PROSODIC SPEECH PRODUCTION AND PERCEPTION DIFFERENCES COMPARING POPULATIONS WITH VARYING LEVELS OF AUTISTIC TRAITS

PROSODIC SPEECH PRODUCTION AND PERCEPTION DIFFERENCES COMPARING POPULATIONS WITH VARYING LEVELS OF AUTISTIC TRAITS

By MONIKA KRIZIC, B.A. (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 (2023)
Hamilton, Ontario (Cognitive Science of Language)

TITLE: Prosodic Speech Production and Perception Differences Comparing Populations with Varying Levels of Autistic Traits

AUTHOR: Monika Krizic, B.A. (Hons) (University of Toronto)
SUPERVISOR: Dr. Daniel Pape
NUMBER OF PAGES: x, 57

Lay Abstract

Autism spectrum disorder (ASD) represents a group of developmental disabilities associated with impairments in communicative abilities, among others. Theories suggest that individuals with higher levels of autistic traits notice small details in the physical properties of sounds, but have trouble distinguishing the more abstract, intended meaning of the same sound patterns. Previous studies found that individual differences in the degree of autistic traits influence one's production and perception of prosody (i.e., the relative highness or lowness of a tone),; individuals with higher levels of autistic are better able to detect fine-grained differences in pitch and time, but not loudness. The present study examined the extent to which speakers with varying levels of autistic traits use prosody during speech production and perception.

This study observed that (1) individuals with higher levels of autistic traits displayed an enhanced *perception* of pitch and loudness, but not time, and (2) that these same participants may exhibit less variability in their *production* of pitch.

Abstract

Autism spectrum disorder (ASD) represents a group of developmental disabilities associated with impairments in social, communicative, and imaginative abilities. Speech impairments associated with ASD can be explained by differences in cognitive processing styles relative to neurotypicals. Previous studies found that individual differences in cognitive processing influence one's production and perception of prosody. For example, Stewart et al. (2018) found that higher levels of autistic character traits indicated by one's Autism Spectrum Quotient (AQ) score (Baron-Cohen et al., 2001) correlated significantly with one's ability to discriminate pitch and time, but no significant correlation between auditory discrimination thresholds for intensity. Additionally, Turnbull (2015; 2019) observed shorter overall word and vowel durations during a task which required participants with varying AQ scores to speak for the benefit of a listener with a hearing impairment.

The present study examined whether prosodic cue-trading in production and perception differs when comparing populations with varying levels of autistic traits, as indicated by their AQ score differences. Furthermore, the study investigated whether these differences exist on a continuum, or rather are categorical, with respect to participants' level of autistic character traits. To achieve this, we analyzed individual variability patterns in 18 participants' speech production and perception.

Results from the *perception* task showed that participants displayed a significant enhanced perception of pitch and intensity, but not duration, when completing a task where participants listened to sentences manipulating the prosodic parameters f0, intensity, duration. Results from the *production* task where participants read sentences designed to elicit background, broad, and narrow focus found no significant effect of AQ across any of the

acoustic parameters measured, although the results for f0 are near the 5% significance level for the f0 condition, suggesting that participants with higher AQ scores may produce lower f0 ranges, and thus, less prosodic variability compared to low AQ participants.

Acknowledgements

This project is very close to my heart—having grown up with family and friends with ASD and watching these loved ones overcome various communicative challenges has been a true inspiration for me. I hope that the results of this project could be beneficial. I would like to express my deepest appreciation to the following people, without whom I would not have been able to complete this research:

To my supervisor, Daniel Pape, thank you for giving me the opportunity to work on this project with you. Thank you for introducing me to prosody research, and for teaching me more than I ever thought I could learn. In doing so, you have helped me strengthen my skills as a researcher and find my passion for this area of research, which I will cherish throughout the rest of my academic career and life. I cannot thank you enough for your knowledge, guidance, and support throughout this experience. It has been an absolute pleasure getting to know you and work with you these last two years!

I would also like to thank my thesis committee members, Gemma Repiso Puigdelliura and Magda Stroinska for their valuable comments on my thesis drafts and for always making themselves available to answer any questions I had.

To the amazing group of research assistants involved in this project, thank you for your hard work and dedication during data analysis. I could not have finished this project without the support of this wonderful team. I truly appreciate the time and effort each of you put into this project with me!

To my mom, Linda Fronteddu, thank you for supporting my decision to pursue my Master's degree. Thank you for always being excited to hear about this project. Thank you for celebrating every little milestone on this journey with me. Thank you for encouraging me from a young age to love learning. Thank you for everything you have done to support me. I love you so much.

To my partner, David Myers, thank you and your family for welcoming me into your home, so I could always be close to the lab. Thank you for cooking me amazing meals, when I otherwise wouldn't have made the time to do so myself. Thank you for staying up late to study with me and listen to me practice conference presentations over and over and over again. Thank you for comforting me during times of uncertainty. I am so grateful to have you in my life.

Finally, thank you to everyone who was involved in this project. Thank you to the participants, my colleagues in the Phonetics Lab, and my friends and family. Thank you to the ARiEAL Research Centre and the Department of Linguistics and Languages at McMaster University. This has been one of the most rewarding, invaluable experiences that I will remember for the rest of my life.

Table of Contents

List of Figures and Tables

List of all Abbreviations and Symbols

ASD – Autism Spectrum Disorder

AQ – Autism Spectrum Quotient

EMB — Extreme Male Brain

EPF — Enhanced Perceptual Function

TD — Typically Developed

VOT – Voice Onset Time

WCC — Weak Central Coherence

1. Introduction

Prosody conveys linguistic and attitudinal contexts on a suprasegmental level (Cole, 2015; Patel et al., 2011). Acoustic cues associated with prosody include fundamental frequency (f0), intensity, and word/syllable duration, which are perceived by listeners as pitch, loudness, and length, respectively (Patel et al., 2011). Multiple studies have attributed individual cues to certain contexts: for example, f0 increase is a signal for word stress (Tyler & Cutler, 2009), duration is relevant for speech segmentation (Matzinger et al., 2021), and intensity is associated with vowel quality (Cutler, 2005). While some researchers provide evidence for f0 being the most important cue for signalling stress, others present support for duration and intensity being just as important and even "traded" for f0; a phenomenon known as *cue-trading* (Patel et al., 2011). An intriguing aspect of cue-trading is that it varies between individuals; for example, listeners can generally use aspects of cue-trading to perceive stress even if the speaker's cue patterns differ from their own (Yu, 2022; Patel et al., 2011). However, the factors that motivate these individual differences remain unanswered. Differences in listeners' prosodic perception may be due to differences in individual cognitive processing styles, and these differences in cognitive processing styles have been associated with varying levels of autistic traits as measured by Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001) score (Bishop et al., 2020; Yu, 2010). Additionally, it is important to investigate prosodic cue-trading in the context of autism spectrum disorders as it was noted as being atypical by Kanner (1943) in both children (Wang et al., 2001) and in adults (Gervais et al., 2004).

The present study aims to examine whether the cue-trading of several prosodic cues differs comparing how people with different levels of autistic traits as indicated by AQ score produce and perceive prosody. More so, the study aims to find out if differences in prosodic cue-trading, which occurs when changes in one or more prosodic cues are offset by changes in other prosodic cues while maintaining the same acoustic percept (Flaherty et al., 2017), exist on a continuum dependent on one's level of autistic character traits.

1.1 Literature Review

1.1.1 Autism Spectrum Disorder (ASD)

Autism Spectrum Disorder (ASD) covers a wide breadth of disorders characterized by deficits rooted in social, communicative, and imaginative function (Wing, 1981; Stewart & Ota, 2008). Cognitive theories of ASD have long argued that the speech impairments experienced by individuals with ASD can be attributed to differences in cognitive processing styles relative to neurotypicals (i.e., those who do not display atypical patterns in thought and function).

The Weak Central Coherence (WCC) theory suggests that individuals with ASD have a cognitive processing style which provides them with the enhanced ability to focus on detail but also a weakened ability to create a meaningful whole from this information (Stewart & Ota, 2008; Yu, 2010). In the context of prosodic cue-trading, WCC would predict enhanced perception of prosodic cues, but an inability to form a meaning (e.g. word stress) from the variation of these cues. Evidence for the WCC found that during a homograph reading task, preceding sentence context had less of an effect on the pronunciation of homographs in ASD participants than the typically developed (TD) group (Happe, 1997). Similarly, it was found that ASD participants were less likely to use a preceding sentence context to produce an appropriate homograph and to understand syntactically and lexically ambiguous sentences which were auditorily presented to them (Jolliffe & Baron-Cohen, 1999). Additionally, Stewart & Ota (2008) found a negative correlation between segment identification toward the real word end of a continuum ranging from nonword to real word and AQ score. In other words, the higher one's AQ score was, the less influenced they were by lexical knowledge in phonetic perception. As well, a 2010 study by Nieuwland, Ditman, & Kuperberg examined event-related potential (ERP) responses to critical words in under-informative statements versus informative scalar statements and found that informative statements elicited larger N400 responses than under-informative statements for those who scored higher on the communication subscale of the AQ. Finally, it was observed that people with higher working memory and higher AQ scores are less affected by phonotactic context in sibilant perception (Yu et al. 2011).

The Enhanced Perceptual Functioning (EPF) theory suggests that those with ASD demonstrate heightened sensitivity and enhanced perception of low-level, purely acoustic stimuli. Specifically, Stewart et. al (2018) found that autistic character traits, as measured by AQ score, correlated significantly with one's ability to discriminate low-level pitch and time, but not intensity. Thresholds in both auditory domains decreased with increasing AQ, suggesting that autistic character traits predict enhanced low-level perception of pitch and time. However, findings in relation to enhanced perception of time are inconsistent: another study by Chen & Peng (2021) involving Mandarin-speaking adolescents with ASD showed a much wider boundary width and lower peakedness score, which refers to the difference between the discrimination of between-category pairs and within-category pairs, during both a voice onset time (VOT) identification and discrimination task.

There are fewer studies that examine ASD differences in production, further motivating our methods in this domain. Turnbull (2015; 2019) found that talkers with higher AQ produced shorter word durations when instructed to speak to someone with a hearing impairment than those with a lower AQ and displayed greater word and vowel reduction. In another task by Bishop et al. (2021), speakers with high AQ produced more compact vowel spaces (and thus less acoustically distinct vowel categories) when directed to produce CVC wordforms embedded in carrier sentences that were designed to control for certain prosodic properties. Similarly, in both a native vowel production and artificial vowel production task, autistic participants displayed less variation in their attempts at imitating the vowels, showing that they have reduced phonetic flexibility (Kissine et al., 2021).

In sum, these differences in speech processing in individuals with ASD might be related to variance in cognitive processing styles relative to their neurotypical counterparts. This is particularly apparent in relation to bias for local processing, which has the potential for weak top-down processing, as the Weak Central Coherence model would predict, or highly developed low-level processing, as the Enhanced Perceptual Functioning model would predict (Yu & To, 2020). There already exists a vast body of research on the connection between speech production and perception of segmental cues, and for individual-variability patterns: for phonetic cue weighting in the production and perception of native and non-native stops see Schertz et al. (2015), for English sibilant-vowel coarticulation see Yu (2019), and in the context of phonetic imitation see Kim & Clayards (2019). However, the perception-production link for suprasegmental cues has received less attention. Thus, this study is interested in the individual-variability patterns of prosodic speech production and perception, in particular with respect to individuals with varying levels of autistic traits.

1.1.2 Lexical Stress

Stress is the accentuation of words within sentences, or of syllables within words (Cutler, 2005). Languages like English allow stress placement inside a word by stressing a certain syllable and producing other syllables unstressed. Longer words may have both primary and secondary stress. Stress can also be applied to syllables within a word to indicate important information about that word and possibly change the meaning. For instance, sometimes syllable stress is used to distinguish the meaning between two identical word forms like in words such as '**RE**bel' as a noun and 're**BEL**' as a verb (stress highlighted with capitals and bold print).

