reading the Mind in the Eyes test (Baron-Cohen et al., 2001a). The AQ was chosen for this study because it has been shown to correlate with speech perception and production measures in previous research (Bishop et al., 2020; Yu, 2010). Importantly, the AQ is not a diagnostic tool and does not indicate whether an individual is on the autism spectrum. Instead, it provides a continuous measure of traits that exist to varying degrees across all individuals.

The AQ includes five subscales: social skills, attention switching, attention to detail, communication and imagination, each comprising 10 questions (Baron-Cohen et al., 2001b). Items are phrased as statements such as “I prefer to do things with others rather than on my own,” and participants responded on a 4-point scale: “definitely agree,” “slightly agree,” “slightly disagree,” and “definitely disagree.” The original scoring method uses a binary system to assign one point for each response indicative of autistic-like traits (yielding scores between 0 and 50) (Baron-Cohen et al., 2001b). Following the approach used by Stewart & Ota (2008) and Yu (2010, 2016), the present study scored AQ responses using a Likert scale ranging from 1 to 4 (yielding scores between 50 and 200). The Likert scale method is used to preserve individual variability in trait expression. Items were reverse-coded (i.e. scoring was inverted for certain items) where necessary so that higher numerical responses consistently reflected a greater degree of autism-associated traits (Stewart & Ota, 2008; Yu, 2010; Yu, 2016). AQ scores were also standardized (z-scored) and used as both a continuous predictor variable (i.e. analyzed as a numeric variable on a continuous scale) and to form groups for Low versus High AQ based on a quartile split.

2.3 Stimuli

The stimuli used in this study were developed by Yu and Lee (2014) and consisted of /V_isV_i-V_iʃV_i/ syllables, where the vowel was either /a/ or /u/. The fricative portion was generated by digitally blending (i.e. interpolating) recordings of /sa/ and /ʃa/ produced by a male native English speaker. The endpoints of the continuum were the natural productions of /s/ and /ʃ/, while the five intermediate steps were created by incrementally mixing the waveforms. The resulting stimuli provided a smooth acoustic progression from the /s/ endpoint to the /ʃ/ endpoint. The intermediate steps consisted of ambiguous waveforms between /s/ and /ʃ/, blending features of both sounds to varying degrees, with the step 4 being the most ambiguous. The seven fricatives were then cross spliced with either an /a/ or /u/ preceding and following vowel segment, which was extracted from the same speaker's productions of /da/ and /du/ syllables, ensuring consistency in voice quality and articulation across the stimuli. A total of 14 audio stimuli were developed, consisting of seven /s-ʃ/ steps for each vowel context (/a/ and /u/).

2.4 Behavioural Task

Participants completed a 2-Alternative Forced Choice (2AFC) identification task conducted using the Gorilla software platform (Figure 1). Participants heard a single stimulus played on a laptop, one of the /V_isV_i - V_iʃV_i/ syllables, and were asked to categorize it as either /s/ or /ʃ/. Participants were given 3.5 seconds to respond after hearing the stimulus. A confidence rating was collected after each response on a scale of 1-10. There were 6 repetitions for each stimulus with a total of 84 trials (7 steps x 2 vowel contexts x 6 repetitions) with stimulus order randomized across participants.

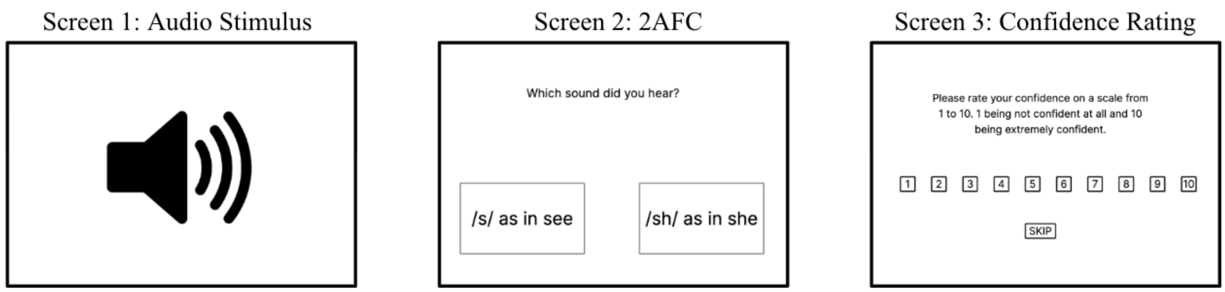

Figure 1.Example trial sequence for the two-alternative forced-choice (2AFC) behavioral task implemented in Gorilla. Screen 1: An audio stimulus (one of the vowel-fricative–vowel tokens) is presented. Screen 2: Participants select whether they heard /s/ ("see") or /ʃ/ ("she"). Screen 3: Participants rate their confidence in their identification on a scale from 1 (not confident) to 10 (extremely confident).

2.5 EEG Task

Participants underwent EEG recording to assess neural responses to the VFV stimuli in a passive oddball paradigm. Participants were instructed to watch a silent movie of their choice and ignore the auditory stimuli. The stimuli codes (e.g., 7a1) represent the VFV continuum (7-step continuum from /s/ to /ʃ/) and vowel context (/a/ or /u/), followed by the step number (1-7) along with continuum, with endpoints labeled as 1 and 7. The stimuli were presented using the Presentation software (Version 23.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com) and were played through ER-1 tubal insert earphones (Etymotic Research earphones designed for auditory presentation in EEG experiments). Two separate blocks were presented, one for the /a/ context (7a1 as standard, 7a4 as deviant) and one for the /u/ context (7u1 as standard, 7u4 as deviant), with the order counterbalanced across presentations. Each participant completed a total of 3000 trials, with 1200 (80%) being the standard and 300 being the deviant (20%) within each block of 1500 trials. Each sound had a duration of 800ms.

2.5.1 EEG Recordings

EEG data were recorded using a BioSemi ActiveTwo system with 64 Ag/AgCl electrodes arranged according to the international 10-20 system. Five external electrodes were used with one placed at each of the following sites: the nasion, the left and right mastoids, and above and beside the left eye to monitor motor movements and blinks. The Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes served as the reference and ground, respectively. Data had a sampling rate of 512 Hz and bandpass filtered at 0.1 to 100 Hz.

2.5.2 EEG Pre-Processing and Analysis

EEG data were processed using EEGLAB (Delorme & Makeig, 2004) in MATLAB (The MathWorks Inc., 2023). Data were epoched[2] from -200 ms to 500 ms relative to stimulus onset and filtered using a bandpass filter from 0.1 to 20 Hz. Independent Component Analysis (ICA) was used to identify and remove any ocular or motor movements. Bad channels were identified visually based on amplitude fluctuations and excessive noise. Such channels were interpolated using the spherical method. Data were referenced using the Reference Electrode Standardization Technique (REST) which re-references the data to a neutral reference at infinity (Yao, 2001; Dong et al., 2017). No baseline correction was applied, as preliminary checks showed that baseline correction artificially shifted the waveforms and risked distorting the data.

Event-related potentials were computed in ERPLAB (Lopez-Calderon & Luck, S. J., 2014). Difference waves (deviant minus standard) were calculated for each vowel context using pop_binoperator in ERPLAB (Lopez-Calderon & Luck, S. J., 2014). The MMN component was quantified as the mean amplitude between 100 and 250 ms post-stimulus onset at the fronto-central electrodes (Fz, FCz, Cz, FC1, FC2, F1, F2, C1, C2). Peak latency within the same window was also extracted for each participant.

[2] Epoching involves segmenting the data into short time windows around the onset of each stimulus presentation.

3. Results

3.1 AQ score

AQ scores for the total sample ranged from 95 to 164 (M = 124.8, SD = 16.8). Analyses were conducted using both continuous and group-based approaches to AQ. For group-based analyses, participants were categorized into Low and High AQ groups based on the top and bottom quartiles (25%), and those in the middle 50% of the distribution were excluded. This was done to see more pronounced speech perception differences for the population scoring higher/lower on the AQ scale and thus allow for clearer AQ comparisons (Krizic et al., 2024). Although 52 participants completed the behavioural task, 8 were excluded from the EEG analysis due to excessive artifacts (had an artifact rejection rate above 20%) or missing external electrodes, resulting in a reduced EEG sample of 44. To maintain consistency across task analyses for groups, quartile-based AQ groupings were determined using the EEG sample[3]. Within the EEG sample, AQ scores ranged from 95 to 155 (M= 122.8, SD = 16.3). The Low AQ group (n = 12) had scores ranging from 95 to 107 (M = 102.5, SD = 4.6) and included 10 female participants and 2 male participants. The High AQ group (n = 12) had scores from 136 to 155 (M = 143.1, SD = 4.6), consisting of 11 female participants and 1 male participant.

Figure 2 illustrates how the total AQ scores were distributed across the five subscales. On average, Attention Switching accounted for 20.3% (SD = 3.8) of total AQ scores, Attention to Detail 21.0% (SD = 3.8), Communication 19.9% (SD = 2.5), Imagination 19.4% (SD = 3.8), and Social Skills 19.5% (SD = 2.5). Although the subscales were relatively balanced overall, Attention Switching and Attention to detail showed greater variability across participants,

[3] Participants were categorized into AQ groups based on quartiles calculated from AQ scores in the EEG sample. The lowest quartile formed the “Low AQ” group and the highest quartile formed the “High AQ” group. These groupings were then consistently applied across behavioural and EEG analysis.

suggesting that these subscales may drive the individual and group level variation observed in AQ related differences.

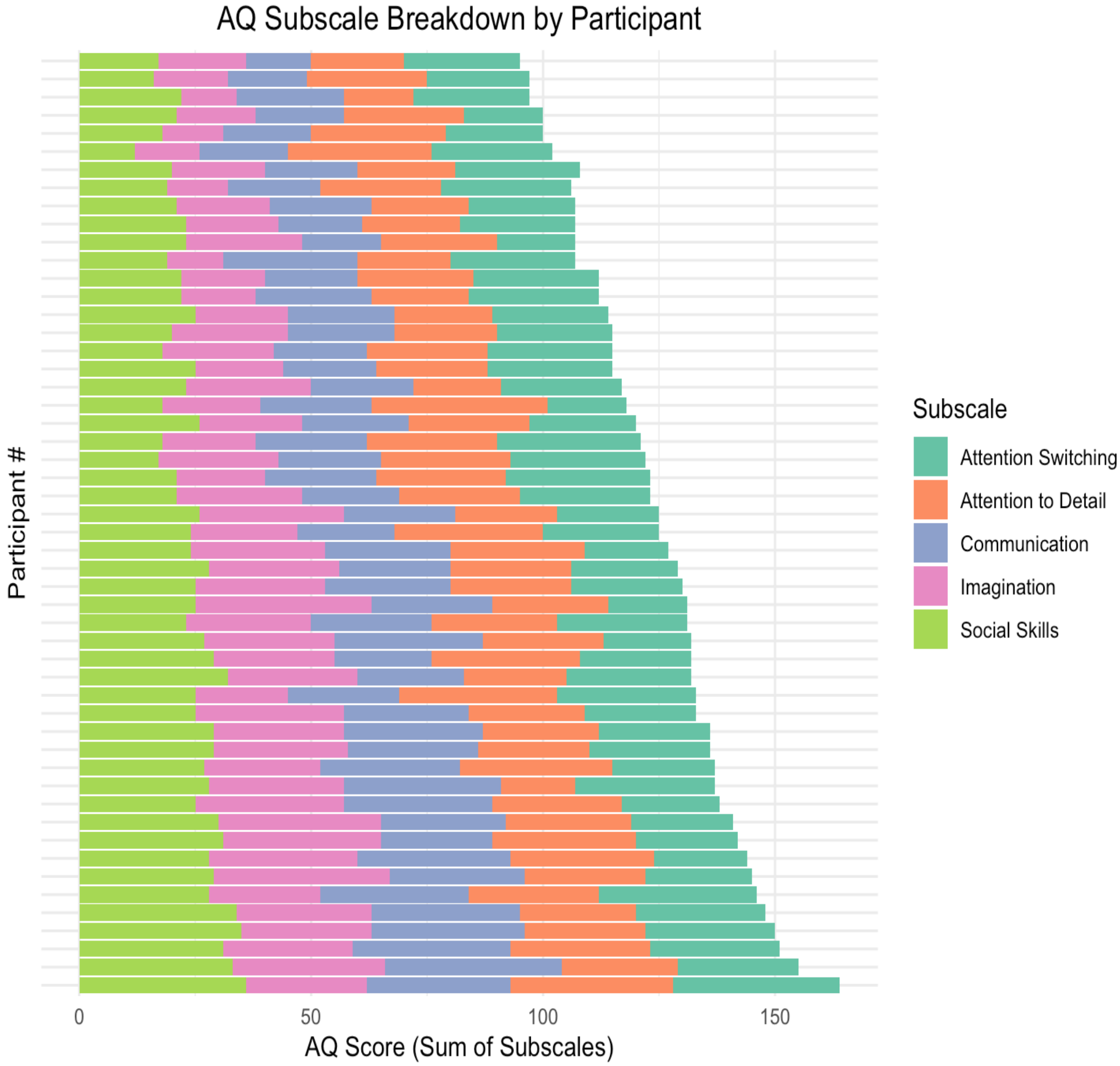

Figure 2. Autism-Spectrum Quotient (AQ) subscale scores by participant. Each horizontal bar represents one participant's total AQ score, broken down by the five subscales: Attention Switching, Attention to Detail, Communication, Imagination, and Social Skills. Scores are displayed as stacked segments, with the sum representing each participant's total AQ score.

3.2 Two-Alternative Forced Choice (2AFC) Identification Task

To examine categorization responses of participants in the 2AFC identification task, a logistic mixed-effects model was used. Analyses were implemented in R statistical software (R Core Team, 2022) using the *glmer()* function from the *lme4* package (Bates, Maechler, & Bolker, 2015). Responses were coded as 1 for /ʃ/ responses and 0 for /s/ responses (following Yu & Lee, 2014). Analyses included trial order (trial_z), continuum step (step_z for linear and $step^2$_z for quadratic), vowel context (/a/ vs. /u/), and AQ (modelled either continuously or as a categorical factor (High vs. Low AQ)). The quadratic term ($step^2$), as done by Yu (2019), was included to capture potential nonlinear response patterns. All continuous predictors were z-scored prior to analysis. Vowel was sum coded (-1 = /u/, +1 = /a/), such that the intercept reflected the grand mean across vowel contexts. Random slopes for step and $step^2$ were included by participant. The optimizer was set to *bobyqa* (Bates, Maechler, & Bolker, 2015) to improve convergence, and the *emmeans* package (Lenth et al., 2025) was used for estimated marginal means and pairwise contrasts.

3.2.1 Continuous AQ

Categorization responses were modelled as: response_bin ~ trial_z + vowel * (step_z + $step^2$_z) * AQ_z + (1 + step_z + $step^2$_z | participant). The model showed a strong main effect of continuum step ($\beta = 4.91$, p = < 2e-16), confirming participants reliably categorized /s/ versus /ʃ/ phonemes along the /s/-/ʃ/ continuum (Figure 3). Trial order exerted a significant effect ($\beta = -0.11$, $p = 0.039$), suggesting that participants produced slightly fewer /ʃ/ responses as the task progressed. A significant main effect of vowel was also observed ($\beta = 0.67$, p = < 2e-16), indicating that, overall, participants produced fewer /ʃ/ responses in the /u/ context relative to the /a/ context.