Although stress is a prosodic phenomenon—it is expressed through changes in fundamental frequency, intensity, and duration (and other cues) of a word, segment, or syllable—it can sometimes have segmental consequences. Specifically, pairs of English words with lexical stress differences usually also differ in vowel quality (Cutler, 1986). Using the example of '**RE**bel' as a noun and 're**BEL**' as a verb again, these word forms don't only differ in the prosodic realizations of their first and second syllables, but these syllables also contain different vowels altogether. In each pair, the stressed syllable contains a full vowel while the unstressed syllable contains a schwa. Thus, these pairs of words differ both segmentally and suprasegmentally. In contrast to other languages (e.g. Spanish), there are few pairs of lexically distinct words in English that differ solely on a suprasegmental dimension (Cutler & Jesse, 2005). In fact, Cutler (1986) found fewer than a dozen. They include: forBEAR—FORbear and TRUSTy—trusTEE. So, pairs such as the aforementioned ones can help us understand the degree to which a person uses prosodic information compared to segmental information during speech production.

Studies on lexical stress production and perception involving participants with ASD have yielded inconsistent results. For example, Paul et. al (2005) administered a set of 12 tasks assessing the production and perception of grammatical and pragmatic prosody in a group of children clinically diagnosed with ASD. They found that the ASD group showed no difference in their ability to disambiguate between nouns and verbs (e.g. REcall—reCALL) compared with the typically developed (TD) group. On the other hand, they found that the ASD group had significantly less percentage of correct responses when they were asked to produce these lexical stress differences. Moreover, a 2010 study by Grossman et. al reported similar findings on the perception of lexical stress: the ASD group showed no difference in their ability to disambiguate between compound nouns and noun phrases than the TD group. They also reported that the ASD group was able to accurately produce the prosodic cues necessary to disambiguate compound nouns from noun phrases, like the TD group, with the only difference being that the ASD group produced atypically longer word durations than the TD group, showing that they possess basic understanding of lexical stress rules, but that ASD speech production differs from that of TD production.

1.1.3 Prosodic Focus Marking

As mentioned in the previous section, stress can be applied to syllables in words or words in sentences. The present section will concentrate on the stress patterns applied to words within sentences, also known as prosodic stress. Prosody refers to certain features (e.g., pitch, duration, and loudness) that are suprasegmental in nature, meaning that they occur above the segment, and how these features contribute to pragmatic or syntactic meaning and interpretation. Prosodic features may occur within natural stress patterns of a given language, but they may also be used deliberately for emphasizing parts of an utterance which the speaker wants to receive selective attention, also known as focus. This thesis concerns the latter.

In a sentence, information can either be in focus, or accented, or it can be out of focus, or unaccented. This is demonstrated in Figure 1, adapted from Roessig et al. (2022).

In Focus / Focused ~ **Accented**	**Out of Focus / Background** ~ **Unaccented**
Context: Liz and Will gathered at Liz's place for dinner. *Will poured Liz more wine.*	
What did Will pour for Liz?	What did Will pour for Liz?
Will poured Liz more **WINE**.	**Will** poured **Liz** more **WINE**.

Figure 1. Examples of words in sentences that appear in and out of focus.

In English, a nuclear pitch accent is usually placed on the constituent receiving focus (Roessig et al. 2022). English, as well as many other languages, makes a distinction between nuclear and prenuclear prominence. Nuclear prominence marks the structural head of a prosodic phrase and its location in a prosodic phrase depends on the pragmatic meaning the speaker intends to convey (Chodroff & Cole, 2018). On the other hand, prenuclear prominence is any pitch accent preceding the nuclear pitch accent in a given intermediate phrase and is often regarded as "secondary" or "optional" as it is generally considered to be unrelated to the information structural representation (Bishop, 2013). When multiple constituents are in focus, this is known as broad focus. Typically, the last argument in the broad focus domain receives the nuclear pitch accent and is referred to as the "focus exponent," while the surrounding constituents receive prenuclear pitch accents (Roessig et al., 2022). Alternatively, narrow focus occurs when a constituent receives selective attention. A constituent in narrow focus will always receive the nuclear pitch accent while the surrounding constituents will be in background focus. Take the following context from our production experiment as an example:

Context: Lynn and Neil went for a late-night walk. *Lynn and Neil gazed at the moon*.

Q: Who did Lynn gaze at the moon with?

A: *Lynn and* **Neil** *gazed at the moon*.

The first two sentences form a contextual passage. The second sentence in italics occurring in the contextual passage contains all target information. In a trial, the first time the target sentence is uttered, every constituent is in broad focus because they are all receiving an equal amount of selective attention. When asked a contextual question about the passage, using the same sentence to answer the question elicits 2 types of focus: background and narrow. In the answer, the bolded word (i.e., Neil) occurs in narrow focus because it receives selective attention—it contains the necessary information to answer the question. As a result, the speaker uses a combination of prosodic parameters to elicit stress on the word for the benefit of the listener (i.e., so that they know it is occurring in narrow focus). Figure 2 adapted from Roessig et al. (2022) illustrates the difference between broad and narrow focus.

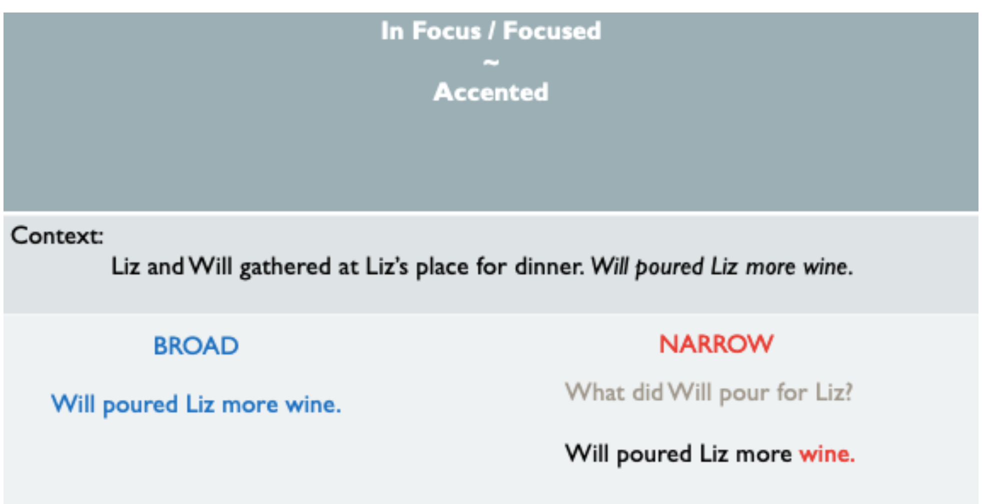

Figure 2. Examples of words in sentences occurring in broad and narrow focus.

On the other hand, when information is out of focus or unaccented, it is considered part of the background (Roessig et al. 2022). The background domain contains the information vital to the listener's knowledge; information that is given or to be expected. Therefore, constituents in the background domain do not receive a nuclear pitch accent and are often completely unaccented, or marked with a low accent (L*) in English. The surrounding constituents in the answer are in background focus because the information has already been given in a previous context and does not help to answer the question at hand.

Although this thesis makes a distinction between background, broad, and narrow focus, it is worth noting that there also exists contrastive focus. Contrastive focus is a type of narrow focus that is used to correct an alternative expression (Roessig et al., 2022). This is illustrated below using the same context from before as an example:

Context: Lynn and Neil went for a late-night walk. *Lynn and Neil gazed at the moon*.

Q: Did Lynn and Thomas gaze at the moon?

A: No, *Lynn and* **Neil** *gazed at the moon*.

In this scenario, "Neil" is in contrastive focus as it is correcting the information that was implied in the question; Lynn gazed at the moon with Neil, not with Thomas.

While a breadth of research establishes that focus in English is prosodically marked by changes in intensity, f0, and duration, many studies look primarily at f0, hence motivating us to examine intensity and duration as well. Furthermore, there are fewer studies that investigate alternative factors that influence the magnitude of prosodic cues such as: sentence position (i.e., initial, medial final), syllable structure (e.g., monosyllabic vs. disyllabic), and focus type (i.e., background, broad, and narrow)—associated with prosodic focus marking. A study by Breen et.

al (2010) found that narrow focus was marked by longer duration, higher mean and maximum f0, and higher maximum intensity compared to background focus in Mainstream American English (MAE). Additionally, they found that across three focus locations (i.e., subject, verb, object), the intended focus location was consistently produced with higher maximum intensity, longer duration, and higher mean and maximum f0.

However, numerous studies have found differences in prosodic focus marking across regional dialects of English. In an eye-tracking study by Arnhold et. al (2020), 42 native speakers of Canadian English followed instructions produced by native speakers of British English to move objects on a screen while their eye movements were tracked, and they found that native British English speakers used rising pitch accent to disambiguate between given and new referents while native speakers of Canadian English used falling accents as a cue for givenness during the task. Moreover, a 2022 study by Kim and Arnhold recruited 38 native speakers of Western[1] Canadian English and had them produce sentences which elicited broad or narrow focus in various locations (i.e., subject, verb, object). They found that while their results for duration and maximum intensity mimicked those of MAE, mean intensity and f0 parameters differed (i.e., mean intensity was lower in the narrow focus condition than the broad focus condition; there was no distinction in f0 range between broad and narrow focus and between broad and background focus).

Further motivating the aims of our study, numerous studies have observed differences in prosodic production and perception between contrasting AQ and ASD populations. In a 2020 study by Bishop, Kuo, & Kim looking at the perception of prosodic prominence using the Rapid

[1] We assume that the speakers spoke a Western Canadian English dialect since all participants were undergraduate students at University of Alberta and native speakers of Canadian English. As well, upon personal communication, the researchers confirmed that most speakers were born and raised in Alberta.

Prosody Transcription (RPT) task, higher scores on the communication subscale of the AQ were inversely related to the likelihood of perceived prominence for words of lower metrical prominence.

1.2 Aims of this Study

The current study presents experiments examining the high-level production and perception of the prosodic cues' fundamental frequency, intensity, duration, and vowel reduction in individuals with varying levels of autistic traits. One of the main goals of this thesis is to compare one's stressed productions against their unstressed productions to examine differences in these productions, and to determine if these differences correlate with AQ scores. Similarly, another goal is to observe participants' accuracies in the matched perception task, and again to examine if there are correlations with the AQ scores. Taken together, these goals motivate the following research questions:

(1) Does the cue-trading of several prosodic cues (e.g., intensity, duration, f0 and its variation, vowel reduction) differ depending on how people with different levels of autistic traits (indicated by AQ score) produce and perceive prosody?

(2) Do differences in prosodic cue-trading exist on a continuum dependent on one's level of autistic traits?

The prediction is that those with high and low AQ scores will produce and perceive prosody differently. We anticipate similar findings to Stewart et al. (2018), that individuals with high AQ will have a higher proportion of correct responses during the perception task than those with low AQ, especially in pitch and time domains, but not intensity. This prediction would also correspond with EPF, that individuals with higher AQ have enhanced perception of low-level acoustic stimuli. Also in line with EPF, we predict that high AQ individuals will have a higher proportion of correct responses compared to the low AQ group when listening to stimuli manipulated below the normal threshold for perception of prosodic differences, which is more difficult to perceive than the normal perception threshold.

Our prediction for the perceptual domain motivates our prediction about performance during the production task: if individuals with high AQ do not display enhanced perception of intensity, it is possible that this may be characteristic of weakened intensity perception in high AQ individuals, thus, we should expect them to display atypical intensity patterns in their production. Specifically, that they will speak with higher intensity overall to compensate for the weakened intensity perception they may experience, and not use intensity as an indicator of narrow focus since they may use intensity less in comparison to f0 and duration when perceiving focus. Another prediction for the production aspect of this study is motivated by findings from Turnbull (2015; 2019) that individuals with high AQ have a smaller clear speech effect—classified by shorter overall word duration and increased vowel reduction—when instructed to speak for the benefit of someone with a hearing impairment. Keeping this in mind, we predict that high AQ participants will produce shorter overall word and syllable duration, and increased vowel reduction in comparison to the low AQ group when asked to elicit different types of focus.

By addressing these questions, we hope that this study will offer in-depth knowledge about deficits in the speech production and perception of individuals with a high level of autistic traits and will serve as an important input to treatment approaches in speech-language pathology and audiology.

2. Research Methodology

2.1. Participants

Participants consisted of 18 McMaster University students (4 males, 14 females). Participants were either recruited using McMaster's Linguistics Research Participation System (SONA) or using poster advertisements displayed around the McMaster campus and social media. We aimed to sample across the whole university, especially for engineering and sciences where we expected higher AQ scores, as several previous studies only recruited participants who were Humanities or Linguistics students, thus questioning the ecological validity of these studies with respect to not only whole student populations, but for ASD differences as well. Our reasoning for sampling across the whole university to get a wider variety of AQ scores also aligns with the reported pattern in program of study; those with higher AQ scores tend to be drawn towards science-related degrees while those with lower AQ scores tend to be drawn towards humanities and social sciences (Baron-Cohen et al., 2001; Stewart & Ota, 2008).

All participants completed a screening questionnaire where we ensured that participants were native speakers of English, meaning that they spoke English from birth, and that they had normal or corrected-to-normal hearing and vision. In addition, they completed a background questionnaire where they indicated their gender, handedness, program and year of study, and age at which they began to speak Canadian English, as well as the AQ test. This study has been reviewed and received ethics clearance from the McMaster Research Ethics Board (MREB).

2.2. Materials and Stimuli

Participants who were considered eligible for the task after filling out the screening questionnaire were asked to sign a consent form after being briefly introduced to the study. In addition, they also completed the background questionnaire and *Autism Spectrum Quotient* (AQ) test (Baron-Cohen et al., 2021).

2.2.1 Autism Spectrum Quotient (AQ)

The AQ test is a self-administered questionnaire developed by Baron-Cohen et al. (2001) to measure the level of autistic traits in neurotypical populations. Although the test was designed using diagnostic criteria, the test itself is non-diagnostic (Stewart, Griffiths, & Grube, 2018). Around the time that the AQ test was developed, there were no instruments available for measuring the degree to which an adult with normal intelligence has the traits associated with ASD. The motivation for developing such a test arose from the need for distinguishing the degree of caseness in an individual and in scientific comparisons, as well as for screening potentially affected individuals and making referrals for full diagnostic assessments. The original study—which observed (1) adults clinically diagnosed with Asperger's syndrome/High Functioning Autism, (2) randomly selected neurotypical controls, (3) students of Cambridge University, and (4) 16 winners of the U.K. Mathematics Olympiad—found that 80% of clinically diagnosed participants scored above a critical minimum of 32+ when scored on a binary scale ranging from 0-50, whereas only 2% of the control group did so. Yu (2020), which used a 4-point Likert scale to preserve information about individual differences, found that the median score of neurotypical participants was 116.5 compared to 102 (Stewart et al., 2008) and 110.05 (Yu, 2010) meanwhile the median of the ASD participants was 132.

The AQ consists of five subscales, with ten questions in each. The questions are framed as such: "I would rather go to the library than a party," and participants' responses are measured on a scale ranging from 'definitely agree' and 'slightly agree' to 'slightly disagree' and 'definitely disagree.' Participants' AQ responses were scored on a Likert scale (1-4), following Stewart & Ota (2008), Yu (2010; 2016), and Kington et al. (2015), because this method preserves information about individual differences.

It is also worth noting that in the first study on the use of the AQ, Baron-Cohen et al. (2001) reported sex differences: male students scored higher than female students overall and across all subscales except for attention to detail but no sex differences in the clinically diagnosed group were observed. This aligns with findings by Ruzich et al. (2015) who performed a comprehensive systematic review of studies that administered the AQ to participants and found a sex difference in neurotypicals taking the questionnaire, but not in those who are clinically diagnosed with ASD. Additionally, these results are in line with the Extreme Male Brain (EMB) theory of autism, which views autism as an extreme of the normal male brain profile—males are significantly better at systemizing than empathizing, whereas the female brain is the opposite (Asperger, 1944; Baron-Cohen, 2002).

2.2.2 Focus Production Task

For this task, participants were required to produce target words in the context of short passages that were presented to them via a PowerPoint slideshow. All participants were recorded using a high-quality microphone (Rode NT1A) positioned around 10 cm away from the speaker's lips, with the microphone being horizontally off-centre by approximately 30–45 degrees and Focusrite Scarlet audio interface. All recordings took place in the soundproof booth in McMaster University's Phonetics Lab. Loudness recording differences were controlled for by the researcher manually adjusting the input gain settings for each participant's recording.

Prior to recording, all participants engaged in a practice session structured similarly to the task but not containing any of the target stimuli to ensure that they understood the task. During this session, the researcher did not produce the stimuli for the participants to ensure that each participant would use their own specific stress pattern for this experiment. When beginning the task, participants produced target words in the context of a short passage. One trial in the focus production task is displayed in Figure 3.

Figure 3. An example of a trial in the focus production task.

This initial production was used to obtain broad focus measurements. Then, they were prompted to press the space bar and a contextual question about the passage appeared, which they were also instructed to read out loud. After being prompted to press the space bar a second time, the answer to the question was provided to them. The answer was always the last sentence of the contextual passage. A cartoon image appeared under the target word as a prompt for participants to produce stress on that word. This stressed production was used to obtain our narrow focus measurements. The surrounding target words, produced with no stress, were used to obtain our background focus measurements.

After completing one block of the whole experiment, the participants were prompted to complete a second block, except this time, they were instructed to complete the task for the benefit of someone who may have a hearing impairment, or whose first language may not be English. This block was motivated by the findings from Turnbull (2015; 2019), where higher AQ-scoring participants elicited a smaller clear speech effect (i.e. they produced shorter word and vowel durations) than lower AQ-scoring participants. Practice trials were provided to ensure the participant was completing the task correctly before recording began.

The target stimuli came from a list created by the researchers. The targets varied in their number of syllables (i.e., monosyllabic condition vs. disyllabic condition), and sentence position (i.e., initial, medial, or final). Additionally, the targets consisted of six male names (3 monosyllabic, 3 disyllabic), six female names (3 monosyllabic, 3 disyllabic), and 6 objects (3 monosyllabic, 3 disyllabic). All syllables followed a CV(C) structure, and whenever possible voiceless phonemes within the stimuli were avoided to allow for accurate f0 tracking during the later analysis. In addition, nine filler trials were included in the task. See Table 1 for the full stimulus list.

Table 1. List of stimuli used during the focus production task separated by syllable condition.

Stimuli List			
Monosyllabic	Liz and Will gathered at Liz's place for dinner. Will poured Liz more wine.	1. Who poured Liz more wine?	**Will** poured Liz more wine.
		2. Who did Will pour more wine for?	Will poured **Liz** more wine.
		3. What did Will pour for Liz?	Will poured Liz more **wine**.
	Lynn and Neil went for a late-night walk. Lynn and Neil gazed at the moon.	1. Who did Neil gaze at the moon with?	**Lynn** and Neil gazed at the moon.

		2. Who did Lynn gaze at the moon with?	Lynn and **Neil** gazed at the moon.
		3. What did Lynn and Neil gaze at?	Lynn and Neil gazed at the **moon**.
	Reid knew his wife, Rose, had a stressful day at work. Reid cooked Rose a meal.	1. Who cooked a meal for Rose?	**Reid** cooked Rose a meal.
		2. Who did Reid cook a meal for?	Reid cooked **Rose** a meal.
		3. What did Reid do for Rose?	Reid cooked Rose a **meal**.
Disyllabic	Lola and Logan went grocery shopping. Lola handed Logan a lemon.	1. Who handed Logan the lemon?	**Lola** handed Logan a lemon.
		2. Who did Lola hand the lemon to?	Lola handed **Logan** a lemon.
		3. What did Lola hand to Logan?	Lola handed Logan a **lemon**.
	Lauren and Ryan went to the petting zoo. Lauren and Ryan pet the llama.	1. Who did Ryan pet the llama with?	**Lauren** and Ryan pet the llama.
		2. Who did Lauren pet the llama with?	Lauren and **Ryan** pet the llama.
		3. What did Lauren and Ryan pet?	Lauren and Ryan pet the **llama**.
	Nolan wanted to buy Lily's concert tickets. Nolan gave Lily some money.	1. Who gave Lily some money?	**Nolan** gave Lily some money.
		2. Who did Nolan give some money to?	Nolan gave **Lily** some money.
		3. What did Nolan give to Lily?	Nolan gave Lily some **money**.

2.2.3 Lexical Stress Task

After completing the Focus Production task, participants next completed a Lexical Stress task. During this task, participants were required to produce target words that occurred in sentences. These sentences were presented to participants via the same PowerPoint slideshow and were accompanied by an animated picture to reinforce the meaning of the target word.

Similar to the structure of the Focus Production task, once the participants completed one block, they were then instructed to complete another block for the benefit of someone who may have a hearing impairment, or whose first language may not be English.

The targets for the lexical stress task came from a list created by the researchers and were roughly based on the study by Cutler (1986). In total, there were fourteen words that varied on both a segmental and suprasegmental dimension (i.e. contained a vowel shift) (e.g. **OB**ject vs. ob**JECT**) and six words that varied on purely a suprasegmental dimension (e.g. forearm vs. FOREarm). In addition, we incorporated four filler words with stress fixed on either the first or second syllable but whose lexical category also changed (e.g. 'model', whose stress occurs on the first syllable when it is used as a noun or a verb). Table 2 provides the full stimulus list.

Table 2. List of stimuli used in the lexical stress task separated by stress pattern opposition and no stress shift.

Stress Pattern Opposition
Record 1. He broke the world **record** (n.) for being the tallest man. 2. She asked to **record** (v.) her professor's lecture.
Object 1. An unknown **object** (n.) was in the box. 2. I can't **object** (v.) to paying for parking.
Reject 1. Their goalie was a **reject** (n.) from another team 2. She tried to **reject** (v.) him but he persisted.
Address 1. I need your **address** (n.) so I can mail you your cheque. 2. There's an issue I must **address** (v.) with you immediately.
Desert 1. Anne lived in a **desert** (n.) area full of cacti. 2. Sue wanted to **desert** (v.) her hometown.
Rebel 1. Robin Hood was a **rebel** (n.) against tyranny. 2. My friends **rebel** (v.) against their parents.
Relay 1. The school **relay** (n.) team trained for the race. 2. The secretary will **relay** (v.) your message for you.
No Shift
Forearm

	1. The man's **forearm** (n.) contained many tattoos. 2. The coach aimed to **forearm** (v.) her team with strategies.
Retail	1. Kate swore she'd never work in **retail** (n.) again. 2. Tom asked his parents to **retail** (v.) how they met.
Discount	1. Many stores offer a **discount** (n.) on items not selling quickly. 2. Many people **discount** (v.) my opinion since I look young.

2.2.4 Perception Task

Once participants completed both production tasks, they then completed the perception task. The target stimuli used in the perception task were recorded by two native English speakers from the Greater Toronto Area (one male, one female). These speakers were recorded completing the focus production task and consented to their recordings being used as auditory stimuli in the perception task. We obtained their broad focus productions (i.e. where each constituent receives the same amount of stress) and manipulated the intensity, f0, and duration of the stressed syllables of all target words. Additionally, we incorporated a *degree-of-manipulation* condition so that stimuli increased by a so-called half or full manipulation of the phonetic parameter in question. The experimental conditions are summarized in Table 3 below:

Table 3. List of manipulations applied to stimuli used during the perception task.

	F0	Intensity	Duration
Baseline	-	-	-
Half	1.5 semitones increase	3 dBs increase	33% longer duration
Full	3 semitones increase	6 dBs increase	66% longer duration

In the f0 condition, we increased targets by 1.5 semitones in the half-manipulation condition and 3 semitones in the full-manipulation condition. In the intensity condition, targets were increased by 3dB in the half-manipulation condition, and 6 dBs in the full-manipulation condition. In the duration condition, the targets were made 33% longer in the half-manipulation condition and 66% longer in the full-manipulation condition. All full-manipulation increases were taken from what a number of studies (e.g. Stewart, Griffiths & Grube, 2018) consider to be within the normal range of prosodic productions, and the half-manipulation increases were included to have a condition that was exactly half of the full manipulation, and thus expectedly below the normal production and thus presumably perception threshold for prosodic differences.

Furthermore, for disyllabic stimuli we incorporated conditions where we took the speakers' narrow focus productions and spliced (1) the stressed syllable of the target word produced in narrow focus onto the target produced in broad focus, and (2) the unstressed syllable of the target word produced in narrow focus onto the target produced in broad focus.

As a final step, all stimuli were loudness-normalized to be presented at 70 dB intensity, to allow for equal auditory representation of all stimuli without some stimuli played louder than others (and thus to avoid biases).