Estimated marginal means were computed for the /a/ and /u/ contexts. As emmeans are expressed on the logit (log-odds) scale, positive values indicate higher odds of producing a /ʃ/ response. The estimated log-odds for /a/ (M = 1.934, SE = 0.348, 95% CI (Confidence Interval) [1.252, 2.615]) were greater than those for /u/ (M = 0.311, SE = 0.343, 95% CI [-0.361, 0.983]).

The pairwise contrast (a-u = 1.62, SE = 0.15, z = 10.78, p = < 0.0001) confirmed that participants were significantly more likely to categorize stimuli as /ʃ/ in the /a/ context compared to /u/.

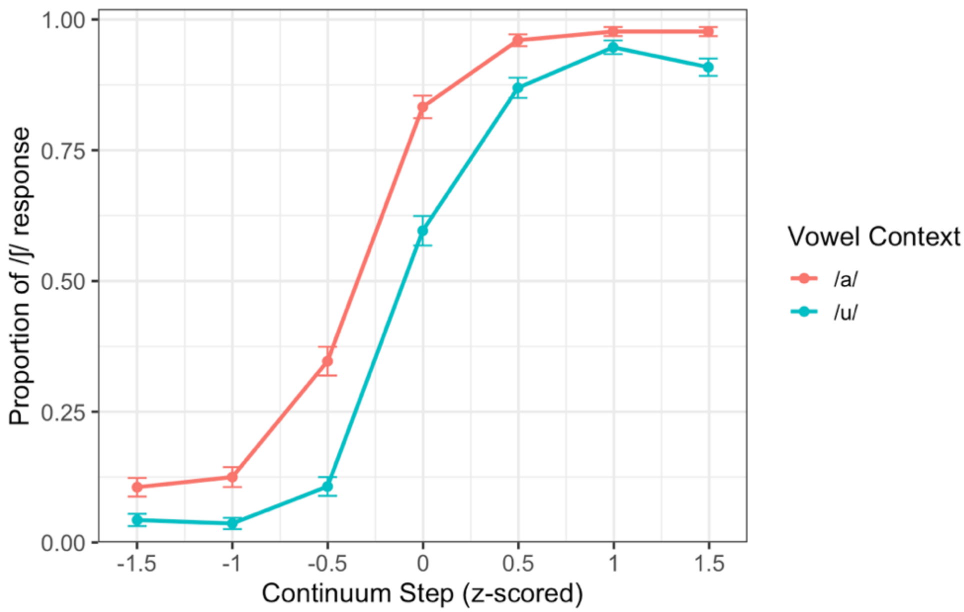

Figure 3. Proportion of /ʃ/ responses across z-scored continuum steps, plotted by vowel context. Red lines represent the /a/ vowel context, and blue lines represent the /u/ vowel context. Error bars represent standard errors.

Additionally, AQ_z was a significant predictor (β = -0.32, p = 0.027), such that individuals with higher AQ scores produced fewer /ʃ/ responses overall across both vowel contexts. The quadratic term ($Step^2$_z: β = 0.09, p = 0.778) and interactions with AQ or vowel did not reach significance.

Overall, these results suggest that categorization was influenced by vowel context and AQ score, with AQ effects manifesting as a response shift towards /ʃ/ rather than as interactions with step or vowel.

Table 1. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with AQ as a continuous predictor.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	1.03081	0.14533	7.093	1.31e-12 ***
Trial_z	-0.11393	0.05511	-2.067	0.0387 *
Vowel$_u$	0.67274	0.07185	9.364	< 2e-16 ***
Step_z	4.90966	0.44201	11.108	< 2e-16 ***
Step^2_z	0.09135	0.32293	0.283	0.7773
AQ_z	-0.32059	0.14450	-2.219	0.0265 *
Vowel$_u$:Step_z	0.02156	0.08618	0.250	0.8024
Vowel$_u$:Step^2_z	0.13840	0.08450	1.638	0.1014
Vowel$_u$:AQ_z	0.01833	0.07383	0.248	0.8039
Step_z:AQ_z	-0.12006	0.41318	-0.291	0.7714
Step^2_z:AQ_z	0.33831	0.20893	1.619	0.1054
Vowel$_u$:Step_z:AQ_z	0.06736	0.10232	0.658	0.5103
Vowel$_u$:Step^2_z:AQ_z	-0.08986	0.09899	-0.908	0.3640

Note: Responses were coded as 1 = /ʃ/, 0 = /s/. Vowel was sum coded (-1 = /u/, +1 = /a/). Step_z is the z-scored linear continuum. Step^2_z is the quadratic term. Trial_z is the z-scored trial order. AQ_z is the z-scored Autism Quotient scores. Significance codes: *** p < .001, ** p < .01, * p < .05.

3.2.2 Group-Based AQ

To further assess whether categorization patterns differed across levels of AQ (i.e. High AQ vs. Low AQ quartiles), analyses were also conducted at the group level. A logistic mixed-effects model with the same fixed-effects structure was fit using AQ group: response_bin ~ trial_z + vowel * (step_z + step^2_z) * AQ_group + (1 + step_z + step^2_z | participant). For the group-based analysis, AQ group was sum-coded (-1 = Low, +1 = High).

The grouped analysis showed the same significant effect of step ($\beta = 4.97$, $p < 5e\text{-}15$). A significant main effect of vowel was also observed ($\beta = 0.62$, $p = 2.34e\text{-}09$), with both groups producing more /ʃ/ responses in the /a/ context than in /u/. Estimated marginal means confirmed this within-group pattern: High AQ (a-u = 1.35, SE = 0.28, *p* = < 0.0001, OR (odds ratio) = 3.87, 95% CI [2.23, 6.72]) and Low AQ (a-u = 1.11, SE = 0.34, *p* = 0.0012, OR = 3.04, 95% CI [1.55, 5.97]).

A main effect of AQ group was also found (β = -0.58, p = 0.0017). Estimated marginal means averaged over vowels showed higher log-odds of a /ʃ/ response for the Low AQ group (M = 2.02, SE = 0.76, 95% CI [0.54, 3.49]) than High AQ (M = 1.10, SE = 0.51, 95% CI [0.10, 2.10]), though the direct pairwise contrast was not significant (High-Low = -0.912, SE = 0.791, z = -1.153, p = 0.249). No significant interactions involving AQ group and vowel or step were observed.

Together, these results indicate that vowel context influenced categorization patterns in both AQ groups, while individuals in the Low AQ group tended to produce more /ʃ/ responses relative to those in the High AQ Group, although the direct pairwise comparison was not significant.

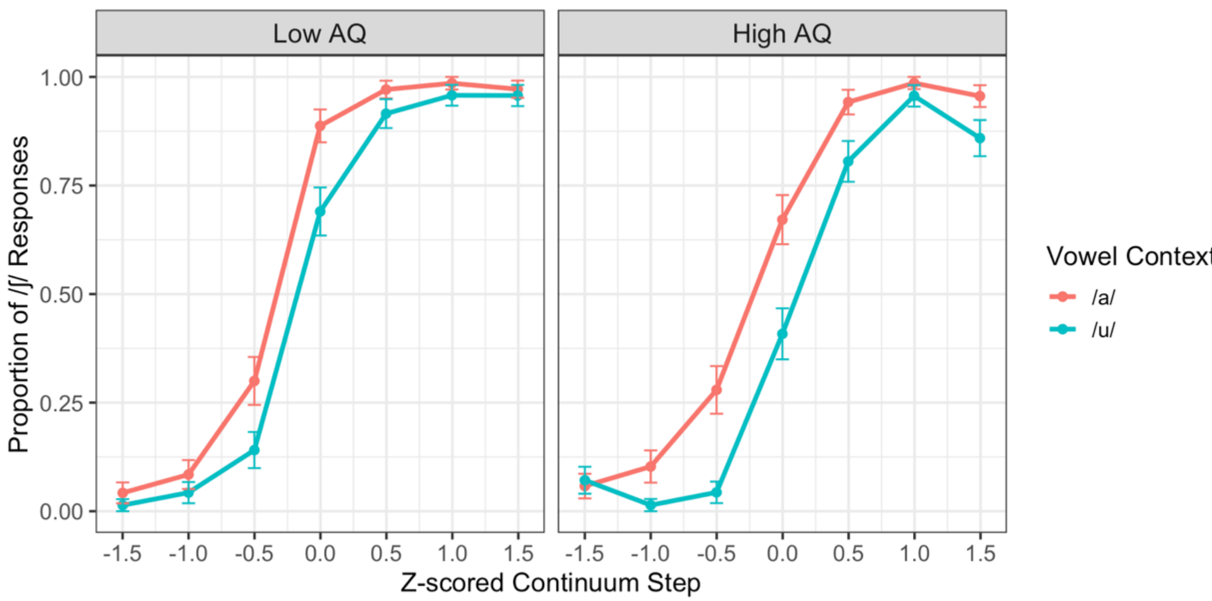

Figure 4. Categorization curves for Low and High AQ groups, plotted by vowel context. Panels show identification curves for the Low AQ group (left) and High AQ group (right). Red lines represent the /a/ vowel context and blue lines represent the /u/ vowel context. Error bars represent standard errors.

Table 2. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with AQ grouped.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	0.80307	0.18774	4.278	1.89e-05 ***
Trial_z	-0.06701	0.08150	-0.822	0.41098
$Vowel_u$	0.62319	0.10435	5.972	2.34e-09 ***
Step_z	4.97177	0.63570	7.821	5.24e-15 ***
$Step^2$_z	0.77272	0.46999	1.644	0.10015
AQ_Group_{Low}	-0.57665	0.18388	-3.136	0.00171 **
$Vowel_u$:Step_z	0.09919	0.12719	0.780	0.43548
$Vowel_u$:$Step^2$_z	-0.00617	0.12756	-0.048	0.96145
$Vowel_u$:AQ_Group_{Low}	0.06463	0.10314	0.627	0.53092
Step_z:AQ_Group_{Low}	-0.84413	0.57983	-1.456	0.14544
$Step^2$_z:AQ_Group_{Low}	0.11861	0.35023	0.339	0.73487
$Vowel_u$:Step_z:AQ_Group_{Low}	0.07894	0.12694	0.622	0.53402
$Vowel_u$:$Step^2$_z:AQ_Group_{Low}	-0.00415	0.12656	-0.033	0.97383

Note: Responses were coded as 1 = /ʃ/, 0 = /s/. Vowel was sum coded (-1 = /u/, +1 = /a/). AQ Group was sum coded (-1 = Low, +1 = High). Step_z is the z-scored linear continuum. $Step^2$_z is the quadratic term. Trial_z is the z-scored trial order. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

3.2.3 Confidence Rating

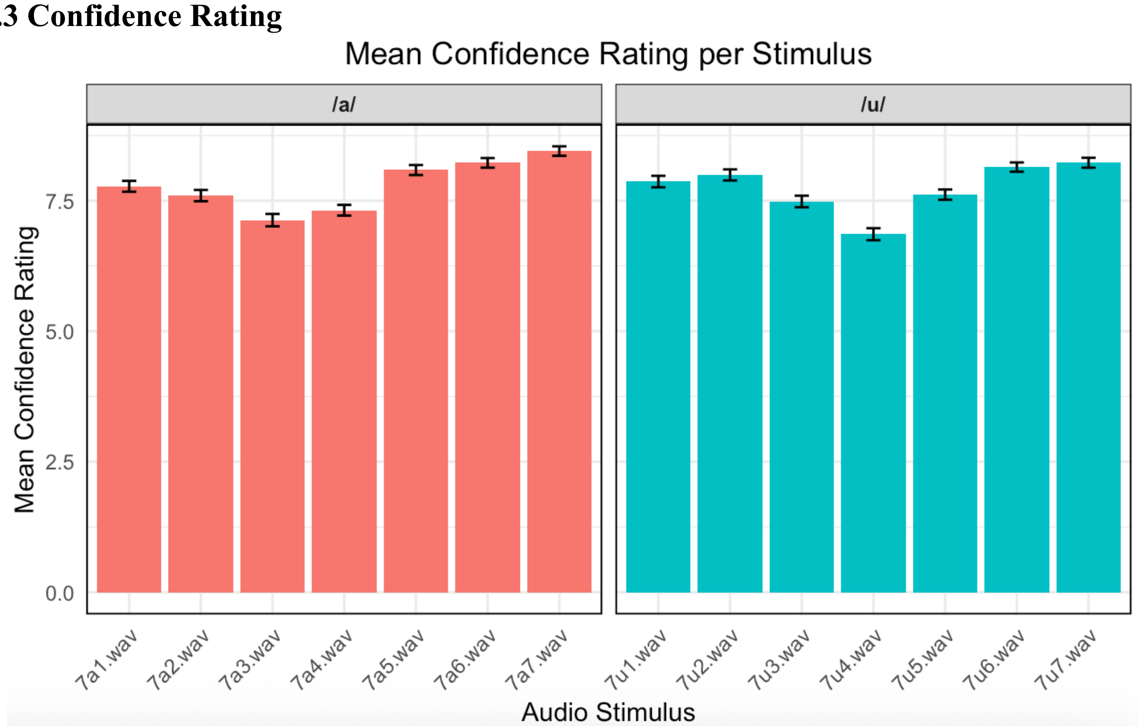

Figure 5. Mean confidence ratings for each audio stimulus in the 2AFC task, separated by vowel context. Red bars represent stimuli from the /a/ context and blue bars represent stimuli from the /u/ context. Error bars represent standard errors.

Figure 5 shows the mean confidence ratings for each stimulus presented in the 2AFC task, separated by vowel context. Confidence ratings were generally high across the continuum but tended to dip near the category boundary. At step 3, confidence was slightly higher in the /u/ context compared to the /a/ context, whereas at step 4 this pattern reversed, with /a/ showing higher confidence than /u/.

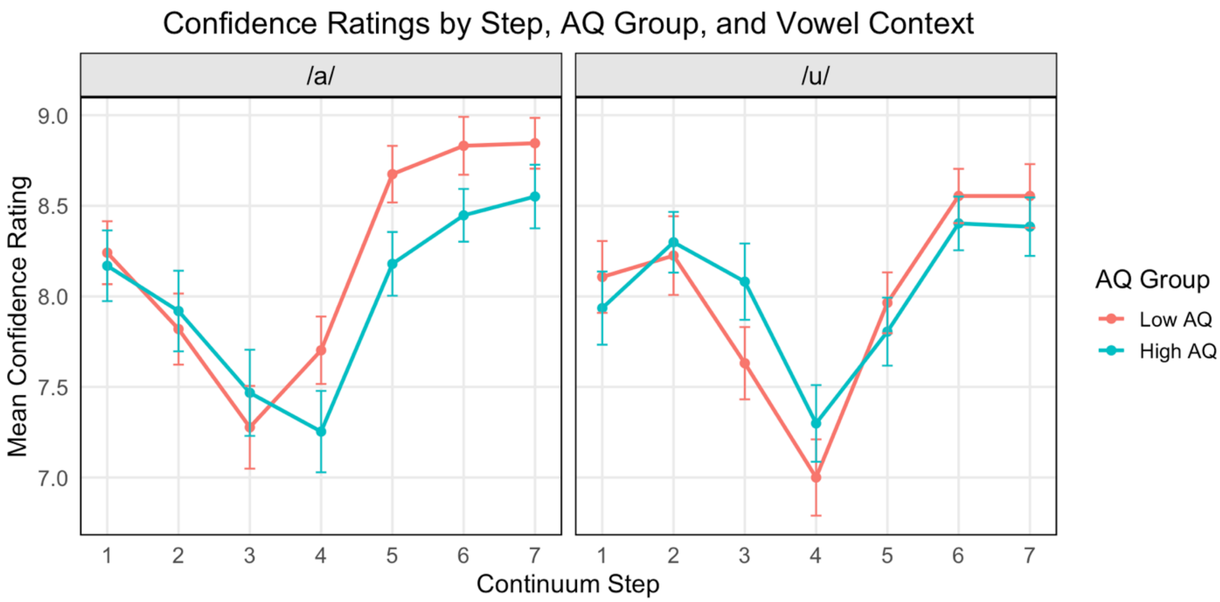

Figure 6. Mean confidence ratings across continuum steps by AQ group and vowel context. Panels show the /a/ (left) and /u/ (right) vowel contexts separately. Red lines represent the Low AQ group, and blue lines represent the High AQ group. Error bars represent standard errors.

As illustrated in Figure 6, Low and High AQ groups displayed a crossover in the /u/ context compared to /a/ at step 4. This suggests a potential vowel-dependent reversal in confidence ratings between groups.

Overall, confidence ratings varied with vowel context and position along the continuum. These descriptive patterns highlight greater variability in confidence near the category boundary.

3.2.4 Point of Subjective Equality (PSE)

To assess whether individual differences in AQ scores influenced phoneme categorization boundaries, perceptual boundary locations (PSEs) were estimated for each participant and vowel context using the *quickpsy* (Linaries & Lopez-Moliner, 2016) package in R (R Core Team, 2022). PSEs identify the 50% point where /s/ and /ʃ/ responses are balanced. From a phonetic standpoint, the 50% point defines the perfect ambiguity region where the listener would categorize the presented stimuli as equally likely being a /s/ or /ʃ/ phoneme. A logistic psychometric function was fit to each individual's identification data as a function of the z-scored continuum step, separately by vowel context. The resulting PSE values were then analyzed using two linear mixed-effects models, one with continuous AQ scores (AQ_z) and one with AQ group (Low vs. High) as predictors. Both models included vowel context (sum coded: -1 = /u/, +1 = /a/), the AQ predictor (either AQ_z or AQ group), and their interactions as fixed effects. Random intercepts for participant were also included.

In the continuous AQ model, no significant effects of vowel ($\beta = 0.50$, $p = 0.496$), AQ_z ($\beta = 0.55$, $p = 0.492$), or their interaction ($\beta = -0.42$, $p = 0.568$) were found. Figure 7 displays the relationship between z-scored AQ scores and PSEs.

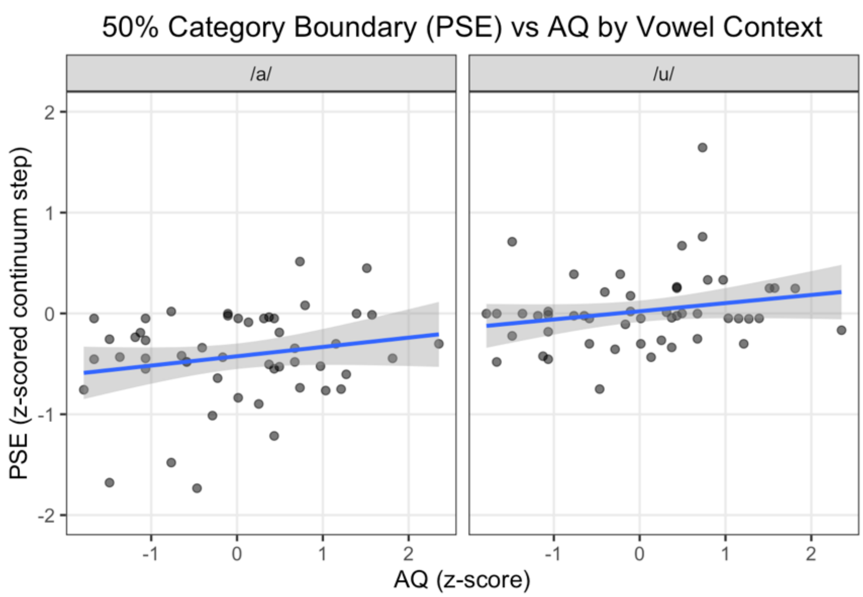

Figure 7. Relationship between z-scored AQ scores and point of subjective equality (PSE) values for the /a/ (left) and /u/ (right) vowel contexts. PSE values represent the 50% category boundary along the fricative continuum. Each point corresponds to an individual participant, with the blue line showing the fitted regression line and shaded areas indicating the confidence interval.

Table 3. Effects of vowel context and continuous AQ (AQ_z) on PSE.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	-0.9811	0.7863	50.0000	-1.248	0.218
$Vowel_u$	0.5025	0.7319	50.0000	0.686	0.496
AQ_z	0.5471	0.7900	50.0000	0.693	0.492
$Vowel_u$:AQ_z	-0.4230	0.7354	50.0000	-0.575	0.568

Note: PSE derived from logistic fits to categorization curves. Vowel was sum coded (-1 = /u/, +1 = /a/). AQ_z is the z-scored Autism Quotient scores. Significance codes: *** p < .001, ** p < .01, * p < .05.

In the group-based model, AQ group was sum coded (-1 = Low, +1 = High). The PSE showed a main effect of vowel (β = 0.22, p = 9.0e-04) and a significant effect of AQ group (β = 0.12, p = 0.024). The vowel and AQ interaction was not significant (β = -0.04, p = 0.505).

Estimated marginal means clarified the nature of these effects. For both groups, the PSE occurred at an earlier point along the continuum in the /a/ context than in the /u/ context (High AQ: a-u = -0.524, SE = 0.175, p = 0.007, 95% CI [-0.887, -0.161]; Low AQ: a-u = -0.357, SE = 0.175, p = 0.054, 95% CI [-0.720, 0.006]). Group comparisons further showed that PSE values between groups did not differ in the /a/ context (High-Low = 0.208, SE = 0.175, p = 0.243, 95% CI [-0.145, 0.561]). However, in the /u/ context, the High AQ group showed significantly higher PSE values than the Low AQ group (High-Low = 0.374, SE = 0.175, p = 0.039, 95% CI [0.021, 0.727]). Higher PSE values correspond to a boundary shift later along the continuum toward /ʃ/ (i.e. High AQ participants placed the category boundary further toward /ʃ/, while Low AQ participants placed it closer to /s/). As shown in Figure 8, vowel context impacted group differences in categorization. For stimuli in the /a/ context, the Low and High AQ groups did not differ in boundary placement. However, in the /u/ context, the High AQ exhibited a boundary shift toward /ʃ/ (i.e. they were more likely to categorize stimuli as /ʃ/ even when the sounds were closer to the /s/ prototype).

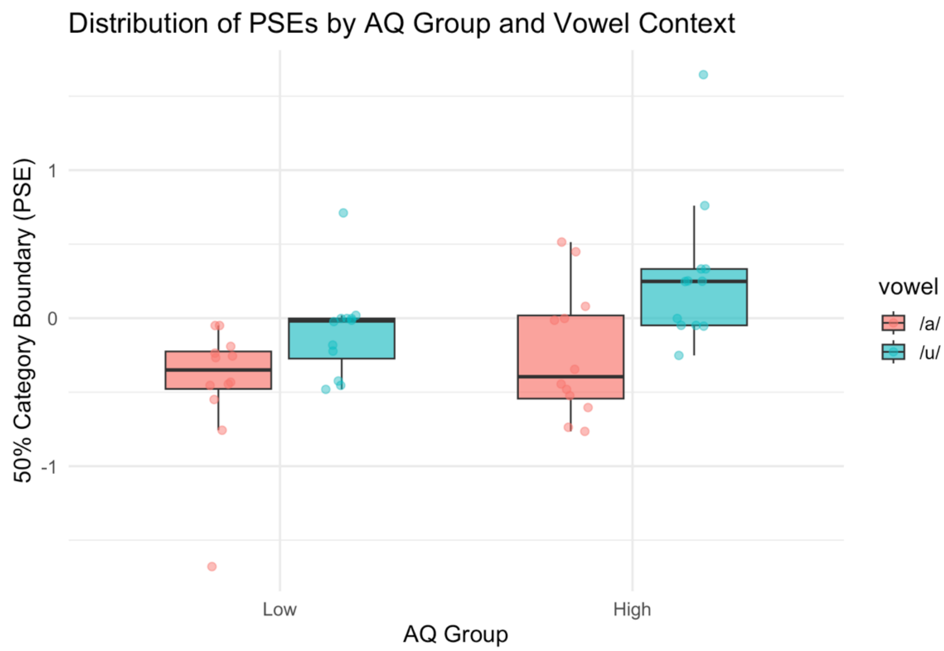

Figure 8. Distribution of phoneme categorization boundaries (PSEs) by AQ group (Low vs. High) and vowel context (/a/ vs. /u/). PSE values represent the 50% category boundary along the fricative continuum. Boxplots show the median, interquartile range, and outliers for the Low AQ and High AQ groups in the /a/ (red) and /u/ (blue) vowel contexts.

Table 4. Effects of vowel context and AQ group (Low vs. High) on PSE.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	-0.12298	0.06200	44.00000	-1.983	0.05359 .
$Vowel_u$	-0.22036	0.06200	44.00000	-3.554	0.00092 ***
AQ_Group_{Low}	0.14546	0.06200	44.00000	2.346	0.02355 *
$Vowel_u$:AQ_Group_{Low}	-0.04165	0.06200	44.00000	-0.672	0.50531

Note: PSE derived from logistic fits to categorization curves. Vowel was sum coded (-1 = /u/, +1 = /a/). AQ group was sum coded (-1 = Low, +1 = High). Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Parallel models were run on the slope of categorization functions, however, no significant differences were found other than the intercept, indicating steep categorization curves.

Table 5. Effects of vowel context and continuous AQ (AQ_z) on categorization slope.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	25.3109	4.5905	50.0000	5.514	1.23e-06 ***
$Vowel_u$	-1.1106	3.1258	50.0000	-0.355	0.724
AQ_z	-0.6971	4.6120	50.0000	-0.151	0.881
$Vowel_u$:AQ_z	5.8533	3.1404	50.0000	1.864	0.068 .

Note: Vowel was sum coded (-1 = /u/, +1 = /a/). AQ_z is the z-scored Autism Quotient scores. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

Table 6. Effects of vowel context and AQ group (Low vs. High) on categorization slope.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	24.3099	7.1629	22.0000	3.394	0.0026 **
$Vowel_u$	0.6935	2.2591	22.0000	0.307	0.762
AQ_Group_{Low}	-9.7475	7.1629	22.0000	-1.361	0.187
$Vowel_u{:}AQ\_Group_{Low}$	2.7143	2.2591	22.0000	1.202	0.242

Note: Vowel was sum coded (-1 = /u/, +1 = /a/). AQ group was sum coded (-1 = Low, +1 = High). Significance codes: *** p < .001, ** p < .01, * p < .05.

3.3 MMN Amplitudes

To examine the relationship between MMN amplitudes and individual variation in AQ scores, linear mixed-effects models were fit to mean amplitude values extracted from the difference wave (deviant minus standard) within the MMN time window (100-250 ms post stimulus onset). All analyses were conducted in R statistical software (R Core Team, 2022) using the *lmer()* (Bates, Maechler, & Bolker, 2015) function from the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017) with restricted maximum likelihood (REML) estimation. The optimizer was set to *bobyqa* (Bates, Maechler, & Bolker, 2015) to improve convergence, and the *emmeans* package (Lenth et al., 2025) was used for estimated marginal means and pairwise contrasts. MMN amplitudes (value) were modelled as a function of vowel context, standardized AQ score (z-scored) (AQ_z), scalp region (frontal vs. central), and their interactions: value ~ vowel context * AQ_z + region + (1 | ERPset). ERPset was included as a random intercept to account for participant-level variability in overall ERP amplitude. In this model, vowel context was sum coded (-1 = /u/ context, +1 = /a/ context). Electrodes were grouped into broader scalp regions and sum coded (-1 = frontal, +1 = central) for analysis to avoid convergence issues and reduced interpretability associated with using individual electrodes as fixed or random effects.

3.3.1 Continuous AQ

Although MMN is typically characterized as a negative deflection peaking around 100-250 ms, the observed difference waves in this study were positive in polarity. Therefore, while the observed effects fall within the MMN time window, they may not reflect a typical MMN and are interpreted as MMN-like responses. The model displayed a significant intercept ($\beta = 0.105$, $p = 0.005$), indicating a reliable ERP response within the MMN time window. No other significant main effects of vowel context ($\beta = 0.006$, $p = 0.647$), AQ_z ($\beta = 0.008$, $p = 0.824$), or region ($\beta = 0.007$, $p = 0.586$) were observed. The vowel context and AQ interaction was also nonsignificant ($\beta = -0.011$, $p = 0.381$).

Estimated marginal means were computed for both vowel contexts and represented the model-predicted ERP amplitude in microvolts (µV). Reliable ERP amplitudes were observed in both vowel contexts: the /a/ context (M = 0.111, SE = 0.037, 95% CI [0.036, 0.186], $p = 0.0046$) and the /u/ context (M = 0.099, SE = 0.037, 95% CI [0.024, 0.174], $p = 0.0106$) were each significantly different from zero. However, pairwise comparison between vowel contexts was not significant (/a/-/u/ = 0.01, SE = 0.03, 95% CI [-0.038, 0.061], $p = 0.647$).

Together, these results indicate that participants produced reliable ERP responses within the MMN time window in both vowel contexts, but that the magnitude of this response did not differ across vowel context, scalp region, or AQ score.

Table 7. Linear mixed-effects model estimates for ERP response amplitude as a function of AQ score, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.104828	0.035243	43.389321	2.974	0.00478 **
$VowelContext_u$	0.005804	0.012667	745.000000	0.458	0.64692
AQ_z	0.007825	0.034979	42.000000	0.224	0.82408
$Region_{Frontal}$	-0.007324	0.013435	745.000000	-0.545	0.58583
$VowelContext_u$:AQ_z	-0.011101	0.012675	745.000000	-0.876	0.38140

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

3.3.2 Group-Based AQ

Figure 9 presents grand average ERP waveforms at electrode FCz for the standard and deviant stimuli as well as the difference wave. Figure 10 displays scalp topographies of the MMN difference wave, averaged over the 100-250 ms time window. To assess whether the ERP response differed across AQ subgroups, a separate linear mixed-effects model was fit using AQ group as a categorical predictor that was sum coded (-1 = Low, +1 = High). The model formula was: value ~ vowel context * AQ_group + region + (1 | ERPset).

The model showed a significant intercept (β = 0.115, p = 0.01), indicating that both groups had reliable ERP responses with the MMN time window. A significant vowel context and AQ interaction was also observed (β = -0.040, p = 0.018), suggesting that the difference between the /a/ and /u/ context varied by AQ group. No other main effects were significant, meaning that region and AQ group alone did not impact ERP response amplitude differences.