Once it was time to begin the task after completing a practice session to familiarize participants with the task and to ensure the volume was at a comfortable level for them, participants heard the following instructions from the principal investigator: "Now you will listen to different speakers saying sentences. Your task is to indicate which word sounded the most focused to you by clicking that word on the screen." The task was programmed using Gorilla, an online experiment builder. Participants wore Sennheiser HD 598 headphones with linear frequency response, connected to a Focusrite Scarlett 2i2 audio interface, while completing the task. The stimuli were presented in a random order for each participant, followed by a full repetition of the task. The data was saved to Gorilla then exported into a Microsoft Excel file.

3. Results

This chapter summarizes the results beginning with the perception task followed by the focus production task.

3.1 AQ Score

AQ scores ranged from 89 to 143 (M = 112.5, SD = 11.8). To explore, for both perception and production, the correlation between prosodic features and AQ scores, we grouped the participants into low-AQ (N = 10) and high-AQ (N = 8) groups by taking the lowest-third and highest-third scores of the total number of participants so that anyone with a score below 104.5 was classified as part of the low-AQ group, and anyone with a score above 124 was classified as being part of the high-AQ group. AQ score for the low-AQ group ranged from 89 to 103 (M = 98.7, SD = 4.7) and was female-dominant with 9 female participants and 1 male participant. AQ score for the high-AQ group ranged from 125 to 143 (M = 130, SD = 6) and contained 5 female participants and 3 male participants. The gender difference we observed in low and high-AQ groups is in correspondence with previous literature that observed the same sex difference in neurotypicals who completed the AQ, but not in individuals clinically diagnosed with ASD (Ruzich et al., 2015).

3.2 Perception Task

Logistic Mixed-Effects Regression models were fit to the data to test our research hypothesis that high-AQ participants perceive the prosodic features f0, intensity, and duration more accurately than low-AQ participants. The models were used to test the hypothesis that high-AQ participants are more accurate in perceiving fine-grained, smaller manipulations in f0, intensity, and duration, reflected in their performance during the below-threshold half-manipulation condition. The logistic regression analysis was executed using the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) packages in R statistical software (R Core Team, 2022) with a Simple Coding scheme. AQ score (high and low) and Degree of Manipulation (half and full) as well as the interaction between them were included as fixed effects, and Listener and Item as random effects. All pairwise comparisons were computed using the *emmeans* package (Lenth, 2023) in R. As well, low AQ score and half Degree of Manipulation were set as the reference levels. The syntax used for analysis is included below:

```
perception_f0_model = glmer(Correct ~ AQ.bin*Half_Full+(1|Listener),  family
= "binomial",  perception_f0)
```

A binary logit model was used due to the collection of our dependent variable—correct responses were recorded using a binary measure of either 0 (for incorrect responses) or 1 (for correct responses). The next three sections outline results in statistical validity for each prosodic cue with proportion of correct perception responses measured with the binary response of 0 or 1 as the dependent variable. Fixed effects of AQ (high or low) and degree of manipulation (half or full), and their interactions, were tested with a significance threshold of $p < 0.05$.

Figure 4 shows the proportion of correct responses during the perception task, split by: (1) the degree to which the stimuli were manipulated (full/half) in each panel, (2) the acoustic parameter (from left to right panel: f0, intensity, and duration), and AQ group (coded by different colours). The red bars represent the high AQ group while the blue bars represent the low AQ group. The black horizontal line represents the chance-level of 33%, since each participant always had three different word options to choose between in a given trial.

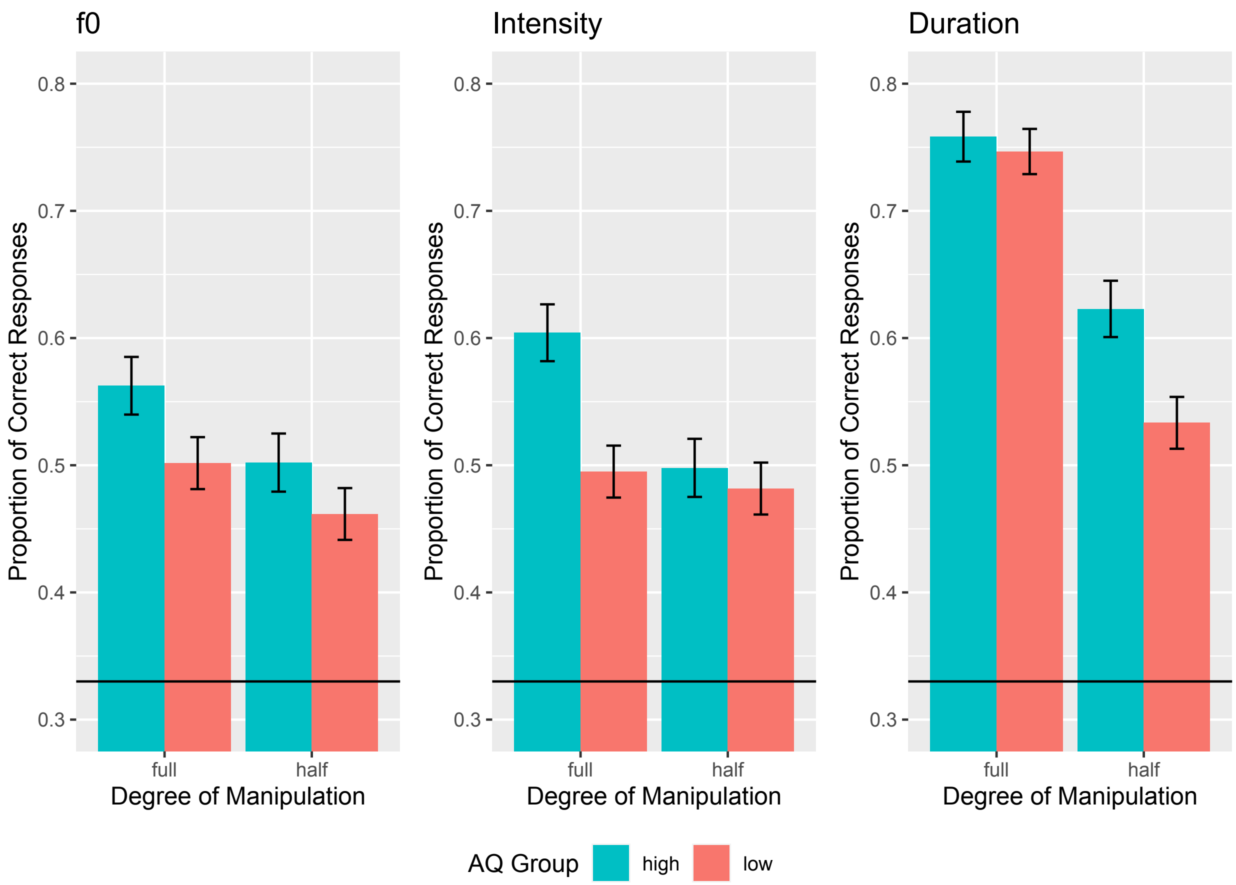

Figure 4. Proportion of correct responses separated by AQ group, degree of manipulation, and prosodic parameter.

Based on previous findings which found that participants with high AQ have enhanced low-level perception of pitch and time, but not intensity (Stewart et al., 2018), we predicted that this accuracy would also manifest in high-level prosodic perception as examined by this study. Looking at Figure 4, it is evident that duration is perceived more accurately than f0 and intensity domains, independent of degree of manipulation and AQ score. As well, there appears to be small differences between duration and intensity, but it is not clear which of these two domains are perceived more accurately.

We predicted that the high-AQ group would have (1) higher proportion of accuracy scores for pitch and time domain, and (2) this higher accuracy should be more pronounced for the half-manipulation condition, since this below-threshold condition is expected to be easier to be perceived compared to the above-threshold full-manipulation condition, due to EPF. This expectation is confirmed in Figure 4 for the pitch and duration domain, for both half and full manipulation conditions. However, an opposite trend is seen for the intensity condition—the high-AQ group perform better in the full manipulation intensity (proportion of correct responses = 0.6) condition than for the half manipulation intensity condition (proportion of correct responses = 0.5), and there is a small difference of 0.02 in the proportion of correct responses between AQ groups for the half manipulation condition. Additionally, going against our predictions, the high AQ group appears to perform better in the full intensity condition than the full f0 condition.

3.2.1 Fundamental Frequency

Figure 5 presents the linear prediction for proportion of correct responses split by AQ group and degree of manipulation. Our results found that there was a significant effect of AQ (z = 2.34, p = 0.02), indicating that individuals with higher AQ scores had a higher proportion of

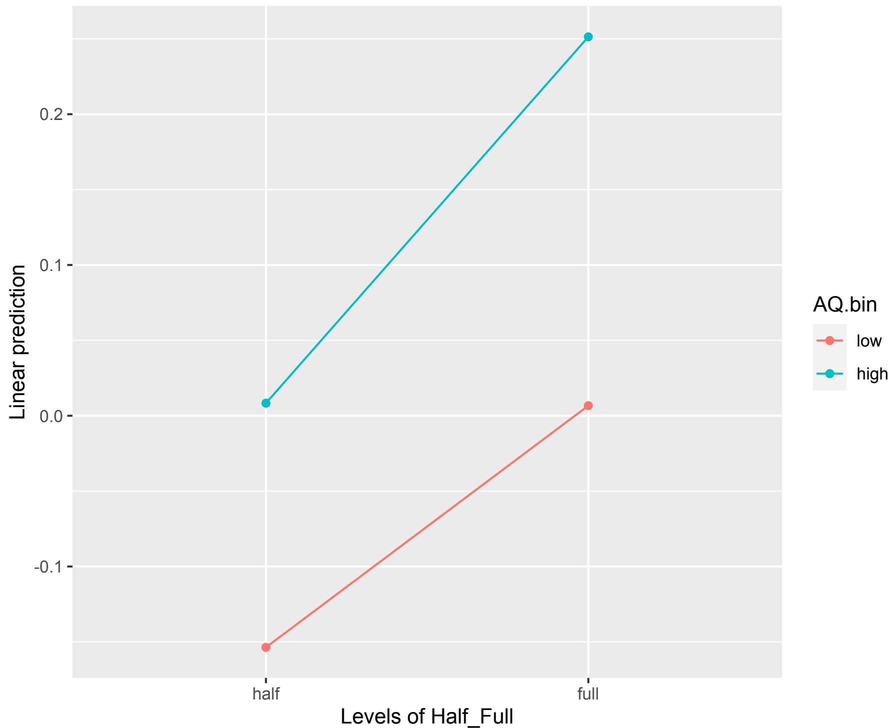

Figure 5. Linear prediction for proportion of correct f0 responses split by AQ group and degree of manipulation.

correct responses to stimuli whose f0 was manipulated (M = 53.23%, SE = 1.61%) than participants with low AQ (M = 48.17%, SE = 1.44%). The model also demonstrated that there was a significant effect of degree of manipulation (z = 2.32, p = 0.02). The interaction plots show that participants are more accurate in perceiving stimuli in the full-manipulation condition containing a 3-semitone increase (M = 54.35%, SE = 1.52%) than stimuli in the half-manipulation condition containing a 1.5-semitone increase (M = 48.89%, SE = 1.52%). The interaction between AQ and manipulation was not significant. Pairwise comparisons found that proportion of correct responses were not significantly different comparing low AQ to high AQ in the half-manipulation condition (β = -0.16, SE = 0.12, z = -1.32, p = 0.55), and also not in the full-manipulation condition (β = -0.26, SE = 0.12, z = -1.99, p = 0.19). Significance and z-scores for all main effects and interactions are listed in Table 4.