Estimated marginal means clarified this interaction. As shown in Figure 11, for the High AQ group, /u/ context (M = 0.213, SE = 0.062, 95% CI [0.086, 0.340], p = 0.002) elicited a larger (more positive) response amplitude than the /a/ context (M = 0.074, SE = 0.062, 95% CI [-0.053, 0.201], p = 0.246). The pairwise comparison confirmed that for the High AQ group /u/ context elicited significantly larger amplitudes than the /a/ context (/a/-/u/ = -0.139, SE = 0.048, 95% CI [-0.233, -0.046], p = 0.004).

For the Low AQ group, neither the /a/ context (M = 0.096, SE = 0.062, 95% CI [-0.031, 0.224], p = 0.132) or /u/ context (M = 0.076, SE = 0.062, 95% CI [-0.051, 0.203], p = 0.231) was significantly different from zero. No reliable difference was observed between contexts (/a/-/u/ = 0.020, SE = 0.048, 95% CI [-0.073, 0.114], p = 0.670).

Group comparisons averaged across vowel contexts indicated larger amplitudes for High AQ (M = 0.143, SE = 0.058, 95% CI [0.024, 0.263]. *p* = 0.021) than Low AQ (M = 0.086, SE = 0.058, 95% CI [-0.033, 0.206], *p* = 0.148). However, this difference was not significant when directly tested (High-Low = 0.057, SE = 0.081, 95% CI [-0.111, 0.225], *p* = 0.489).

These results suggest that High AQ participants produced a reliable ERP response within the MMN time window, characterized by a stronger response amplitude in the /u/ context than the /a/ context amplitude. On the other hand, Low AQ participants did not display this response difference across contexts.

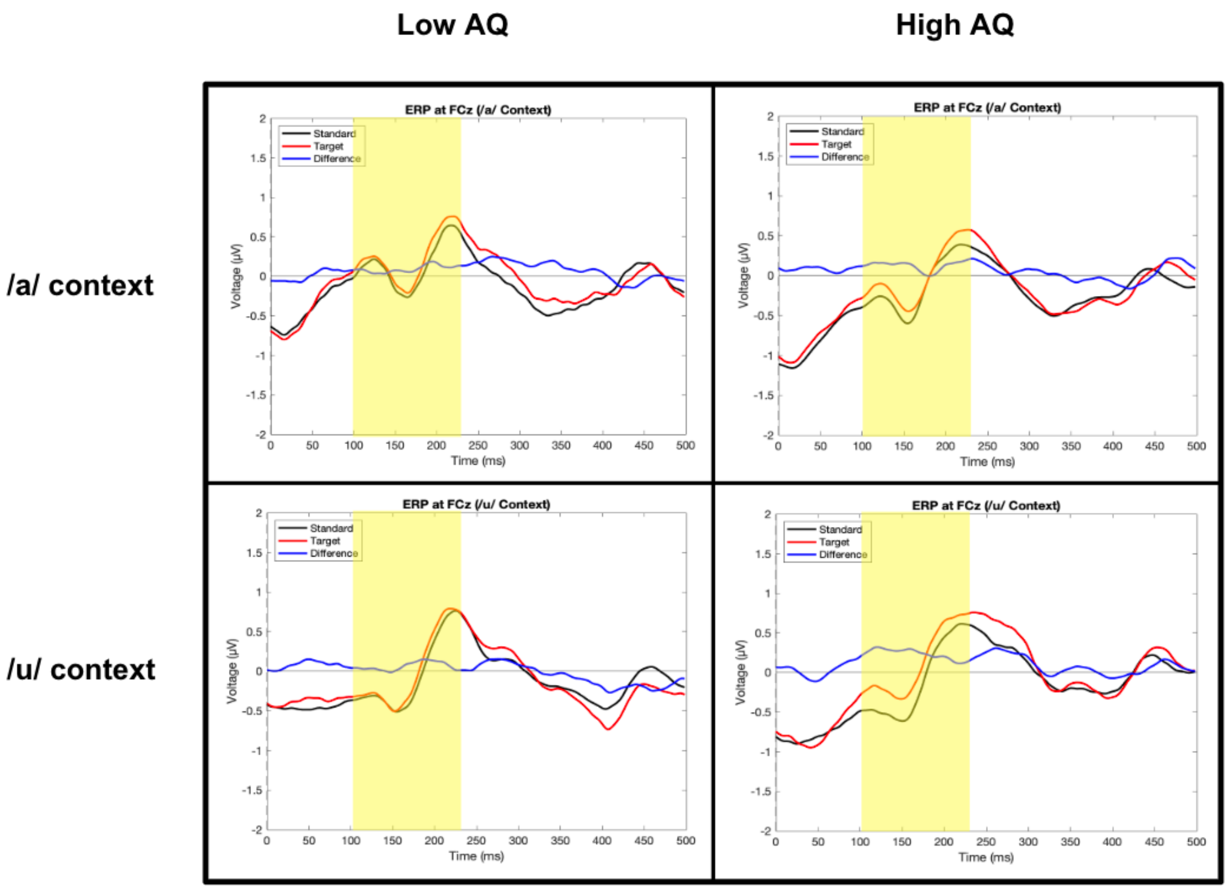

Figure 9. Event-related potentials (ERPs) at electrode FCz for the /a/ (top row) and /u/ (bottom row) vowel contexts in Low AQ (left column) and High AQ (right column) groups. Waveforms represent the grand average across participants within each group. Black lines represent standard stimuli, red lines represent deviant stimuli, and blue lines represent the difference wave (deviant minus standard). The shaded yellow region indicates the MMN analysis time window (100–250 ms). Time is shown in milliseconds (ms), and voltage is plotted in microvolts (μV). Negative voltage is plotted downwards.

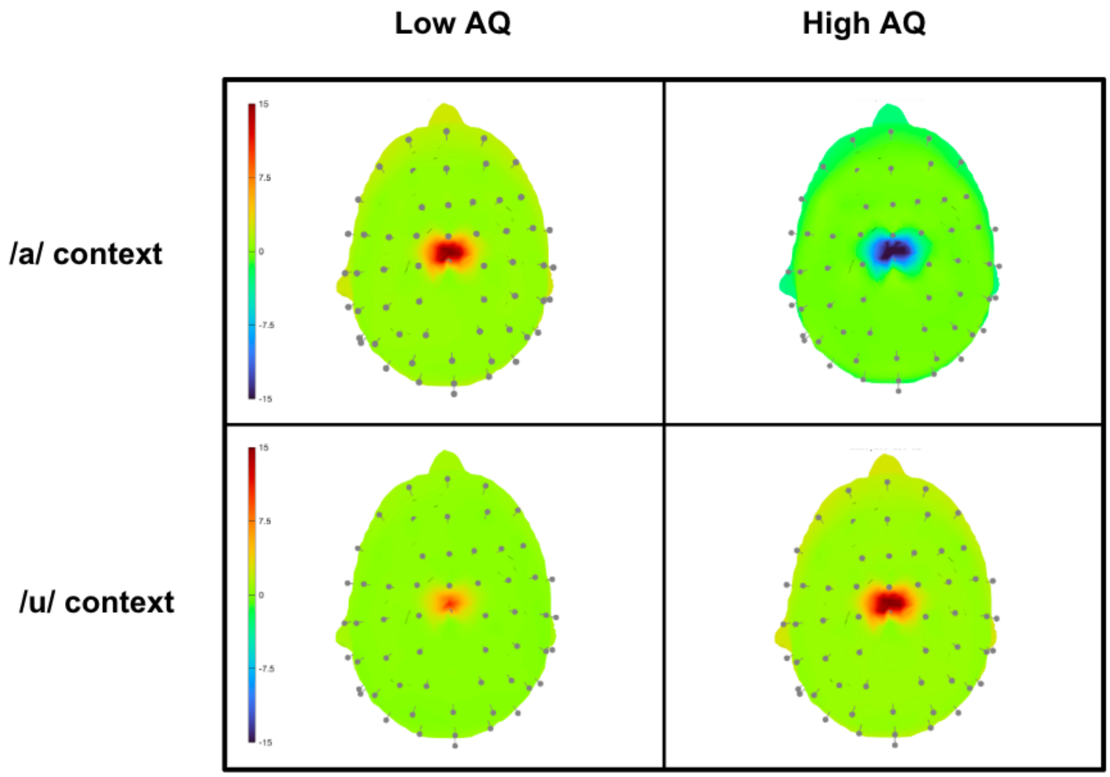

Figure 10. Scalp topographies of the MMN difference wave (deviant minus standard) in the 100–250 ms time window for the /a/ (top row) and /u/ (bottom row) vowel contexts, shown separately for Low AQ (left column) and High AQ (right column) groups. Maps represent the grand average across participants within each group. Warmer colours (red/orange/yellow) indicate positive voltages, and cooler colours (shades of blue) indicate negative voltages.

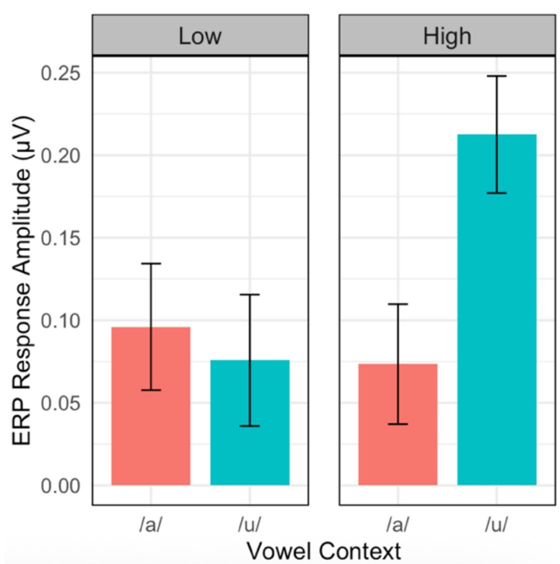

Figure 11. Mean ERP response amplitudes (µV) in the 100–250 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ (left) and High AQ (right) groups. Error bars represent standard errors.

Table 8. Linear mixed-effects model estimates for ERP response amplitude as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.114802	0.040971	22.955313	2.802	0.0101 *
$VowelContext_{u}$	-0.029699	0.016810	405.000000	-1.767	0.0780
AQ_Group_{Low}	0.028537	0.040538	22.000000	0.704	0.4888
$Region_{Frontal}$	0.001115	0.017829	405.000000	0.063	0.9502
$VowelContext_{u}$:AQ_Group_{Low}	-0.039833	0.016810	405.000000	-2.370	0.0183 *

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

3.4 MMN Latency

The same linear mixed-effects modelling approach used for ERP response amplitudes was applied to ERP response latency. In both the continuous and group-based analyses, ERP response latency, defined as the time point (in ms) at which the peak of the difference wave occurred within the MMN window (100 -250 ms post-stimulus), was modelled using the same fixed effects structure: vowel context, AQ score (either standardized AQ_z or grouped AQ_group), scalp region (region), and their interaction. All model specifications, including factor coding, optimizer settings, and region grouping, were consistent with the amplitude analyses.

3.4.1 Continuous AQ

The model displayed a significant intercept (β = 175.36, p = < 2e-16), indicating a predicted peak latency of approximately 175 ms. Figure 12 displays how ERP response latency shifts with standardized AQ for both vowel contexts. When AQ was modelled as a continuous predictor no significant main effects were observed for vowel context (β = -2.12, p = 0.130), AQ score (β = 3.62, p = 0.273, or region (β =-0.67, p = 0.651). There was also no significant interaction between AQ and vowel context (β = -2.59, p = 0.065).

Estimated marginal means were computed for both vowel contexts and represented the model-predicted ERP latencies in milliseconds (ms). ERP response latencies were predicted to be approximately 173 ms for the /a/ context (M = 173, SE = 3.58, 95% CI [166, 180], p = < 0.0001) and 177 ms for the /u/ context (M = 177, SE = 3.58, 95% CI [170, 185], p = < 0.0001) (significance values indicated that ERP latencies were significantly different from zero). The pairwise comparison between vowel contexts was not significant (/a/-/u/ = -4.24, SE = 2.80, 95% CI [-9.74, 1.26], p = 0.130).

These results confirm a reliable ERP response around 175 ms but provide no strong evidence for latency differences across vowel contexts, regions or AQ levels.

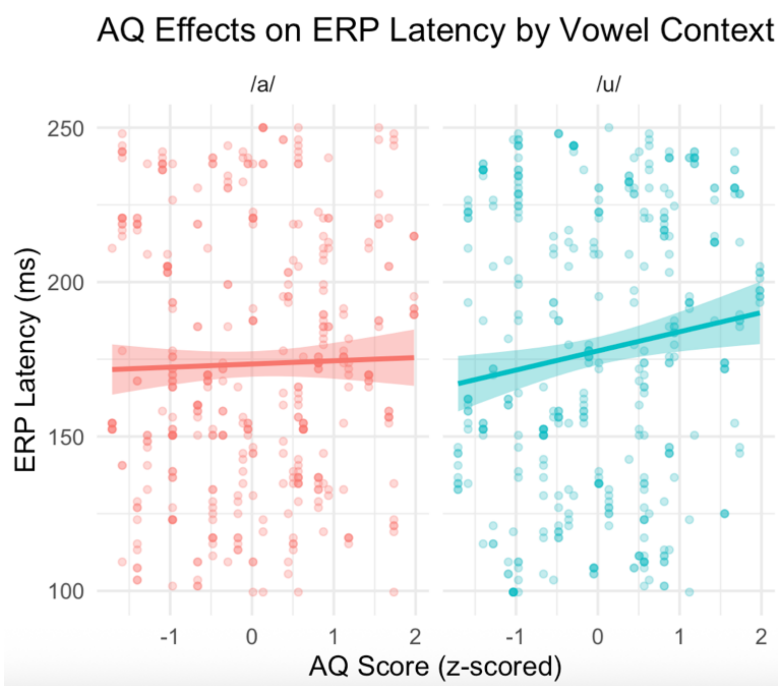

Figure 12. Relationship between z-scored AQ scores and ERP latency (ms) for the /a/ (left) and /u/ (right) vowel contexts. Each point represents an individual trial, with solid lines showing the fitted regression line and shaded areas indicating the confidence interval.

Table 9. Linear mixed-effects model estimates for ERP Response latency as a function of AQ score, vowel context and scalp region

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	175.3596	3.2943	43.9628	53.231	<2e-16 ***
VowelContext_u	-2.1208	1.4005	745.0000	-1.514	0.1304
AQ_z	3.6217	3.2589	42.0000	1.111	0.2727
$\text{Region}_{\text{Frontal}}$	-0.6732	1.4855	745.0000	-0.453	0.6505
VowelContext_u:AQ_z	-2.5860	1.4014	745.0000	-1.845	0.0654

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

3.4.2 Group-Based AQ

Figure 13 shows the latency differences across AQ groups and vowel contexts. A significant main effect of vowel context was observed (β = -3.85, p = 0.038). No significant effect of AQ group (β = 3.49, p = 0.417), region (β = 0.48, p = 0.804) or interaction between vowel context and AQ group was observed (β = 3.23, p = 0.080).

Estimated marginal means showed that in the High AQ group, latencies were significantly longer in the /u/ context (M = 192, SE = 6.56, 95% CI [179, 205], p = < 0.0001) than /a/ (M =178, SE = 6.56, 95% CI [164, 191], p = < 0.0001) (/a/-/u/= 14.16, SE = 5.22, 95% CI [-24.4, -3.9], p = 0.0069). In contrast, the Low AQ group showed no difference across vowel contexts (/a/: M = 177, SE = 6.56, 95% CI [164, 191], p < 0.001, /u/: M = 178, SE = 6.56, CI [165, 192], p = < 0.0001).

Group comparisons averaged over vowel contexts indicated longer latencies for the High AQ group (M = 185, SE = 6.02, 95% CI [172, 196], p = < 0.0001) relative to the Low AQ group (M = 178, SE = 6.02, 95% CI [165, 190], p = < 0.0001, but this difference was not significant (High-Low = 6.99, SE = 8.46, 95% CI [-10.5, 24.5], p = 0.417).

These results suggest that High AQ participants demonstrated a reliable latency increase from the /a/ context to the /u/ context, while Low AQ participants did not.

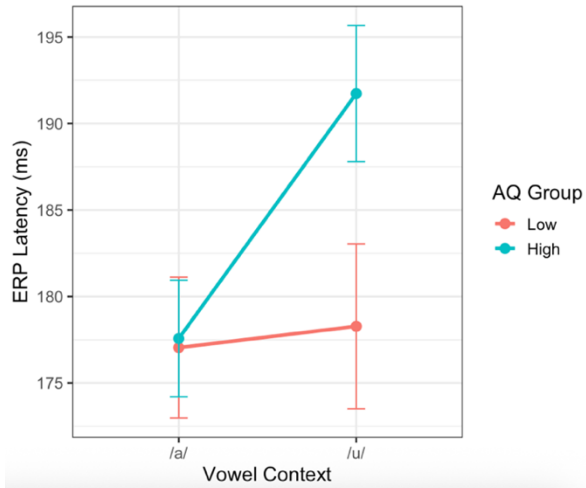

Figure 13. Mean ERP latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ (red) and High AQ (blue) groups. Points represent the grand average across participants within each group and context. Error bars represent standard errors.

Table 10. Linear mixed-effects model estimates for ERP response latency as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	181.3185	4.2786	23.0575	42.378	<2e-16 ***
VowelContext_{u}	-3.8474	1.8439	405.0000	-2.087	0.0376 *
$\text{AQ\_Group}_{\text{Low}}$	3.4948	4.2287	22.0000	0.826	0.4174
$\text{Region}_{\text{Frontal}}$	0.4849	1.9558	405.0000	0.248	0.8043
$\text{VowelContext}_{u}\text{:AQ\_Group}_{\text{Low}}$	-3.2326	1.8439	405.0000	-1.753	0.0803

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

3.5 Exploratory Results

3.5.1 Gender Differences in Behavioural Categorization

To examine possible gender differences in fricative categorization, the same logistic mixed-effects model described in Section 3.2.1 was re-fit, replacing the continuous AQ predictor with Gender, which was sum coded (-1 = female, +1 = male) and retaining all other interactions. The model formula was: response_bin ~ trial_z + vowel * (step_z + step^2_z) * gender + (1 + step_z + step^2_z | participant). Given the imbalanced gender distribution (Females: N = 45, Males: N = 7), gender effects are treated as exploratory.

Consistent with Section 3.2.1, there was a significant main effect of vowel (β = 1.144, p = 1.47e-15) and step (β = 4.531, p = 6.10e-15), indicating reliable categorization along the /s/-/ʃ/ continuum (Figure 14). There were also significant interactions between vowel and gender (β = -0.495, p = 5.48e-04) and vowel, step^2 and gender (β = -0.334, p = 0.018), suggesting that the influence of vowel context differed across male and female participants.

Estimated marginal means showed that in the female group, the probability of a /ʃ/ response was higher in the /a/ context (M = 0.864, SE = 0.044, 95% CI [0.753, 0.929]) than in the /u/ context (M = 0.633, SE = 0.086, 95% CI [0.455, 0.781]). The male group showed a similar pattern with a much higher probability of /ʃ/ responses in the /a/ context (M = 0.939, SE = 0.040, 95% CI [0.798, 0.983]) than in /u/ (M = 0.366, SE = 0.126, 95% CI [0.166, 0.625]. Pairwise contrasts confirmed these patterns, with females 3.66 times more likely to categorize a stimulus as /ʃ/ in the /a/ context than the /u/ context (OR = 3.66, SE = 0.59, 95% CI [2.67, 5.03], p = < 0.0001). Males showed a much stronger vowel-driven shift, being 26.57 times more likely to categorize as /ʃ/ in /a/ compared to /u/ (OR = 26.57, SE = 14.60, 95% CI [9.04, 78.15], p = < 0.0001).

Together, these exploratory results suggest that both male and female participants exhibited context-dependent categorization, but the boundary shift between the /a/ and /u/ contexts appeared stronger in males. However, these gender-related effects should be interpreted cautiously, given the imbalance in group sizes.

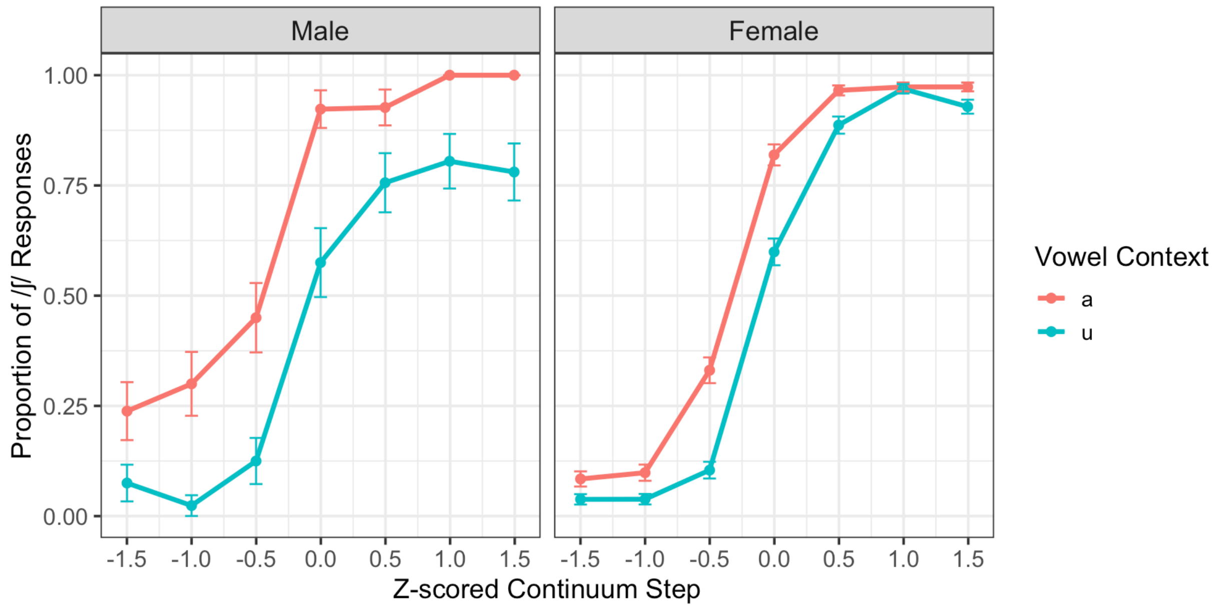

Figure 14. Proportion of /ʃ/ responses across the fricative continuum by vowel context and gender. Panels show identification curves for male (left) and female (right) participants. Red lines represent the /a/ vowel context and blue lines represent the /u/ vowel context. Error bars represent standard errors.

Table 11. Logistic mixed-effects model estimates for all predictors in the 2AFC identification task with gender as a predictor.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	1.14277	0.36379	3.141	0.001682 **
Trial_z	-0.11481	0.05533	-2.075	0.037970 *
Vowel_{u}	1.14449	0.14343	7.980	1.47e-15 ***
Step_z	4.53121	0.58079	7.802	6.10e-15 ***
Step^2_z	0.08464	0.33658	0.251	0.801448
$\text{Gender}_{\text{Female}}$	0.05351	0.29761	0.180	0.857312
Vowel_{u}:Step_z	0.19305	0.16351	1.181	0.237745
Vowel_{u}:Step^2_z	0.35429	0.14097	2.513	0.011962 *
Vowel_{u}:$\text{Gender}_{\text{Female}}$	-0.49545	0.14335	-3.456	0.000548 ***
Step_z:$\text{Gender}_{\text{Female}}$	0.53696	0.56475	0.951	0.341711
Step^2_z:$\text{Gender}_{\text{Female}}$	0.09548	0.25946	0.368	0.712866
Vowel_{u}:Step_z:$\text{Gender}_{\text{Female}}$	-0.23946	0.16313	-1.468	0.142127
Vowel_{u}:Step^2_z:$\text{Gender}_{\text{Female}}$	-0.33384	0.14051	-2.376	0.017508 *

Note: Responses were coded as 1 = /ʃ/, 0 = /s/. Vowel was sum coded (-1 = /u/, +1 = /a/). Gender was sum coded (-1 = female, +1 = male). Step_z is the z-scored linear continuum. Step^2_z is the quadratic term. Trial_z is the z-scored trial order. Significance codes: *** $p < .001$, ** $p < .01$, * $p < .05$.

3.5.2 P200 Response

Visual inspection of the ERP waveforms (Figure 9) suggested a positivity around the 200 ms range on the deviant trials that could overlap with, and potentially distort, MMN measurements. As such an early positivity can alter the deviant-standard subtraction, exploratory analysis of the P200 response was conducted to assess whether vowel context and AQ influence this component in ways that may explain the observed MMN pattern.

The deviant waveforms for both /a/ and /u/ were analyzed (as opposed to the difference wave) to avoid standard-trial adaptation effects (reduced responses to frequently repeated standard stimuli) and to ask whether the deviant positivity varied by vowel and AQ.

3.5.3 P200 Amplitude

To examine the relationship between P200 amplitudes and individual variation in AQ scores, linear mixed-effects models were fit to mean amplitude values extracted from the deviant waveform within the P200 time window (180-250 ms post stimulus onset). P200 amplitudes (value) were modelled exactly as the MMN models described earlier, with vowel context sum coded (-1 = /u/ context, +1 = /a/ context), AQ (z-scored or grouped (AQ group sum coded: -1 = Low, +1 = High)), scalp region sum coded (-1 = frontal, +1 = central) and their interactions. Random intercepts by participants were also included.

When AQ was treated as a continuous predictor, a reliable ERP response was observed within the P200 time window as indicated by the significant intercept (β = 0.429, p = 8.73e-08). The model also displayed a significant main effect of vowel context (β = -0.034, p = 0.020) and region (β = -0.108, p = 7.81e-12), while the main effect of AQ was not significant. A significant interaction between vowel context and AQ_z (β = -0.056, p = 1.67e-04) was observed, suggesting that the amplitude difference between vowel contexts varied as a function of AQ score.

Estimated marginal means confirmed that reliable P200 amplitudes were observed in both vowel contexts where /a/ produced slightly smaller P200 amplitudes (M = 0.395, SE = 0.068, 95% CI [0.257, 0.532], p = < 0.0001) compared to /u/ (M = 0.493, SE = 0.068, 95% CI [0.326, 0.600], p = < 0.0001). The pairwise contrast confirmed that amplitudes in the /a/ context were significantly reduced relative to /u/ (/a/-/u/ = -0.069, SE = 0.029, 95% CI [-0.126, -0.011], p = 0.020).

Table 12. Linear mixed-effects model estimates for P200 response amplitude as a function of AQ score, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.42898	0.06660	42.51441	6.442	8.73e-08 ***
$VowelContext_u$	-0.03428	0.01467	745.00000	-2.336	0.019762 *
AQ_z	-0.02710	0.06644	42.00000	-0.408	0.685429
$Region_{Frontal}$	-0.10823	0.01556	745.00000	-6.953	7.81e-12 ***
$VowelContext_u$:AQ_z	-0.05556	0.01468	745.00000	-3.783	0.000167 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

When AQ was grouped, a significant intercept was once again observed (β = 0.435, p = 1.80e-04). The model also showed main effects of vowel context (β = 0.085, p = 7.62e-06) and region (β = -0.096, p = 2.13e-06), but no significant effect of AQ group. Crucially, there was a significant vowel context and AQ group interaction (β = -0.063, p = 8.44e-04), suggesting that differences between both vowel contexts varied by AQ group.

Estimated marginal means showed that in the High AQ group, P200 amplitudes were smaller for the /a/ context (M = 0.268, SE = 0.14, 95% CI [-0.020, 0.556], p = 0.066) and not significant, compared to /u/ (M = 0.566, SE = 0.14, 95% CI [0.278, 0.854], p = 0.0005). The High AQ group also displayed a significant pairwise contrast between contexts (/a/-/u/ = -0.298, SE = 0.053, 95% CI [-0.403, -0.193], p = < 0.0001). By contrast, in the Low AQ group, amplitudes were similar across both vowel contexts and were significant (/a/: M = 0.432 SE = 0.14, 95% CI [0.143, 0.720], p = 0.005, /u/: M = 0.476, SE = 0.14, 95% CI [0.188, 0.764], p = 0.002). However, the pairwise difference between contexts was not significant (/a/-/u/ = -0.044, SE = 0.053, 95% CI [-0.149, 0.061], p = 0.408).

Overall, these results indicate that vowel-dependent P200 differences are more pronounced in the High AQ group, as seen in Figure 15.

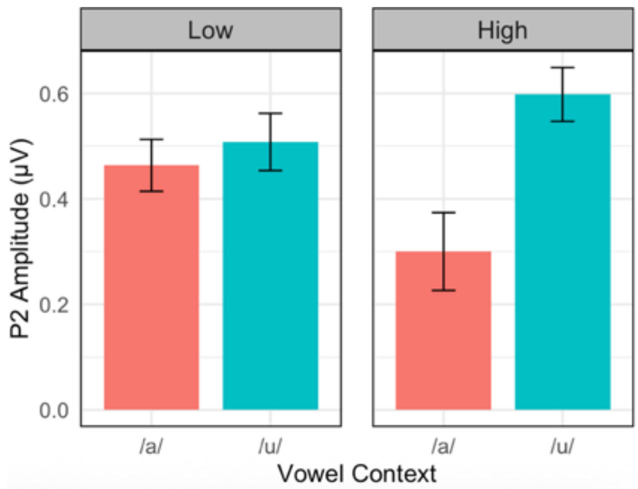

Figure 15. Mean P200 response amplitudes (µV) in the 180–250 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ (left) and High AQ (right) groups. Error bars represent standard errors.

Table 13. Linear mixed-effects model estimates for P200 response amplitude as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	0.43537	0.09704	22.20899	4.487	0.000180 ***
$VowelContext_u$	-0.08548	0.01885	405.00000	-4.535	7.62e-06 ***
AQ_Group_{Low}	-0.01833	0.09681	22.00000	-0.189	0.851536
$Region_{Frontal}$	-0.09618	0.01999	405.00000	-4.810	2.13e-06 ***
$VowelContext_u:AQ\_Group_{Low}$	-0.06340	0.01885	405.00000	-3.363	0.000844 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

3.5.4 P200 Latency

The same linear mixed-effects modelling approach used for P200 amplitudes was applied to P200 latency. P200 response latency, defined as the time point (in ms) at which the peak of the deviant waveform occurred within the P200 time window, was modelled using the same fixed-effects structure: vowel context sum coded (-1 = /u/ context, +1 = /a/ context), AQ (z-scored or grouped (AQ group sum coded: -1 = Low, +1 = High)), scalp region sum coded (-1 = frontal, +1 = central) and their interactions. Random intercepts by participants were also included.