Table 4. Significance scores for all main effects and interactions between AQ score and degree of manipulation for f0 responses.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	0.02817	0.04343	0.649	0.5165
AQ.bin2	0.20331	0.08685	2.341	0.0192 *
Half_Full2	0.20164	0.08685	2.322	0.0202 *
AQ.bin2: Half_Full2	0.08268	0.17369	0.476	0.6341

3.2.2 Intensity

Figure 6 presents the linear prediction for proportion of correct responses to stimuli whose intensity was manipulated by AQ group and degree of manipulation. Our results found that there was a significant effect of AQ score (z = 2.53, p = 0.01), implying that high-AQ participants are more accurate in perceiving differences in intensity (M = 55.1%, SE = 1.6%)

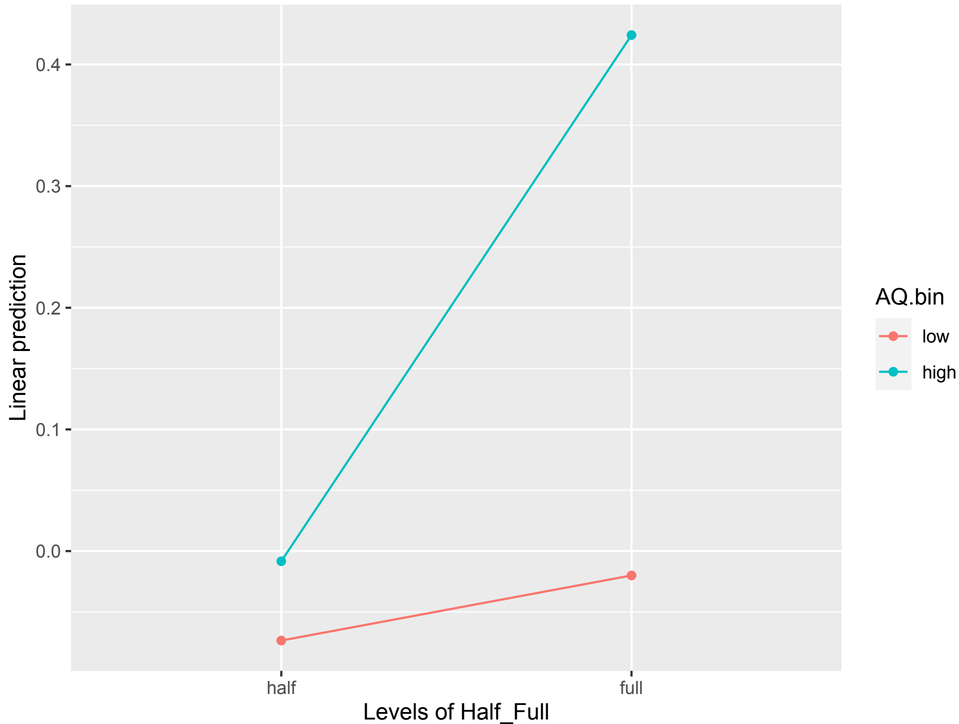

Figure 6. Linear prediction for proportion of correct intensity responses split by AQ group and degree of manipulation.

than those with low AQ (M = 48.83%, SE = 1.44%). Additionally, the model also found a significant effect of degree of manipulation (z = 2.78, p = 0.005), suggesting that participants are significantly more accurate in the full-manipulation condition when stimuli were increased by 6 dBs (M = 54.35%, SE = 1.52%) than in the half-manipulation condition when stimuli received a 3 dB increase (M = 48.89%, SE = 1.52%). The model also showed that AQ group interacted with degree of manipulation (z = 2.17, p = 0.03), and Figure 6 shows their interaction pattern. Pairwise comparisons demonstrated that the proportion of correct responses was only significantly different between AQ groups in the full-manipulation condition (β = -0.44, SE = 0.13, z = -3.32, p = 0.005), but not in the half-manipulation condition (β = -0.07, SE = 0.13, z = -0.49, p = 0.96), with the larger differences demonstrated in the interaction plot. Significance and z-scores for all main effects and interactions are listed in Table 5.

Table 5. Significance scores for all main effects and interactions between AQ score and degree of manipulation for intensity responses.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	0.08055	0.05028	1.602	0.10910
AQ.bin2	0.25470	0.10055	2.533	0.01131 *
Half_Full2	0.24301	0.08729	2.784	0.00537 **
AQ.bin2: Half_Full2	0.37903	0.17458	2.171	0.02992 *

3.2.3 Duration

Figure 7 presents the linear prediction for proportion of correct responses to stimuli whose duration was manipulated by AQ group and degree of manipulation. Upon visual inspection of Figure 7, it would appear that participants with a higher level of autistic traits are

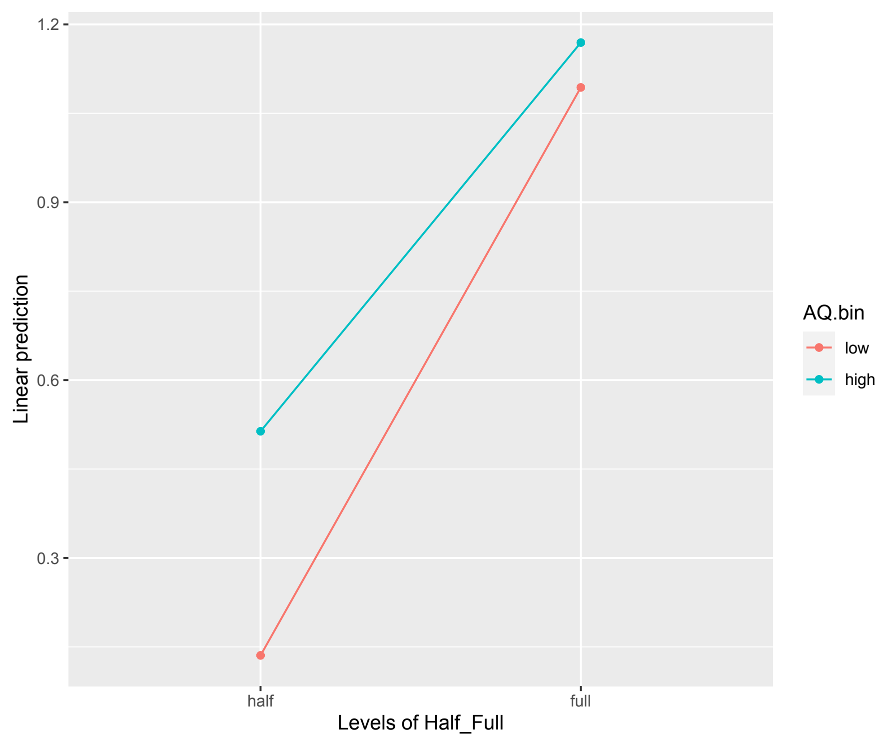

Figure 7. Linear prediction for proportion of correct duration responses split by AQ group and degree of manipulation.

more accurate in detecting below-threshold differences in duration, aligning with our research hypothesis. Our model, however, found no significant effect of AQ score (z = 1.41, p = 0.16), implying that participants classified as high AQ do not demonstrate a difference when compared to participants classified as low AQ. Additionally, the model found a significant effect of degree of manipulation ($z = 8.46$, $p = <2e-16$), suggesting that participants have a significantly higher rate of accuracy in perceiving stimuli in the full-manipulation condition where stimuli were 66% longer in duration ($M = 75.18\%$, $SE = 1.31\%$), than in the half-manipulation condition where stimuli were 33% longer in duration ($M = 57.31\%$, $SE = 1.51\%$). Pairwise comparisons found that the proportion of correct responses between AQ groups was neither significantly different in the half-manipulation condition (($\beta = -0.38$, $SE = 0.18$, $z = -2.1$, $p = 0.15$), nor in the full-manipulation condition ($\beta = -0.08$, $SE = 0.19$, $z = -0.39$, $p = 0.98$). Significance and z-scores for all main effects and interactions are listed in Table 6.

Table 6. Significance scores for all main effects and interactions between AQ score and degree of manipulation for duration responses.

	Estimate	Std. Error	z-value	Pr (>|z|)
(Intercept)	0.72806	0.08025	9.072	<2e-16 ***
AQ.bin2	0.22682	0.16036	1.414	0.157
Half_Full2	0.80684	0.09540	8.458	<2e-16 ***
AQ.bin2: Half_Full2	-0.30241	0.19067	-1.586	0.113

3.3 Focus Production Task

This section will go over the results for the production task beginning with the results for f0, followed by intensity, then duration. Please note that in this thesis, we only report the results for the sentence-medial position –where we expected to see the most stable results due to the possibility of prosodic information getting lost or influenced by external factors in initial and final sentence positions—across all three focus conditions (background, broad, and narrow). As well, due to time limitations, we only report the results for 14 participants (4 males, 10 females).

Linear Mixed-Effects Regression models were fit to the data to test our research hypothesis that participants with higher levels of autistic traits as indicated by AQ score will differ in their production of prosodic cues f0, intensity, and duration to indicate different levels of focus (background, broad, and narrow). The linear regression analysis was executed using the *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) packages in R (R Core Team, 2022). AQ score (high and low) and Focus Type (background, broad, and narrow) as well as the interaction between them were included as fixed effects, and Speaker and Target Word as random effects. All pairwise comparisons were computed using the *emmeans* package (Lenth, 2023) in R. As well, low AQ score and background focus were set as the reference levels. The syntax used for analysis is included below:

```
production_intensity_model = lmer(stressed.intensity.noout_z ~ focus *
AQ.total.bin +(1|speaker)+(1|target), control=lmerControl(optimizer="bobyqa",
optCtrl=list(maxfun=2e5)),  production_intensity)
```

The next three sections outline results in statistical validity for each prosodic cue. For f0 measures, the dependent variable was the normalized f0 range in semitones, the baseline for the normalization was the f0 across each speaker's entire recording. For intensity measures, the dependent variable was the normalized intensity range in dBs across each speaker's entire recording where the baseline for normalization was the speaker's intensity over the entire recording. For duration, the dependent variable was the duration of stressed syllables in milliseconds.

3.3.1 Fundamental Frequency

Figure 8 shows the normalized f0 range in semitones during the production task divided by AQ group and focus.

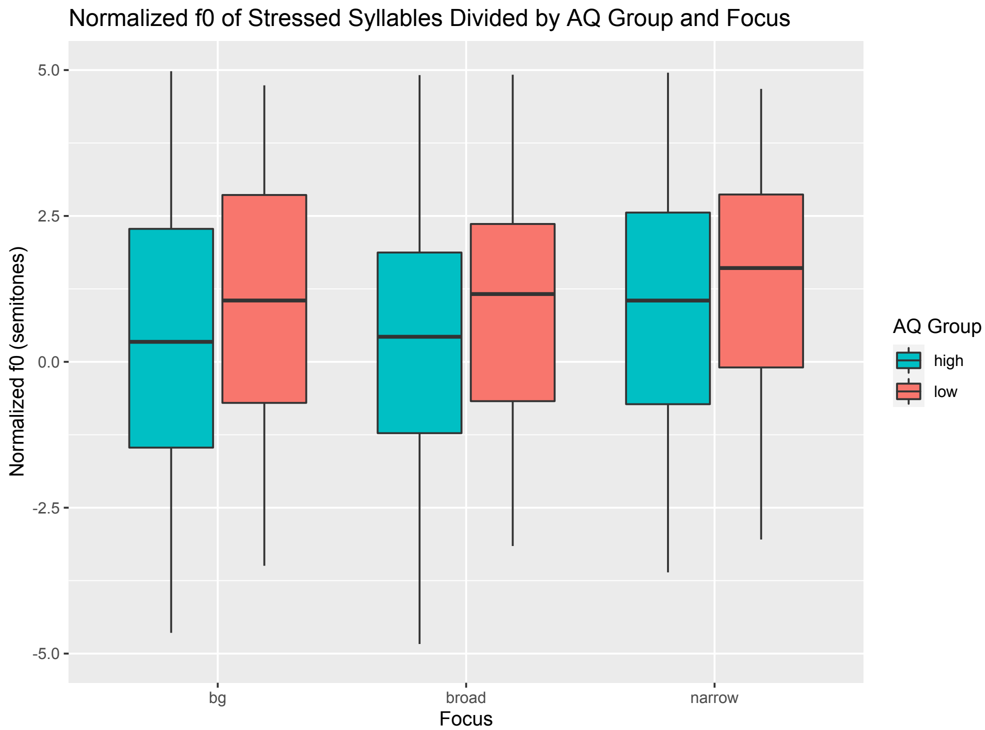

Figure 8. Normalized f0 (semitones) of stressed syllables separated by AQ group and focus condition.

Figure 9 presents the linear prediction for normalized f0 range divided by AQ group and focus. Our model found a significant effect of narrow focus (t = 3.27, p = 0.001), suggesting that narrow focus is consistently produced with higher f0 range (M = 0.86, SD = 1.68) when compared to background focus (M = -0.0008 SD = 1.87), and this is the case for both participant groups. Although no significant effect of AQ score or interactions between AQ score and focus type was found in the model, AQ nearly misses significance (p = 0.054) and it is worth exploring in a future study if AQ reaches significance with more speakers or items. Significance and t-values for all main effects and interactions are listed in Table 7.