When AQ was as a continuous predictor, a significant intercept was observed with a predicted P200 peak latency of approximately 217 ms (β = 216.73, p = < 2e-16). Figure 16 shows how P200 latency shifts with standardized AQ for both vowel contexts. The model showed a significant main effect of vowel context (β = -1.127, p = 0.015), indicating latency differences between vowel contexts, as well as a significant main effect of region (β = 1.219, p = 0.013). AQ was not a significant predictor. However, a significant interaction between vowel context and AQ was also observed (β = 1.612, p = < 4.90e-04), suggesting that differences in latency between vowel contexts varied with individual differences in AQ.

Estimated marginal means showed that latencies were shorter in the /a/ context (M = 216, SE = 1.66, 95% CI [212, 219], p = < 0.0001) compared to /u/ (M = 218, SE = 1.66, 95% CI [215, 221], p = < 0.0001). The pairwise contrast confirmed this difference (/a/-/u/ = -2.25, SE = 0.92, 95% CI [-4.06, -0.448], p = 0.015).

The results indicate that, on average, the /a/ context elicited slightly faster P200 latencies than the /u/ context, and that this latency difference was amplified with higher AQ scores.

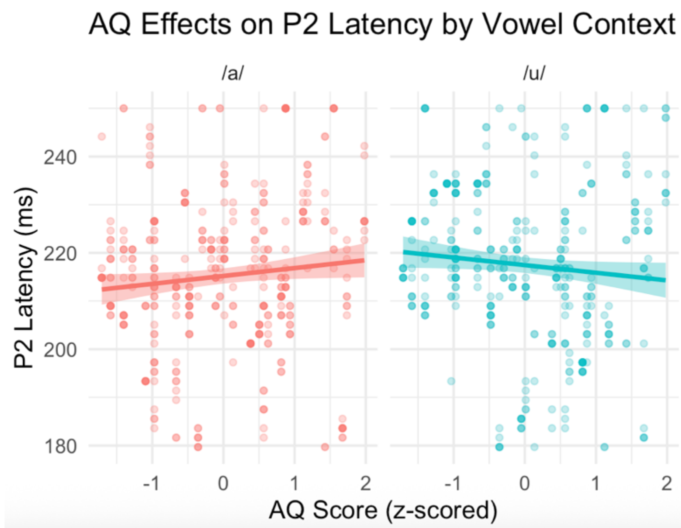

Figure 16. Relationship between z-scored AQ scores and P200 latency (ms) for the /a/ (left) and /u/ (right) vowel contexts. Each point represents an individual trial, with solid lines showing the fitted regression line and shaded areas indicating the confidence interval.

Table 14. Linear mixed-effects model estimates for P200 response latency as a function of AQ score, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	216.73215	1.59698	42.88488	135.713	< 2e-16 ***
$VowelContext_u$	-1.12701	0.46006	745.00000	-2.450	0.01453 *
AQ_z	0.02658	1.58968	42.00000	0.017	0.98674
$Region_{Frontal}$	1.21886	0.48797	745.00000	2.498	0.01271 *
$VowelContext_u$:AQ_z	1.61185	0.46035	745.00000	3.501	0.00049 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

Figure 17 displays the P200 latency differences across AQ groups and vowel contexts. The AQ group model displayed a significant intercept (β = 218.9, p = < 2e-16), main effect of vowel context (β = 1.741, p = 0.006) and a significant vowel context and AQ group interaction (β = 2.491, p = 9.51e-05). The main effect of AQ group and region were not significant.

Estimated marginal means showed that within the High AQ group, mean latencies did not differ between the /a/ (M = 220, SE = 3.21, 95% CI [213, 226], p = < 0.0001) and /u/ context (M = 218, SE = 3.21, 95% CI [212, 225], p = < 0.0001). Pairwise contrasts showed no significant differences between vowel contexts. In contrast, the Low AQ group exhibited shorter latencies in the /a/ context (M = 215, SE = 3.21, 95% CI [208, 221], *p* = < 0.0001) compared to the /u/ context (M = 223, SE = 3.21, 95% CI [217, 230], *p* = < 0.0001). Pairwise contrasts confirmed this difference in the Low AQ group (/a/-/u/ = -8.46, SE = 1.79, 95% CI [-11.98, -4.95], *p* = < 0.0001).

These results suggest that Low AQ participants showed a reliable vowel context related latency shift, with slower responses in the /u/ context compared to the /a/ context. By contrast, High AQ participants did not exhibit a significant vowel context effect.

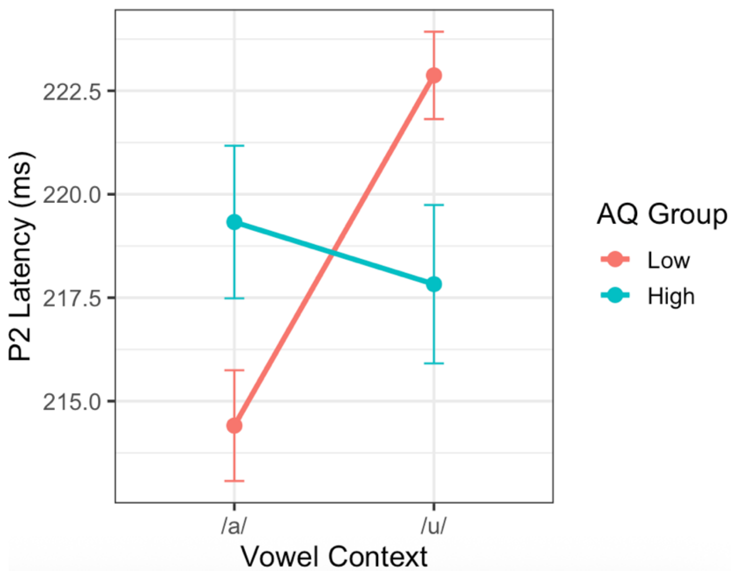

Figure 17. Mean P200 latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ (red) and High AQ (blue) groups. Points represent the grand average across participants within each group and context. Error bars represent standard errors.

Table 15. Linear mixed-effects model estimates for P200 response latency as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	218.90936	2.18828	22.46580	100.037	< 2e-16 ***
$\text{VowelContext}_{\text{u}}$	-1.74065	0.63193	405.00000	-2.754	0.00614 **
$\text{AQ\_Group}_{\text{Low}}$	-0.03161	2.17684	22.00000	-0.015	0.98854
$\text{Region}_{\text{Frontal}}$	0.89859	0.67026	405.00000	1.341	0.18078
$\text{VowelContext}_{\text{u}}\text{:AQ\_Group}_{\text{Low}}$	2.49115	0.63193	405.00000	3.942	9.51e-05 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

3.5.5 Late MMN

An exploratory analysis was conducted in a later time window (300-500 ms) to assess a late MMN response. The same linear mixed-effects model used for the earlier MMN time window was re-fit with identical coding where amplitude and latency (value) were modelled with vowel context sum coded (-1 = /u/ context, +1 = /a/ context), AQ (z-scored or grouped (AQ group sum coded: -1 = Low, +1 = High)), scalp region sum coded (-1 = frontal, +1 = central) and their interactions. Random intercepts by participants were also included.

3.5.6 Late MMN Amplitude

When AQ was modelled as a continuous predictor, the model displayed a significant main effect of vowel context ($\beta = 0.031$, $p = 0.017$), and a significant vowel context and AQ interaction ($\beta = -0.0364$, $p = 0.005$). Neither the main effect of AQ or region was significant.

Estimated marginal means showed that the /u/ context elicited more negative amplitudes (M = -0.068, SE = 0.039, 95% CI [-0.146, 0.010], $p = 0.088$) compared to the /a/ context (M = -0.007, SE = 0.039, 95% CI [-0.085, 0.072], $p = 0.868$). The pairwise contrast confirmed that this difference was significant (/a/-/u/ = 0.061, SE = 0.026, 95% CI [0.011, 0.111], $p = 0.018$).

These results suggest that late MMN amplitudes differed reliably across vowel contexts, with the /u/ context producing a more negative response than the /a/ context.

Table 16. Linear mixed-effects model estimates for Late MMN response amplitude as a function of AQ score, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	-0.03706	0.03677	43.30762	-1.008	0.31909
$VowelContext_u$	0.03056	0.01283	745.00000	2.382	0.01749 *
AQ_z	-0.01027	0.03651	42.00000	-0.281	0.77998
$Region_{Frontal}$	-0.01130	0.01361	745.00000	-0.831	0.40652
$VowelContext_u$:AQ_z	-0.03642	0.01284	745.00000	-2.837	0.00468 **

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

When AQ was grouped, the model showed a significant effect of vowel context (β = ;8, p = 0.05) and a significant vowel context and AQ group interaction (β = -0.057, p = 4.48e-04). The main effects of AQ group and region were not significant. Figure 18 shows the mean Late MMN response amplitudes for each vowel context separated by AQ group.

Estimated marginal means indicated that in the High AQ group, amplitudes did not differ significantly between the /a/ context (M = 0.056, SE = 0.064, 95% CI [-0.188, 0.075], p = 0.267) and /u/ context (M = 0.006, SE = 0.064, 95% CI [-0.137, 0.126], p = 0.0001). The pairwise contrast did not reveal a significant difference between vowel contexts in the High AQ group. In contrast, in the Low AQ group, the /u/ context produced significantly more negative amplitudes (M = -0.142, SE = 0.064, 95% CI [-0.274, -0.012], p = 0.034) compared to the /a/ context (M = 0.064, 95% CI [-0.097, 0.166], p = 0.595). The pairwise contrast confirmed this difference in the Low AQ group (/a/-/u/ = 0.178, SE = 0.046, 95% CI [0.088, 0.267], p = 0.0001).

These results indicate that the effect of vowel context on late MMN amplitude depended on AQ group, where Low AQ participants showed a reliable difference between vowel contexts, while High AQ participants did not.

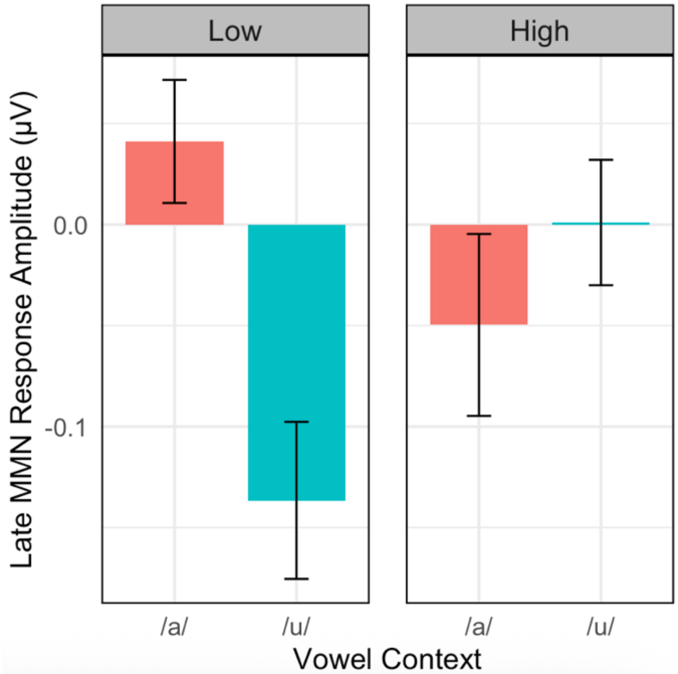

Figure 18. Mean Late MMN response amplitudes (µV) in the 300–500 ms time window for the /a/ and /u/ vowel contexts, shown separately for Low AQ (left) and High AQ (right) groups. Error bars represent standard errors.

Table 17. Linear mixed-effects model estimates for Late MMN response amplitude as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	-0.04259	0.04274	22.80486	-0.997	0.329412
$VowelContext_{u}$	0.03175	0.01613	405.00000	1.968	0.049797 *
AQ_Group_{Low}	0.01168	0.04236	22.00000	0.276	0.785382
$Region_{Frontal}$	-0.01970	0.01711	405.00000	-1.151	0.250385
$VowelContext_{u}$:AQ_Group_{Low}	-0.05710	0.01613	405.00000	-3.539	0.000448 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in microvolts (µV). Positive values indicate more positive-going amplitudes. Significance codes: *** p < .001, ** p < .01, * p < .05.

3.5.7 Late MMN Latency

When using the parallel model for latency, the continuous AQ analysis showed a significant intercept (β = 401.90, p = < 2e-16), indicating that the predicted peak latency was around 402 ms. There was no significant main effect of vowel context, AQ score or region. However, a significant vowel context and AQ interaction emerged (β = 4.26, p = 0.013), suggesting that the latency difference between vowel contexts varied as a function of AQ score (shown in Figure 19).

Estimated marginal means indicated that the /a/ context (M = 403, SE = 4.44, 95% CI [394, 411], p = < 0.0001) and the /u/ context (M = 401, SE = 4.44, 95% CI [392, 410], p = < 0.0001) elicited comparable peak latencies, however, the difference between vowel contexts was not significant.

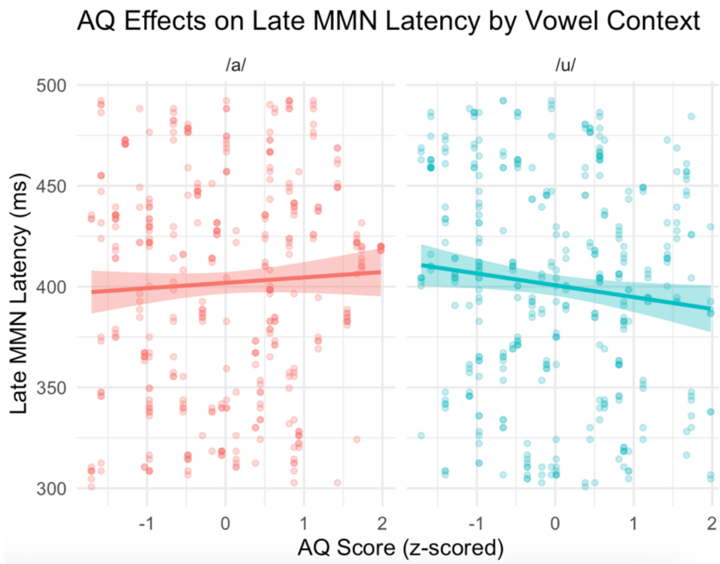

Figure 19. Relationship between z-scored AQ scores and Late MMN latency (ms) for the /a/ (left) and /u/ (right) vowel contexts. Each point represents an individual trial, with solid lines showing the fitted regression line and shaded areas indicating the confidence interval.

Table 18. Linear mixed-effects model estimates for Late MMN response latency as a function of AQ score, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	401.8980	4.0981	43.8824	98.070	<2e-16 ***
VowelContext_{u}	0.6116	1.7074	745.0000	0.358	0.7203
AQ_z	-1.5931	4.0559	42.0000	-0.393	0.6965
$\text{Region}_{\text{Frontal}}$	1.8884	1.8109	745.0000	1.043	0.2974
VowelContext_{u}:AQ_z	4.2609	1.7084	745.0000	2.494	0.0128 *

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ_z is the z-scored Autism Quotient scores. Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

In the grouped analysis, the model once again displayed a significant intercept (β = 402.73, p = < 2e-16). There were no main effects of vowel context, AQ group or region.

However, a significant vowel context and AQ group interaction emerged ($\beta = 7.80$, $p = 3.96$e-04), suggesting that the latency difference between vowel contexts depended on AQ group.

Estimated marginal means showed that in the High AQ group, the /a/ context elicited longer latencies (M = 408, SE = 8.45, 95% CI [391, 425], $p = < 0.0001$) than the /u/ context (M = 390 ms, SE = 8.45, 95% CI [372, 407], $p = < 0.0001$). The pairwise contrast confirmed this effect for the High AQ group (/a/-/u/ = 18.4, SE = 6.17, 95% CI [6.31, 30.58], $p = 0.003$). By contrast, in the Low AQ group, the /u/ context latencies were longer (M = 413, SE = 8.45, 95% CI [396, 430], $p = < 0.0001$) than the /a/ context (M = 400, SE = 8.45, 95% CI [383, 417], $p = < 0.0001$). The pairwise contrast also confirmed this effect for the Low AQ group (/a/-/u/ = -12.7, SE = 6.17, 95% CI [-24.89, -0.61], $p = 0.039$).

These results suggest that late MMN latency varied as a function of AQ group (shown in Figure 20). High AQ participants showed faster latencies in the /u/ context relative to /a/, whereas Low AQ participants showed the opposite pattern with slower latencies in /u/ compared to /a/.

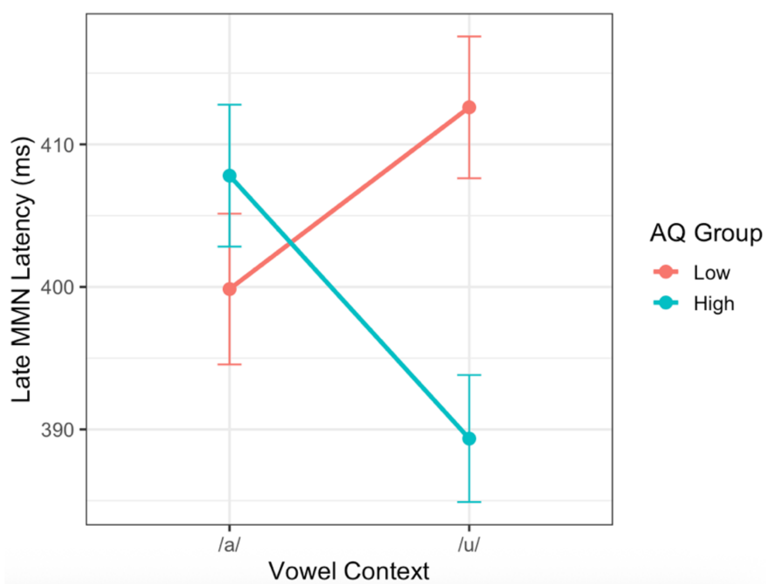

Figure 20. Mean Late MMN latency (ms) for the /a/ and /u/ vowel contexts, shown separately for Low AQ (red) and High AQ (blue) groups. Points represent the grand average across participants within each group and context. Error bars represent standard errors.

Table 19. Linear mixed-effects model estimates for Late MMN response latency as a function of AQ group, vowel context and scalp region.

	Estimate	Std. Error	df	t value	Pr (>|t|)
(Intercept)	402.7269	5.5880	22.8635	72.070	<2e-16 ***
$VowelContext_u$	1.4242	2.1829	405.0000	0.652	0.514503
AQ_Group_{Low}	-3.8204	5.5344	22.0000	-0.690	0.497230
$Region_{Frontal}$	0.9732	2.3153	405.0000	0.420	0.674480
$VowelContext_u:AQ\_Group_{Low}$	7.7989	2.1829	405.0000	3.573	0.000396 ***

Note: Vowel context was sum coded (-1 = /u/ context, +1 = /a/ context. Region was sum coded (-1 = frontal, +1 = central). AQ group was sum coded (-1 = Low, +1 = High). Coefficients are reported in milliseconds (ms). Significance codes: *** p < .001, ** p < .01, * p < .05.

4. Discussion

The study investigated how individual differences in cognitive-behavioural traits, as measured by the AQ, shape behavioural and neural responses to VFV syllables. Specifically, the study investigated how coarticulatory vowel context influenced fricative categorization and whether these effects varied with AQ scores. A combination of a 2AFC behavioural task and EEG recordings of MMN was used to assess perceptual and neural sensitivity to /s/-/ʃ/ contrasts across different vowel contexts. The research provides insights into the mechanisms underlying perceptual variability and adds to emerging evidence that cognitive-behavioural traits influence how individuals perceive and process coarticulated speech cues.

4.1 Behavioural Findings

Consistent with previous work (Mann & Repp, 1980; Yu, 2019; Yu & Lee, 2014), fricative categorization was influenced by vowel context where participants produced fewer /ʃ/ responses in the /u/ context compared to /a/. As Yu (2010) noted, lip protrusion associated with rounded vowels like /u/ lowers the center of gravity of adjacent fricatives like /s/, making them acoustically more similar to /ʃ/ in naturally coarticulated speech. Hence, this coarticulatory effect results in spectral frequency lowering, which in turn shifts the perceptual boundary between /s/ and /ʃ/ (Yu, 2010; Yu & Lee, 2014). While these coarticulatory effects originate from acoustic variation, they may be shaped by higher-level perceptual processes. Mitterer (2006) replicated that perception of fricatives is influenced by surrounding vowels, particularly when the vowel is rounded. Mitterer (2006) argued that this compensation for coarticulation was not driven by low-level auditory processing but instead reflected phonological mediation where the perceived identity of the vowel (informed by acoustic cues) shaped how the fricative was categorized. These findings suggest that coarticulatory influences on fricative perception, such as those observed here, may engage higher-level perceptual processes that rely on phonological representations, rather than solely on acoustic similarity. Consequently, individual differences in higher-level processing, such as those related to AQ, may influence how each listener attends to and resolves these perceptual ambiguities.

While the vowel context reliably shaped fricative categorization across participants, individual differences influenced this effect. In particular, group-based analyses suggested that Low AQ participants tended to produce more /ʃ/ responses than High AQ participants overall, although the direct pairwise comparison was not significant. The PSE analysis offered further insight. The continuous AQ model showed that higher AQ scores were associated with fewer /ʃ/ responses overall, while the group-based model showed that the Low AQ group shifted toward /ʃ/ at an earlier continuum step than the High AQ group, particularly in the /u/ context. Yu (2010) found that individuals with lower AQ scores tended to under-compensate for acoustic coarticulatory effects such as those induced by rounded vowels. In the current study, this under-compensation may have led Low AQ participants to interpret coarticulated /s/ tokens as more /ʃ/-like, particularly in the /u/ context. This pattern suggests that Low AQ individuals were more likely to categorize phonemically ambiguous fricatives as /ʃ/, consistent with reduced compensation for the spectral effects of vowel rounding. The under-compensation in the Low

AQ group compared to the High AQ group contradicts other theoretical accounts (Weak Central Coherence (WCC) Theory and Theory of Mind perspectives), which predict less compensation for High AQ individuals due to reduced contextual integration. The present findings, in this study, align with Yu (2010) but highlight the need to reconcile conflicting interpretations across the literature. Additionally, in line with the boundary shift interpretation, confidence changed with contextual ambiguity (higher confidence for /u/ at step 3 and higher confidence for /a/ at step 4) and was not predicted by AQ. This supports the view that AQ effects arise at the category boundary, not from shifts in overall confidence.

In contrast, High AQ individuals, may have a more heightened focus on acoustic detail and better able to account for coarticulatory variation, consistent with the EPF model (Mottron et al., 2006). The heightened perceptual sensitivity may support more accurate normalization for context, allowing these individuals to maintain stable phoneme boundaries even when there are contextual shifts. Yu (2010) similarly proposed that because high AQ listener focus on low-level auditory patterns, it allows them to adjust for contextual variation more effectively.

Gender differences were also examined as part of these individual differences analyses. Given the gender distribution the sample (Females: n = 45, Males: n = 7), any gender contrasts were highly exploratory. The only gender-related pattern was a vowel-dependent curvature effect (vowel x step^2 x gender and vowel x gender) and did not include a gender and step interaction that would indicate a boundary shift. By contrast, AQ-related differences appeared in overall response tendencies, and in group contrasts, earlier boundary placement. Taken together, it suggests that the present findings reflect AQ-related variation rather than a gender effect. Importantly, females in this sample had higher AQ scores than males, which reverses typical population trends found in research (Baron-Cohen et al., 2001b; Yu, 2010). Nonetheless, the gender analysis was still reported as potential gender differences are theoretically relevant for fricative perception and commonly examined in literature (Bishop et al., 2020). Additionally, including gender in the analyses provides transparency and allows for comparisons with prior work.

In this study, AQ-related differences in the behavioural task were most evident in the analysis of perceptual boundaries. The distinction highlights the subtle nature of individual variability in speech perception. While all participants demonstrated sensitivity to vowel context, only the PSE analysis showed systemic differences in how listeners adjusted their category boundaries. This aligns with rational models of speech perception, which predict that contextual variability and uncertainty increase reliance on prior phonetic category knowledge and lead to boundary-level shifts (Feldman, Griffiths, & Morgan, 2009). As the boundary-level shifts driven by contextual variability are more pronounced at categorical boundaries, methods like PSE may be informative for detecting subtle differences in perception that are not seen in identification accuracy. Yu and Lee (2014) found that although group-level discrimination and identification scores appeared consistent, individual-level analyses displayed considerable variation with some participants showing more perceptual compensation to coarticulatory context than others. These findings support the results of this study where group comparisons based on PSE showed subtle AQ-related differences in boundary placement that were not evident in the overall identification results.

4.2 Neural Findings

4.2.1 ERP Response Amplitude

When AQ was modelled as a continuous predictor, ERP analyses of the deviant-standard difference waves showed significant neural responses to the stimuli across both vowel contexts.

However, AQ scores, when treated as a continuous variable, did not significantly impact the amplitude of these ERP responses. Instead, consistent with the exploratory checks above, AQ effects were interactional, influencing the size of the vowel context difference rather than overall amplitude of the response. Interestingly, group-based analyses displayed a significant interaction between AQ and vowel context. High AQ participants showed significantly greater ERP difference wave amplitudes in the /u/ context compared to the /a/ context. The Low AQ group did not show a significant difference across contexts. One possible interpretation is that high AQ individuals may have heightened sensitivity to detailed acoustic variation, consistent with the Enhanced Perceptual Functioning (EPF) model (Mottron et al., 2006). The /u/ context has greater spectral overlap between /s/ and /ʃ/, and individuals with higher AQ traits may exhibit stronger neural responses as they focus on the more low-level auditory details of sounds. Rather than reflecting reduced compensation, the larger ERP response in the /u/ condition may signal heightened bottom-up detection of acoustic change. While findings across studies have been mixed (possibly due to the differences in the deviant stimuli used) some have reported increased auditory ERP response amplitudes in individuals with ASD in some contexts (such as frequency changes), supporting the idea that those with ASD have a greater sensitivity to acoustic differences (Clara et al., 2023; Ferri et al., 2003; Lepistö et al., 2005). Importantly, the amplitude of early auditory ERP responses has been linked not only to change detection but also to the quality of phonological representations (Pakarinen et al., 2007), suggesting that the enhanced responses observed in high AQ individuals may reflect a higher sensitivity to phonetic detail.

4.2.2. ERP Latency

In terms of latency, the ERP responses occurred around 175 ms post-stimulus onset, consistent with prior work on early auditory discrimination (Garrido et al., 2009). While no significant effects were found when AQ was modelled as a continuous predictor, group-based analyses showed a main effect of vowel context. Participants in the High AQ group exhibited slower ERP responses in the /u/ context compared to /a/. The delay likely reflects increased processing demands due to the greater acoustic ambiguity of the rounded vowel context, suggesting that individuals with higher AQ traits may be less efficient at resolving coarticulatory variation. Taken together with the amplitude findings, this suggests that while high AQ individuals show heightened neural sensitivity to phonetic detail, this increased sensitivity comes with slower processing of coarticulatory variation. The latency findings align with the WCC theory, which discusses a reduced tendency to integrate contextual cues, leading to more reliance on bottom-up acoustic input (Happé & Frith, 2006; Stewart & Ota, 2008). It also complements the EPF account, which emphasizes heightened sensitivity to fine-grained auditory detail at the expense of integrating global context (Mottron et al., 2006). Within the predictive coding framework, MMN latency (or in this case ERP response latency) reflects the speed at which the brain detects a mismatch between sensory input and internal predictions (Garrido et al., 2009). A longer latency suggests that the process of updating the auditory model requires more time or effort in individuals with higher AQ traits. The observed latency patterns are consistent with prior findings of delayed MMN responses in individuals with ASD and extends them to the broader AQ spectrum (Huang et al., 2017; Kasai et al.,2005; Kuhl et al., 2004; Roberts et al., 2011; Vlaskamp et al., 2017; Yu et al., 2015). Overall, these findings highlight how individual-level variability in perceptual processing can influence the timing of early auditory responses, particularly under conditions of phonetic uncertainty.

4.3.3 P200 Overlap and Late MMN

Furthermore, the ERP analysis showed that, although the observed difference waves were positive in polarity, they consistently fell within the typical MMN time window (100-250 ms), suggesting the presence of MMN-like responses. This polarity shift was also evident in the topographic maps, which displayed a fronto-central positivity rather than the expected negativity typically associated with MMN. One likely explanation for this complication may be overlapping ERP components such as P200. MMN is partly overlapped by other ERP components within similar time windows which can make MMN amplitude and latency ambiguous (Pekkonen et al., 1993). Since the MMN analysis includes deviant-standard waves, an enhanced P200 to the deviant may have masked or inverted an MMN-like negativity, resulting in a positive-going waveform within the MMN time window. To investigate this overlap account, two exploratory checks were included: (1) P200 on the deviant waveform (180-250 ms), and (2) a later difference-wave window (300-500 ms).

First, P200 measured on the deviant waveform was reliable for both vowel contexts and was larger for /u/ than for /a/ at approximately 217 ms. In addition, an AQ and vowel interaction indicated stronger amplitude differences between contexts as AQ increased. Prior work suggests that the P200 response may reflect early phonemic categorization and discrimination similar to MMN, which indexes the detection of auditory change (Brunellière et al., 2009). Brunellière et al. (2009) found that a stable and phonemically distinct contrast in French elicited both MMN and P200 responses, while a less perceptually distinct contrast elicited MMN alone. The presence of a P200 in response to the stable contrast may suggest that P200 is involved in early phonemic discrimination, in addition to deviance detection with MMN. This distinction highlights that while MMN can signal that a change has been detected, the engagement of P200 appears to reflect the recruitment of higher-level processing mechanisms when the contrast involves phonemically meaningful categories rather than solely acoustic deviations. In the present study, the deviant stimuli consisted of phonemically ambiguous VFV syllables, which may have required higher-level processing of the phonemic categories to resolve ambiguity. A clear P200 response observed across both AQ groups suggests that listeners accessed phonemic-level representations accurately despite ambiguity. Therefore, it is possible that overlapping activity from P200 may have influenced the morphology or polarity of the observed difference wave.