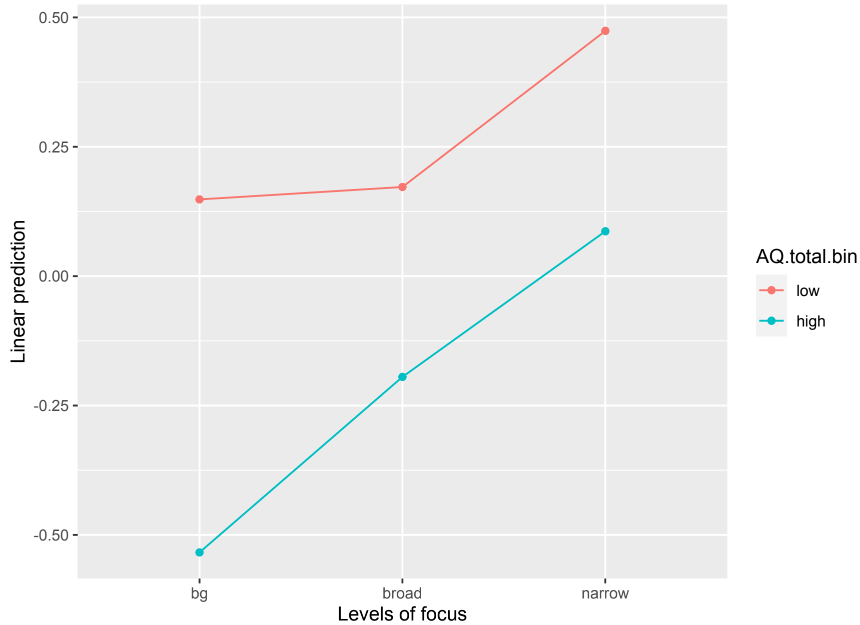

Figure 9. Linear prediction for normalized f0 range separated by AQ group and focus condition.

Table 7. Significance scores for all main effects and interactions between AQ score and levels of focus for f0 range production.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	0.02546	0.16831	9.44160	0.151	0.8829
Focus2	0.18157	0.14130	200.83197	-1.285	0.2003
Focus3	0.47319	0.14490	200.39425	3.266	0.00128**
AQ.total.bin2	-0.47880	0.22450	12.03498	-2.133	0.0542
Focus2: AQ.total.bin2	0.31522	0.28247	200.71227	1.116	0.26578
Focus3: AQ.total.bin2	0.29488	0.28963	200.29884	1.018	0.30984

3.3.2 Intensity

Figure 10 shows the normalized intensity range in dBs divided by AQ group and focus. Relating back to previous literature which found that individuals with higher levels of autistic traits have enhanced perception of pitch and time, but not intensity (Stewart et al., 2018), we wondered if this may be due to a deficit in intensity perception and if this could be characteristic of high-AQ individuals. Due to the link between speech production and perception, we expected

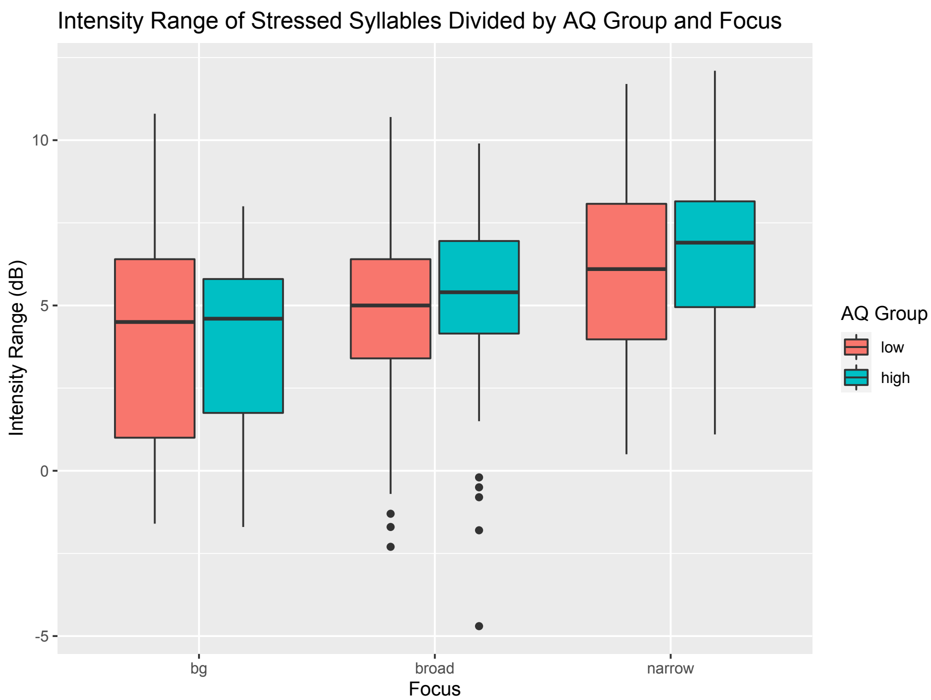

Figure 10. Normalized intensity range (dB) of stressed syllables separated by AQ group and focus condition.

participants with high AQ to produce louder intensity than those with low AQ to possibly compensate for weaker intensity perception. Upon visual inspection of Figure 10, both participant groups appear to be producing higher intensity in their speech as they move from background focus and approach narrow focus. Contrary to our prediction, both participant groups appear to be nearly identical in background and broad focus conditions. In the narrow focus condition, it appears that high AQ participants are producing slightly louder intensity. The following section covers the results that our linear regression model found and determines whether or not this is a statistically significant finding.

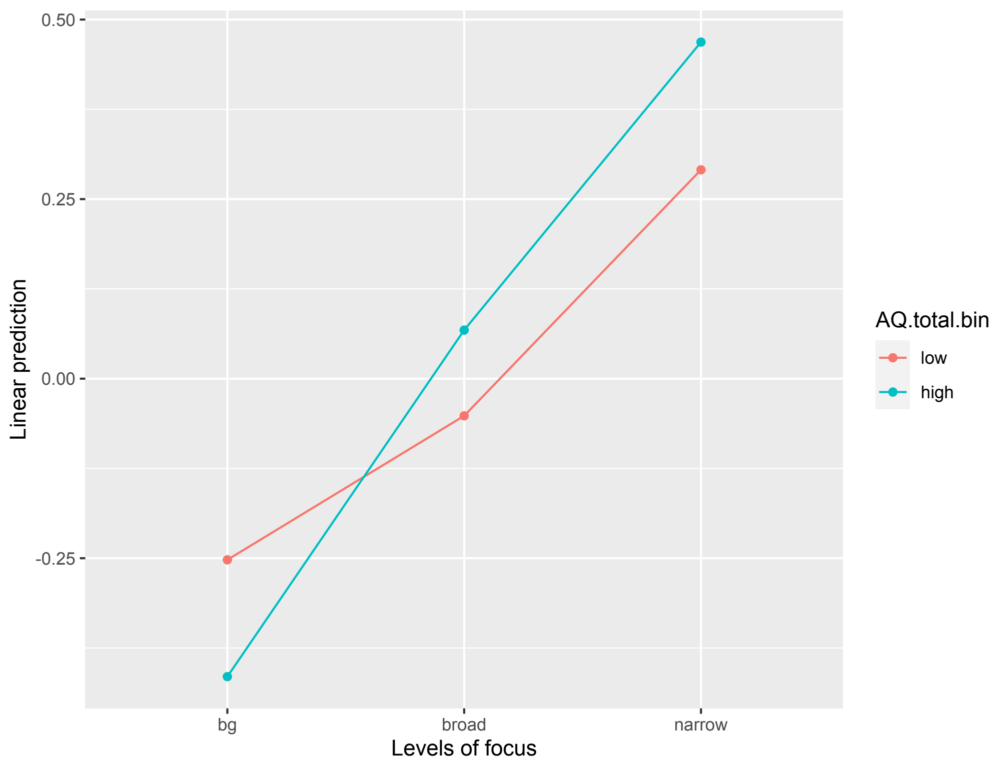

Figure 11. Linear prediction for normalized intensity range separated by AQ group and focus condition.

The linear prediction for intensity range by AQ group and focus is presented in Figure 11. The statistical model found a significant effect of both broad focus ($t = 2.454$, $p = 0.015$) as well as narrow focus ($t = 5.034$, $p = 1.05e-06$), suggesting that broad focus ($M = 5.07$, $SD = 2.96$) and narrow focus ($M = 6.24$, $SD = 2.72$) are both consistently produced with higher intensity range than background focus across both AQ groups. Significance and t-values for all main effects and interactions are listed in Table 8. No interaction was found between focus type and AQ group.

Table 7. Significance scores for all main effects and interactions between AQ score and levels of focus for intensity range production.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	0.01803	0.15620	12.17588	0.115	0.91000
Focus2	0.34149	0.13916	205.17720	2.454	0.015 *
Focus3	0.71325	0.14168	204.82953	5.034	1.05e-06***
AQ.total.bin2	0.04487	0.25192	12.49689	0.178	0.862
Focus2: AQ.total.bin2	0.28200	0.27833	205.18804	1.013	0.312
Focus3: AQ.total.bin2	0.34048	0.28336	204.83571	1.202	0.231

3.3.3 Duration

Earlier research has found that speakers with higher AQ scores displayed a smaller clear speech effect when instructed to speak for the benefit of a listener who has a hearing impairment, and this smaller clear speech effect was observed in shorter overall word durations and greater vowel reduction (Turnbull 2015; 2019). This finding motivated our prediction that speakers with

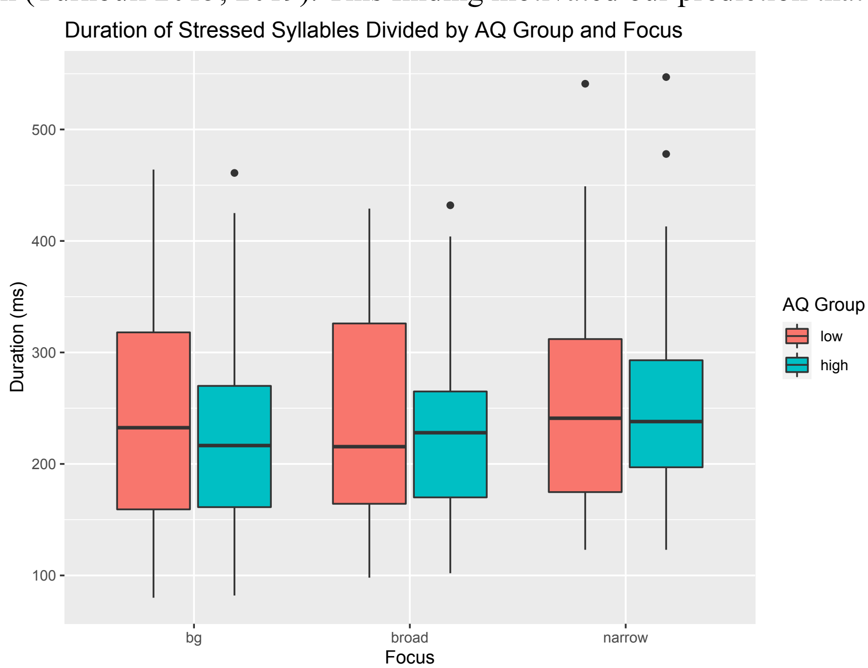

Figure 12. Duration (ms) of stressed syllables separated by AQ group and focus condition.

high AQ would produce speech with shorter durations across all focus conditions. Figure 12, which shows the duration of stressed syllables divided by AQ group and focus, shows that both AQ groups produce nearly identical duration in each of the three focus conditions, but that the groups may differ slightly in their broad focus production, with the high AQ group producing slightly longer duration. When background focus is set to the reference level, the model showed that broad focus ($t = -0.06$, $p = 0.96$) and narrow focus ($t = 1.9$, $p = 0.06$) were not statistically significant. The model also showed that there was no significant effect of AQ ($t = -0.49$, $p = 0.64$). Significance and t-values for all main effects and interactions are displayed in Table 9.

Table 8 . Significance scores for all main effects and interactions between AQ score and levels of focus for duration production.