Second, MMN responses frequently show two peaks, an early one (150-250 ms) and a later one (300-500 ms) termed as a late MMN or late discriminative negativity (Cheour et al., 1999). Late MMN was examined as these responses have been linked to higher-level processing of complex linguistic changes (Bishop et al., 2011; Korpilahti et al., 2001) and may indicate MMN responses not obscured by the P200. In the later 300-500 ms window, the difference waves showed a vowel-dependent negativity (more negative for /u/ than for /a/) at approximately 402 ms and a AQ and vowel interaction. These converging patterns are consistent with the idea that the early MMN window positivity may be P200 dominated, while an MMN-like negativity emerges later.

4.2.4 Summary of Neural Patterns

Taken together, the amplitude and latency findings suggest that individuals with higher AQ traits engage differently when faced with coarticulatory ambiguity, exhibiting both stronger but slower neural responses. The enhanced ERP difference wave amplitude observed in the /u/ context suggests a heightened neural sensitivity to auditory deviance, while the delayed latency suggests increased processing demands. In other words, while individuals with higher AQ traits are more sensitive to deviant stimuli, the processing of that deviance takes longer. These patterns are consistent with the EPF framework, which suggests enhanced low-level perceptual processing, and with WCC theory, which describes reduced top-down integration. Rather than indicating a deficit, these findings reflect one possible processing strategy that prioritizes detailed acoustic analysis over speed when faced with ambiguous speech (Yu, 2010).

4.3 Limitations

Several limitations of this study must be acknowledged. The behavioural task and EEG recordings were conducted in a non-soundproof, shared lab space. To minimize external noise, sessions were conducted in rooms with doors closed, and signs were posted to indicate that an experiment was in progress. While these measures helped to reduce disruptions, the shared nature of the lab still posed unavoidable challenges in maintaining consistent auditory conditions across sessions. Similarly, EEG recordings were susceptible to room noise and potential electromagnetic interference from nearby electronic equipment. Although standard preprocessing and artifact rejection steps were applied, these factors may have contributed to increased noise levels in the EEG data. In some cases, excessive artifacts and poor signal quality led to the exclusion of participants, further reducing the final EEG sample size and limiting statistical power.

Furthermore, the sample size for the EEG task may constrain the generalizability of the findings as it limits statistical power for detecting subtle individual difference effects. The gender distribution was also skewed, with a greater proportion of female participants. Given that AQ scores can vary by gender and that certain traits may manifest differently across genders (Baron-Cohen et al., 2001b; Yu, 2010), this imbalance may have influenced the observed effects and limits the applicability of the findings to more diverse populations. Additionally, several participants reported familiarity with the AQ questionnaire, which may have introduced bias based on how they wished to be perceived in relation to the AQ continuum.

Moreover, the contrast in results between continuous and group-based AQ analyses raise questions regarding measurement sensitivity. Group-based comparisons showed stronger patterns, likely due to clearer contrasts between individuals at the high and low ends of the AQ distribution, however, this approach may oversimplify the representation of AQ as a dimensional trait. Nonetheless, it is common practice in phonetics and broader AQ literature to adopt a high vs. low grouping strategy for comparisons, with several approaches to participant classification reported (see Stevenson & Hart, 2017). In contrast, continuous modelling preserves individual variability but is more vulnerable to the effects of sample size and distribution. As it was not feasible to include participants spanning the full spectrum of possible AQ scores, gaps in score representation may have reduced sensitivity of the continuous analysis. These findings suggest that categorical grouping may be more effective when sample size and variability are limited, but future research should continue to explore the most appropriate methods for modelling individual differences.

4.4 Conclusion

This study examined how individual differences in cognitive-behavioural traits, as measured by AQ scores, influence perception of coarticulated VFV syllables. By combining behavioural and neural measures, the findings demonstrate that perceptual and processing differences in response to phonetic variability are associated with broader cognitive-behavioural traits. The observed results suggest that individuals with higher AQ traits may rely more heavily on fine-grained acoustic information and engage in more detailed perceptual analysis, particularly in the context of increased ambiguity. These tendencies are reflected in both their behavioural categorization patterns and in the amplitude/latency of neural responses. Such differences align with broader theoretical accounts, such as the EPF and WCC models, which propose that cognitive style influences perceptual processing at multiple levels. These findings contribute to a broader understanding of the interaction between cognition, behaviour, and perception, emphasizing that speech processing is not uniform across individuals. Understanding this variability has important implications not only for models of speech perception, but also for how auditory processing is conceptualized in neurodiverse populations. By bridging cognitive theory with empirical observation, this research highlights the value of looking at individual-level differences in advancing our understanding of speech perception and its variability.

References

Ahtam, B., Braeutigam, S., & Bailey, A. (2020). Semantic Processing in Autism Spectrum Disorders Is Associated With the Timing of Language Acquisition: A Magnetoencephalographic Study. *Frontiers in Human Neuroscience*, *14*. https://doi.org/10.3389/fnhum.2020.00267

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001a). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–251.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001b). 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

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, *14*(2), 402–416. https://doi.org/10.1111/j.1467-7687.2010.00990.x

Bishop, J., Kuo, G., & Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from Rapid Prosody Transcription. *Journal of Phonetics*, *82*, 100977. https://doi.org/10.1016/j.wocn.2020.100977

Boddaert, N., Belin, P., Chabane, N., Poline, J.-B., Barthélémy, C., Mouren-Simeoni, M.-C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2003). Perception of Complex Sounds: Abnormal Pattern of Cortical Activation in Autism. *American Journal of Psychiatry*, *160*(11), 2057–2060. https://doi.org/10.1176/appi.ajp.160.11.2057

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience*, *15*(2), 226–235. https://doi.org/10.1162/089892903321208169

Booth, R., & Happé, F. (2010). "Hunting with a knife and … fork": Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, *107*(4), 377–393. https://doi.org/10.1016/j.jecp.2010.06.003

Brunellière, A., Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2009). Behavioral and electrophysiological evidence for the impact of regional variation on phoneme perception. *Cognition*, *111*(3). https://doi.org/10.1016/j.cognition.2009.02.013

Campbell, J. A., McSherry, H. L., & Theodore, R. M. (2018). Contextual Influences on Phonetic Categorization in School-Aged Children. *Frontiers in Communication*, *3*. https://doi.org/10.3389/fcomm.2018.00035

Cheour, M., Ceponiene, R., Hukki, J., Haapanen, M. L., Näätänen, R., & Alho, K. (1999). Brain dysfunction in neonates with cleft palate revealed by the mismatch negativity. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *110*(2), 324–328. https://doi.org/10.1016/s1388-2457(98)00005-4

Cheour, M., H.T. Leppänen, P., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical Neurophysiology*, *111*(1), 4–16. https://doi.org/10.1016/S1388-2457(99)00191-1

Chodroff, E., & Wilson, C. (2020). Acoustic–phonetic and auditory mechanisms of adaptation in the perception of sibilant fricatives. *Attention, Perception & Psychophysics*, *82*(4), 2027–2048. https://doi.org/10.3758/s13414-019-01894-2

Daniloff, R. G., & Hammarberg, R. E. (1973). On defining coarticulation. *Journal of Phonetics*, *1*(3), 239–248. https://doi.org/10.1016/S0095-4470(19)31388-9

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dong, L., Li, F., Liu, Q., Wen, X., Lai, Y., Xu, P., & Yao, D. (2017). MATLAB Toolboxes for Reference Electrode Standardization Technique (REST) of Scalp EEG. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00601

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, *116*(4), 752–782. https://doi.org/10.1037/a0017196

Ferri, R., Elia, M., Agarwal, N., Lanuzza, B., Musumeci, S. A., & Pennisi, G. (2003). The mismatch negativity and the P3a components of the auditory event-related potentials in autistic low-functioning subjects. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *114*(9), 1671–1680. https://doi.org/10.1016/s1388-2457(03)00153-6

Fuhrmeister, P., & Myers, E. B. (2021). Structural neural correlates of individual differences in categorical perception. *Brain and Language*, *215*, 104919. https://doi.org/10.1016/j.bandl.2021.104919

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. https://doi.org/10.1016/j.clinph.2008.11.029

Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *5*(2), 701–714. https://doi.org/10.1016/j.rasd.2010.11.006

Happé, F., & Frith, U. (2006). The Weak Coherence Account: Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Happé, F. G. E. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, *15*(Pt 1), 1–12. https://doi.org/10.1111/j.2044-835X.1997.tb00721.x

Huang, D., Yu, L., Wang, X., Fan, Y., Wang, S., & Zhang, Y. (2018). Distinct patterns of discrimination and orienting for temporal processing of speech and nonspeech in Chinese children with autism: An event-related potential study. *The European Journal of Neuroscience*, *47*(6), 662–668. https://doi.org/10.1111/ejn.13657

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders*, *37*(9), 1679–1690. https://doi.org/10.1007/s10803-006-0299-3

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. https://doi.org/10.1111/j.1467-7687.2007.00644.x

Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., Itoh, K., Koshida, I., Iwanami, A., Nakagome, K., Fukuda, M., Yamasue, H., Yamada, H., Abe, O., Aoki, S., & Kato, N. (2005). Delayed automatic detection of change in speech sounds in adults with autism: A magnetoencephalographic study. *Clinical Neurophysiology*, *116*(7), 1655–1664. https://doi.org/10.1016/j.clinph.2005.03.007

Kong, E. J., & Edwards, J. (2016). Individual differences in categorical perception of speech: Cue weighting and executive function. *Journal of Phonetics*, *59*, 40–57. https://doi.org/10.1016/j.wocn.2016.08.006

Korpilahti, P., Krause, C. M., Holopainen, I., & Lang, A. H. (2001). Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain and Language*, *76*(3), 332–339. https://doi.org/10.1006/brln.2000.2426

Krizic, M. (2023). *Prosodic Speech Production and Perception Differences Comparing Populations with Varying Levels of Autistic Traits* [Thesis]. https://macsphere.mcmaster.ca/handle/11375/29048

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, *8*(1), F1–F12. https://doi.org/10.1111/j.1467-7687.2004.00384.x

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, *1066*(1–2), 147–157. https://doi.org/10.1016/j.brainres.2005.10.052

Lenth, R. V., Banfai, B., Bolker, B., Buerkner, P., Giné-Vázquez, I., Hervé, M., Jung, M., Love, J., Miguez, F., Piaskowski, J., Riebl, H., & Singmann, H. (2025). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (Version 1.11.2-8) [Computer software]. https://cran.r-project.org/web/packages/emmeans/index.html

Linares, D., & López-Moliner, J. (2016). quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. *The R Journal*, *8*(1), 122. https://doi.org/10.32614/rj-2016-008

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Ludlow, A., Mohr, B., Whitmore, A., Garagnani, M., Pulvermüller, F., & Gutierrez, R. (2014). Auditory processing and sensory behaviours in children with autism spectrum disorders as revealed by mismatch negativity. *Brain and Cognition*, *86*, 55–63. https://doi.org/10.1016/j.bandc.2014.01.016

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*(4), 369–390. https://doi.org/10.1017/S0142716400009796

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics*, *28*(3), 213–228. https://doi.org/10.3758/BF03204377

The MathWorks Inc. (2023). *MATLAB* (Version 9.13.0 (R2023b)) [Computer software]. The MathWorks Inc. https://www.mathworks.com

Mayerle, M. C. C. da S., Riesgo, R., Gregory, L., Borges, V. M. S., & Sleifer, P. (2023). Mismatch Negativity in Children and Adolescents with Autism Spectrum Disorder. *International Archives of Otorhinolaryngology*, *27*(2), e218–e225. https://doi.org/10.1055/s-0043-1768209

Mitterer, H. (2006). On the causes of compensation for coarticulation: Evidence for phonological mediation. *Perception & Psychophysics*, *68*(7), 1227–1240. https://doi.org/10.3758/BF03193723

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Pekkonen, E., Jousmäki, V., Partanen, J., & Karhu, J. (1993). Mismatch negativity area and age-related auditory memory. *Electroencephalography and Clinical Neurophysiology*, *87*(5), 321–325. https://doi.org/10.1016/0013-4694(93)90185-X

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Recasens, D. (2018). Coarticulation. In M. Arnoff (Ed.), *Oxford Research Encyclopedia of*

Linguistics. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.416

Roberts, T. P. L., Cannon, K. M., Tavabi, K., Blaskey, L., Khan, S. Y., Monroe, J. F., Qasmieh, S., Levy, S. E., & Edgar, J. C. (2011). Auditory magnetic mismatch field latency: A biomarker for language impairment in autism. *Biological Psychiatry*, *70*(3), 263–269. https://doi.org/10.1016/j.biopsych.2011.01.015

Stevenson, J. L., & Hart, K. R. (2017). Psychometric Properties of the Autism-Spectrum Quotient for Assessing Low and High Levels of Autistic Traits in College Students. *Journal of Autism and Developmental Disorders*, *47*(6), 1838–1853. https://doi.org/10.1007/s10803-017-3109-1

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition*, *109*(1), 157–162. https://doi.org/10.1016/j.cognition.2008.07.010

Vlaskamp, C., Oranje, B., Madsen, G. F., Møllegaard Jepsen, J. R., Durston, S., Cantio, C., Glenthøj, B., & Bilenberg, N. (2017). Auditory processing in autism spectrum disorder: Mismatch negativity deficits. *Autism Research*, *10*(11), 1857–1865. https://doi.org/10.1002/aur.1821

Yao, D. (2001). A method to standardize a reference of scalp EEG recordings to a point at infinity. *Physiological Measurement*, *22*(4), 693–711. https://doi.org/10.1088/0967-3334/22/4/305

Yu, A. C. L. (2010). Perceptual Compensation Is Correlated with Individuals' "Autistic" Traits: Implications for Models of Sound Change. *PLOS ONE*, *5*(8), e11950. https://doi.org/10.1371/journal.pone.0011950

Yu, A. C. L. (2019). On the nature of the perception-production link: Individual variability in English sibilant-vowel coarticulation. *Laboratory Phonology*, *10*(1), Article 1. https://doi.org/10.5334/labphon.97

Yu, A. C. L., & Lee, H. (2014). The stability of perceptual compensation for coarticulation within and across individuals: A cross-validation study. *The Journal of the Acoustical Society of America*, *136*(1), 382–388. https://doi.org/10.1121/1.4883380

Yu, A. C. L., & To, C. K. S. (2020). Atypical context-dependent speech processing in autism. *Applied Psycholinguistics*, *41*(5), 1045–1059. https://doi.org/10.1017/S0142716420000387

Yu, L., Fan, Y., Deng, Z., Huang, D., Wang, S., & Zhang, Y. (2015). Pitch Processing in Tonal-Language-Speaking Children with Autism: An Event-Related Potential Study. *Journal of Autism and Developmental Disorders*, *45*(11), 3656–3667. https://doi.org/10.1007/s10803-015-2510-x