	Estimate	Std. Error	df	t-value	Pr (>|t|)
(Intercept)	0.005382	0.363290	5.829113	0.015	0.9887
Focus2	-0.004175	0.076578	265.020926	-0.055	0.9566
Focus3	0.156580	0.082369	265.164496	1.901	0.0584
AQ.total.bin2	-0.101647	0.209567	11.782021	-0.485	0.6365
Focus2: AQ.total.bin2	0.070400	0.153156	265.021101	0.460	0.6461
Focus3: AQ.total.bin2	0.132127	0.164753	265.172114	0.802	0.4233

4. Discussion

Given the fact that previous studies report mixed results for the production and perception of prosody in individuals with varying levels of autistic traits, the present research was developed to investigate the cue-trading of the prosodic cues f0, intensity, and duration amongst participants with varying AQ scores in the context of background, broad, and narrow focus types. This chapter will compare the results from the production and perception studies across each acoustic parameter, and how these results relate to the research questions outlined in the aims of our study in Chapter 1.2.

4.1 Fundamental Frequency

Previous studies looking at the perception of f0 involving participants with varying levels of autistic traits have observed enhanced discrimination of f0 (Stewart et al., 2018), aligning with the Enhanced Perceptual Function theory (Mottron et al., 2006) which proposes that individuals with higher levels of autistic traits are better able to detect the small, fine-grained differences in low-level, purely acoustic stimuli. The present study also observed this—when participants were presented with stimuli containing differences in f0, participants with higher AQ scores performed with a higher proportion of accuracy ($M = 53.23\%$, $SE = 1.61\%$) than those with lower AQ scores ($M = 48.17\%$, $SE = 1.44\%$), thus a significant effect of AQ score was found ($z = 2.34$, $p = 0.02$). These findings, therefore, support the Enhanced Perceptual Function theory.

Previous studies looking at the production of f0 involving participants with varying levels of autistic traits have mixed findings, thus making it difficult for us to formulate a research hypothesis. Some studies have found that participants diagnosed with ASD are often perceived as speaking monotonously or machine-like, and therefore with less variability in their production of prosody (Baltaxe & Simmons, 1985; Lord, Rutter, & Le Couteur, 1994), while others have observed more variability in their usage of prosodic cues, especially f0 (Diehl et al., 2008). The present study found no significant results for the production of f0 in individuals with varying levels of autistic traits. However, it is possible that the lack of significance is due to the sample size given our trend approaching significance. Thus, more participants should be included in future research to determine whether AQ score is correlated with f0 change and variability. In our study, the factor AQ score nearly reaches significance ($t = -2.13$, $p = 0.054$), for the produced f0 differences, showing that participants with higher AQ score produce lower f0 ranges than those with low AQ score— or in other words participants with higher AQ scores have lower variability in their f0 range, although the result nearly missed the significance level ($p = 0.054$). This finding corresponds to what has been observed about speech in autistic individuals in that they are perceived to be more monotonous, machine-like, or bizarre (Baltaxe & Simmons, 1985; Lord, Rutter, & Le Couteur, 1994), and thus lacking variability in their f0 production. Alternatively, this finding contrasts with fewer but more recent studies that have observed a higher variability in autistic speakers, especially for f0 (Diehl et al., 2008).

In summary, the results from both the perception and production tasks align with results from previous studies and our research hypothesis. Our results from the perception task align with the EPF as well as Stewart et al. (2018)'s finding that individuals with higher AQ scores display enhanced perception of pitch. In line with the Weak Central Coherence theory, we expected that if our participants exhibited enhanced perception of f0, that they would be less likely to use f0 as an indicator of narrow focus, and therefore produce lower f0 ranges than participants with lower AQ. This is because individuals with high AQ might have a harder time grasping that an increase in f0 is signalling a change in focus and therefore, they may use it as an indicator of focus less than a participant with a lower AQ score would. In light of both cognitive theories presented here, having an enhanced perception of f0 but the inability to view it as an indicator of focus, and therefore, using f0 to signal focus less than others, may have social consequences. On one hand, as the WCC predicts, that such individuals may experience difficulty determining which constituent a speaker is signalling focus which could lead to a misunderstanding of the speaker's intentions. For production, this could mean that a speaker with higher AQ may experience difficulty using f0 in their own speech to signal focus which could lead to them being misunderstood by their audience.

4.2 Intensity

Stewart et al. (2018) observed that participants with higher levels of autistic traits display an enhanced perception of pitch and duration, but not intensity. However, it remains unclear whether people with higher levels of autistic traits experience an enhanced or impaired perception of intensity. Several studies have reported cases of hyperacusis, a condition where one experiences reduced tolerance to sound where sounds presented at the normal hearing threshold are too loud and sounds presented above the normal hearing threshold cause discomfort or pain, in both clinical and personal accounts of autism (O'Connor, 2012; Tyler et al., 2014). In fact, a 1999 study by Rosenhall et al. discovered that 18% of autistic children were unable to tolerate an 80 dB nHL broadband click stimulus. Considering this reduced intensity tolerance, it would stand to reason that people with higher levels of autistic traits might display enhanced perception of intensity, but as a matter of fact, several studies have reported either normal (Alcántara et al., 2012) or even impaired (Kargas et al., 2015) perception of intensity relative to typically developed controls. There are fewer studies that investigate the production of intensity in individuals with varying degrees of autistic traits, as many studies look primarily at f0 since it is commonly associated with being the most relevant cue in signalling focus. The findings from the present study showed that individuals with higher levels of autistic traits displayed an enhanced ability to perceive intensity, aligning more with clinical and personal accounts of hyperacusis in such individuals. Since we originally hypothesized that participants with higher AQ scores would perform less accurately than the low AQ group during the perception task, this in turn, would affect their production of intensity. We further hypothesized that if we were correct in assuming that they would perform less accurately during the perception task, they would utilize intensity to signal focus more than the low AQ group. The reasoning for this prediction is that if they are less accurate in perceiving differences in intensity, they may not understand the need for modulating their voice intensity, especially in certain social settings where speaking loudly may not be widely accepted, but they struggle to grasp this concept as the WCC would predict. Previous literature involving intensity production in autistic individuals report mixed results: aligning with our hypothesis, Olivati et al. (2017) found that ASD individuals displayed significantly higher maximum and minimum intensity than the TD group, however, Diehl et al. (2013) found that when producing question-like speech, there were differences in intensity between ASD and TD groups and Filipe et al. (2014) found no intensity differences between groups when observing falling and rising intonations.

When participants were presented stimuli containing differences in intensity and asked to indicate which word sounded the most focused to them, participants who scored higher on the AQ performed more accurately (M = 55.1%, SE = 1.6%) than participants with low AQ (M = 48.83%, SE = 1.44%), thus a significant effect of AQ in the perception of intensity was found ($z = 2.53$, $p = 0.01$). Additionally, a significant effect of degree of manipulation was found ($z = 2.78$, $p = 0.005$), suggesting that both participant groups are consistently more accurate in perceiving stimuli containing a 6 dB increase (M = 54.35%, SE = 1.52%) than stimuli containing a 3 dB increase (M = 48.89%, SE = 1.52%). Our statistical model also found an interaction between AQ score and degree of manipulation, but only for the full-manipulation condition where stimuli differences were increased by 6 dB ($\beta = -0.44$, SE = 0.13, $z = -3.32$, $p = 0.005$). These findings are much different than what we originally hypothesized. Firstly, participants with higher AQ scores performed more accurately than participants with lower AQ scores. Secondly, with the Enhanced Perceptual Function theory in mind, higher AQ participants should perform more accurately than the low AQ group in the half-manipulation condition where the stimuli were increased by 3 dB as it is below the normal threshold for perception and should therefore be more difficult to perceive than the full-manipulation condition.

Looking at the results from our production task, we did not observe any effects or interactions of AQ, but rather, we observed an effect of both broad focus ($t = 2.454$, $p = 0.015$) and narrow focus ($t = 5.034$, $p = 1.05e-06$), suggesting that both AQ groups produce significantly higher intensity ranges in both broad (M = 5.07, SD = 2.96) and narrow focus (M = 6.24, SD = 2.72) when compared to background focus.

4.3 Duration

Findings from previous perception studies that looked at duration and involved participants with varying levels of autistic traits have returned mixed results. On one hand, Stewart et al. (2018) found that individuals with higher levels of autistic traits displayed enhanced perception of duration, meanwhile Chen & Peng (2021) observed a wider boundary width and lower peakedness scores in participants with ASD, suggesting they were less accurate than the typically developed group at perceiving durational differences. Production studies involving speakers with varying levels of autistic traits observed a smaller clear speech effect characterized by shorter overall word and vowel durations in participants with higher AQ when they were instructed to speak for the benefit of a listener with a hearing impairment (Turnbull 2015; 2019). We hypothesized that our participants scoring higher on the AQ scale would display enhanced perception of duration, regardless of the degree of manipulation, when compared to the low AQ group. Additionally, we hypothesized that participants with higher AQ scores would use duration less when eliciting different focus types, in line with Turnbull's studies. This relates to both cognitive theories presented in chapter one. On one hand, the Enhanced Perceptual Function theory would predict that participants with higher AQ scores would perform more accurately in the perception task than the low AQ group due to their enhanced ability to discriminate fine-grained acoustic information. On the other hand, the Weak Central Coherence theory might predict that because of one's higher level of autistic traits, they might have a harder time grasping that an increase in duration is signalling a change in focus, and therefore, they may use it as an indicator of focus in their production less than a participant with a lower AQ score would.

Beginning with the results from the perception task, no significant effect of AQ was found. However, a significant effect was found for degree of manipulation ($z = 8.46$, $p = <2e\text{-}16$), suggesting that participants from both AQ groups performed significantly better during the full manipulation condition where duration was increased 66% ($M = 75.18\%$, $SE = 1.31\%$), than the half-manipulation condition where stimuli was increased by 33% ($M = 57.31\%$, $SE = 1.51\%$). This finding differs from previous results finding correlations between autistic character traits and perception of f0 and duration, but not intensity (Stewart et al., 2018). It is possible that the results from the present study and Stewart et al.'s study differ due to differences in regional dialects between participants. Although the study by Stewart et al. (2018) utilized the same recruitment criteria as the present study, that all participants must have English as their first language, the 2018 study took place recruited native British English speakers. In Arnhold et al.'s 2020 eye-tracking study, it was observed that native speakers of Canadian English did not use the same prosodic cues as native British English speakers when disambiguating between given and new referents produced by a native British English speaker. Therefore, it is to be expected that speakers of different dialects may use prosody in different ways.

Additionally, the results for the production task found no significant effect of AQ or degree of manipulation, and there were no interactions between the fixed effects. These results differ from Kim & Arnhold's 2022 study, which also recruited speakers of Canadian English, where they found that (1) native speakers of Canadian English produced increased duration and maximum intensity when eliciting broad and narrow focus but (2) produced lower mean intensity in the narrow focus condition than the broad focus condition and (3) made no distinction in f0 range between focus types. Again, this could be due to regional dialectal differences as Kim & Arnhold's study took place in Alberta, so it can be assumed that native speakers of Western Canadian English were recruited. Since the present study took place in the Greater Toronto Area, which is one of the most culturally diverse regions in the world, we cannot guarantee that participants were not influenced by an L2. A follow-up study with a more specific recruitment criteria (e.g., native monolingual speakers of Southern Ontario English) may produce more robust results not only for duration, but across all prosodic parameters. At the same time, however, limiting the recruitment criteria in such a way makes it more difficult to generalize the findings to the current population residing in the GTA.

Another element to consider that may produce more robust results is focus type. The present study looked at narrow focus as our stress element. Roessig et al. (2022) found the difference between broad vs. narrow and narrow vs. contrastive focus types to be more subtle and inconsistent, compared to broad vs. contrastive focus, which returned an increase in all acoustic parameters measured when participants were instructed to answer questions in the context of an interactive, animated game. It would be interesting and beneficial to conduct a follow-up study that analyzes high-level production and perception of prosodic cues such as Roessig et al. (2022). The present results may benefit from conducting a similar follow-up study while comparing different AQ groups and examine the findings with Weak Central Coherence theory in mind as it would predict that although individuals with high AQ can detect fine-grained differences in prosodic cues, they struggle to gather the "bigger picture" meaning from that acoustic information. Therefore, we might expect to see more robust differences when discriminating between acoustic measures and adding another focus type, but high AQ participants might struggle to gather the meaning from these focus types, and thus not implement the same acoustic cues in their production of varying focus types, or if they do, to a smaller degree than those with lower AQ scores.

To summarize, the present study found an enhanced *perception* of the acoustic parameters f0 (pitch) and intensity, but not duration in individuals with higher levels of autistic traits when compared to those with lower AQ scores. In terms of the production of prosodic cues, the present study found no significant effect of AQ across any of the acoustic parameters measured, although the results for f0 are near the 5% significance level for the f0 condition. Abnormalities in the production and perception of prosodic features are amongst some of the most notable characteristic in autistic individuals and can severely hinder one's social development. The findings discussed in this chapter, and how they relate to existing literature contribute to our understanding of how populations with varying levels of autistic traits use prosody to communicate. How these findings may be used to benefit the community will be addressed in the next chapter.

5. Conclusion

The present study sought to determine whether autistic character traits in typically developing individuals are related to the cue-trading of prosodic cues intensity, duration, and f0 for the perception and production of background, broad, and narrow focus. Additionally, the study aimed to find out if the differences observed in the perception and production of prosodic cues existed on a continuum dependent on one's level of autistic traits. On a perceptual level, the data demonstrate correlations between autistic character traits and the ability to perceive differences in f0 and intensity, but not duration. From a production standpoint, the effect of AQ nearly reaches significance during f0 production, suggesting that individuals with higher levels of autistic traits may produce a lower range of f0 differences compared to low AQ participants, and further analysis with more speakers and items would be beneficial in proving this. In other words, high AQ participants show less variation in their production of signalling the different prosodic foci, and thus would signal the concept of focus less strongly than low AQ participants do.

Our results, which differ between each prosodic cue analyzed and from the results of previous studies, show that the correlations that we and others have observed with respect to autistic character traits are not demonstrative of a general perceptual enhancement (Stewart et al. 2018) or a certain production atypicality, but instead, that AQ is specific to certain prosodic cues. In a systematic review by Asghari et al. (2021) that evaluated 39 eligible production studies containing 910 and 850 ASD and TD groups, respectively, they observed that (1) the studies reported variability in the production of prosodic parameters between participant groups and (2) that these findings were dependent on the age group of participants and the type of task used. Specifically, they observed the trend that tasks which required participants to interact with others or complete problem-solving tasks yielded lower statistical significance while tasks that involved general narration yielded statistical significance when comparing participant groups. Additionally, they also observed that studies that showed a statistically significant difference between ASD and TD participants involved adults and adolescents, not children. Looking at results in relation to perception, Bishop, Kuo, and Kim (2020) observed an interaction between accent type and autistic traits, such that an individual’s perceived prominence of higher pitch levels (ranging from L* to L+H*) weakened as AQ score increased. Thus, further research involving various demographics and tasks is needed to better understand these implications; atypical expressions of prosody often impact people’s impressions of the speaker, thus hindering their social development (Kissine et al., 2021).

There may also be some clinical implications associated with the present study. Since vocal productions start as early as within a child’s first year of life, understanding differences in prosody perception and production involving participants at various stages of life with varying levels of autistic traits may assist with early diagnosis of ASD. This study provides a knowledge increase about speech production and perception differences for individuals with a high level of autistic traits, and possibly those clinically diagnosed with ASD, which may serve as an important input to treatment approaches in speech-language pathology and audiology. In speech-language pathology, prosody-based therapies that are more individualized and target each patient’s needs can be developed, should someone decide they want to build on expression and comprehension of prosodic features. Furthermore, audiology may benefit from the results by implementing relevant and dominant segmental and suprasegmental cues into their cochlear implant hardware and software algorithms to better help those with hearing impairments associated with autism.

References

Asghari, S.Z., Farashi, S., Bashirian, S., & Jenabi, E. (2021). Distinctive prosodic features of people with autism spectrum disorder: a systematic review and meta-analysis study. *Scientific Reports, 11*.

Alcántara, J. I., Cope, T. E., Cope, W., & Weisblatt, E. J. (2012). Auditory temporal-envelope processing in high-functioning children with autism spectrum disorder. *Neuropsychologia, 50*(7), 1235-1251.

Arnhold, A., Porretta, V., Chen, A., Verstegen, S. A. J. M., Mok, I., & Järvikivi, J. (2020). (Mis) understanding your native language: Regional accent impedes processing of information status. *Psychonomic Bulletin & Review, 27*(4), 801–808.

Asperger, H. (1944). Die „Autistischen psychopathen" im kindesalter. *Archiv für psychiatrie und nervenkrankheiten, 117(*1), 76-136.

Baltaxe, C.A.M. & Simmons, J.Q. (1985). Prosodic Development in Normal and Autistic Children. In: Schopler, E. & Mesibov, G.B. (Eds.) *Communication Problems in Autism*. (pp. 95-125). Springer.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5-17.

Bates, D., Mächler, M., Bolker, B.M., & Walker, S.C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1).

Bishop, J. (2013). *Prenuclear accentuation: Phonetics, phonology, and information structure* [Unpublished doctoral dissertation]. University of California, Los Angeles.

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*.

Bishop, J., Zhou, C., Antolovic, K., Grebe, L., Hwang, K.H., Imaezue, G., et al. (2021). Brief report: Autistic traits predict spectral correlates of vowel intelligibility for female speakers. *Journal of Autism and Developmental Disorders.*

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. Acoustic correlates of information structure. *Language and Cognitive Processes, 25*, 1044-1098.

Chen, F. & Peng, G. (2021). Categorical perception of pitch contours and voice onset time in Mandarin-speaking adolescents with autism spectrum disorders. *Journal of Speech, Language, and Hearing Research, 64*, 4468-4484.

Chodroff, E., & Cole, J. (2018). Information structure, affect, and prenuclear prominence in American English. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 2018-September*, 1848-1852.

Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience, 30*(1-2), 1-31.

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech, 29*(3).

Cutler, A. & Jesse, A. (2021). Word Stress in Speech Perception. In Pardo, J.S., Nygaard, L.C., Remez, R.E. & Pisoni, D.B. (Eds.), *The handbook of speech Perceptio*n (2nd ed). (pp. 239-265). John Wiley & Sons, Inc.

Cutler, A. (2005). Lexical Stress. In Pisoni, D.B. & Remez, R.E. (Eds.), *The handbook of speech perceptio*n (pp. 264-289). Blackwell Publishing.

Diehl, J. J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics, 34*(1), 135-161.

Diehl, J.J., Bennetto, L., Watson, D., Gunlogson, C., & McDonough, J. (2008). Resolving ambiguity: a psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain and Language, 106*(2), 144–152.

Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: Perceptual and acoustic measurements. *Journal of autism and developmental disorders*, 44, 1972-1981.

Flaherty, M., Dent, M.L., & Sawusch, J.R. (2017). Experience with speech sounds is not necessary for cue trading by budgerigars (Melopsittacus undulatus). *PLoS One, 12*(5).

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthelemy, C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience, 7*.

Grossman, R.B., Bemis, R.H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech, Language and Hearing Research, 53*(3), 778-93.

Happe, F.G.E. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15*, 1-12.

Jolliffe, T. & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition, 71*(2), 149-185.

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

Kargas, N., López, B., Reddy, V., & Morris, P. (2015). The relationship between auditory processing and restricted, repetitive behaviors in adults with autism spectrum disorders. *Journal of autism and developmental disorders, 45*, 658-668.

Kim, D., & Clayards, M. (2019). Individual differences in the link between perception and production and the mechanisms of phonetic imitation. *Language, Cognition and Neuroscience, 34*(6), 769–786.

Kim, Jiseung & Anja Arnhold (2022). Prosodic focus marking in Canadian English. In *Proc. Speech Prosody 2022* (pp. 97-100).

Kingston, J., Rich, S., Shen, A., & Sered, S. (2015). Is perception personal? *International Congress of Phonetic Sciences.*

Kissine, M., Geelhand, P., Philippart De Foy, M., Harmengnies, B., & Deliens, G. (2021). Phonetic inflexibility in autistic adults. *Autism Research, 14*(6), 1186-1196.

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.

Lenth, R. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.7, https://CRAN.R-project.org/package=emmeans

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive development disorders. *Journal of Autism and Developmental Disorders, 24*(5).

Matzinger, T., Ritt, N., & Tecumseh Fitch, W. (2021). The Influence of Different Prosodic Cues on Word Segmentation. *Psychology of Language, 12*.

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*. 27-43.

Nieuwland, M.S., Ditman, T., & Kuperberg, G.R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324-346.

O'Connor, K. (2012). Auditory processing in autism spectrum disorder: a review. *Neuroscience & Biobehavioral Reviews, 36*(2), 836-854.

Olivati, A. G., Assumpção, F. B., & Misquiatti, A. R. N. (2017, April). Acoustic analysis of speech intonation pattern of individuals with Autism Spectrum Disorders. In *CoDAS* (Vol. 29). Sociedade Brasileira de Fonoaudiologia.

Patel, R., Niziolek, C., Reilly, K., & Guenther, F.H. (2011). Prosodic Adaptations to Pitch Perturbation in Running Speech. *Journal of Speech, Language, and Hearing Research, 54*(4), 1051-1059.

Paul, R., Augustyn, A., Klin, A. & Volkmar, F. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of autism and developmental disorders, 35*, 205-220.

R Development Core Team. (2022) R: a language and environment for statistical computing. In *R: A Language and Environmental for Estatistical Computing*. https://www.r-project.org/

Roessig, S., Winter, B., Mucke, D. (2022). Tracing the phonetic space of prosodic focus marking. *Frontiers in Artificial Intelligence, 5*.

Rosenhall, U., Nordin, V., Sandström, M., Ahlsen, G., & Gillberg, C. (1999). Autism and hearing loss. *Journal of autism and developmental disorders, 29*, 349-357.

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6*, 2.

Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics, 52*, 183-204.

Stewart, M.E., Griffiths, T.D., & Grube, M. (2018). Autistic traits and enhanced perceptual representation of pitch and time. *Journal of Autism and Developmental Disorders, 48*, 1350-1358.

Stewart, M.E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition, 109*, 157-162.

Turnbull, R. (2015). Patterns of individual differences in reduction: Implications for listener-oriented theories. In ICPhS.

Turnbull, R. (2019). Listener-oriented phonetic reduction and theory of mind. *Language, Cognition and Neuroscience, 34*(6), 747-768.

Tyler, M.D. & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America, 126*, 367-376.

Tyler, R. S., Pienkowski, M., Roncancio, E. R., Jun, H. J., Brozoski, T., Dauman, N., ... & Moore, B. C. (2014). A review of hyperacusis and future directions: part I. Definitions and manifestations. *American journal of audiology, 23*(4), 402-419

Wang, A., Dapretto, M., Hariri, A., Sigman, M. & Brookheimer, S. (2001). Processing affective and linguistic prosody in autism: An fMRI study. *Neuroimage, 13*(6).

Wing, L. (1981). Asperger's syndrome: A clinical account. *Psychological Medicine, 11*(1), 115-129.

Yu, A.C.L. (2010). Perceptual compensation is correlated with individuals' "autistic" traits: Implications for models of sound change. *PLoS ONE, 5*(8).

Yu, A.C.L. (2022). Perceptual Cue Weighting Is Influenced by the Listener's Gender and Subjective Evaluations of the Speaker: The Case of English Stop Voicing. *Frontiers in Psychology, 13*.

Yu, A.C.L. (2016). Vowel-dependent variation in Cantonese /s/ from an individual perspective. *The Journal of the Acoustical Society of America, 139*.

Yu, A. C., Grove, J., Martinovic, M., & Sonderegger, M. (2011, August). Effects of working memory capacity and "autistic" traits on phonotactic effects in speech perception. In *Proceedings of the international congress of the phonetic sciences xvii, Hong Kong: International congress of the phonetic sciences* (pp. 2236-2239).

Yu, A.C.L. & To, C.K.S. (2020). Atypical context-dependent speech processing in autism. *Applied Psycholinguistics, 41*, 1045-1059.

Yu, M., Sommers, B., Yin, Y., & Yan, G. (2019). Effects of implicit prosody and semantic bias on the resolution of ambiguous Chinese phrases. *Frontiers in Psychology, 10*.