The influence of sentence context on the categorical perception of English stop VOT in individuals with a high Autism Spectrum Quotient (AQ)

Gloria Yu
Candidate number: KFYW8
University College London

PLIN0025 Long Essay/Project (20/21)
Supervisor: Dr James White
Submitted on 21st April, 2021
Word count: 6265

Abstract

Autism Spectrum Disorder (ASD) is associated with a cognitive processing style where low-level stimuli are well-processed, sometimes at the expense of high-level processing (Happé & Frith, 2006; Meilleur *et al.*, 2015; Pellicano & Burr, 2012). The present study investigates this cognitive theory in speech perception. The study is a forced-choice word-identification task of three minimal pairs in English: *beer/pier, tie/dye, cage/gauge.* The Voice Onset Time (VOT; low-level stimuli) of the word-initial stops was manipulated to vary along 9-step continua (10-70ms). The words were heard in three sentence contexts (high-level stimuli): neutral, voiced bias, and voiceless bias. The study compares individuals with a high Autism Spectrum Quotient (AQ; Baron-Cohen *et al.,* 2001) to those with low- and moderate-AQ. It replicates the findings of Miller *et al.*, (1984) that sentence context shifts the VOT of the category boundary, for all participants. It did not find a reduced effect of sentence context nor reduced categorical perception in the high-AQ group.

1 Introduction

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition (APA, 2013); it is associated with a style of cognitive processing where "low-level" details of stimuli's physical properties are well-processed, but where processing "high-level" mental representations is challenging (Stewart *et al.,* 2018).

In typically developing (TD) individuals, some aspects of speech perception, like voicing categories in stops, are categorical (Lisker & Abramson, 1964). These categories are influenced by listeners' prior knowledge of native phonemic categories (Kuhl *et al.,* 2008) and sentence context (Cole & Jakimik, 1978; Marslen-Wilson 1975); both of which are considered higher-level mental representations. Given that autistics typically struggle with higher-level processing, it is possible to ask: (i) is the effect of sentence context on phonemic categorisation significantly reduced in autistic individuals?; (ii) is phonemic perception less categorical?

One reason speech perception is of interest for ASD is that the condition currently has no comprehensive language profile. Although communication challenges are part of its diagnostic criteria, only pragmatic language deficits have been identified as a unifying factor in ASD (Lord & Paul, 1997). Structural language development — phonology, morphology, syntax, semantics — is extremely heterogeneous in ASD (Boucher, 2012), with the most variability in development before age six (Pickles, Anderson & Lord, 2014).

This heterogeneity is reflected in studies on receptive phonetics and phonology in ASD. On one hand, some studies found reduced phonetic and phonological awareness. For example, Key, Yoder, and Stone's (2016) perception task found reduced consonant differentiation in the ASD group compared to TD controls. Similarly, Matsuzaki *et al.*'s (2019) oddball paradigm study used vowel stimuli to find delayed auditory discrimination in autistic individuals, which correlated with their overall language skills.

On the other hand, some studies have found (receptive) phonetics and phonology to be relative strengths (Tager-Flusberg, 2000; Loucas *et al.*, 2008). Belmonte *et al.* (2013) identified a subgroup of autistic individuals with receptive skills that were more well-developed than their expressive skills. Several studies have even suggested that speech perception difficulties in autism are uncommon and show great variability (Boucher, 2012; Haesen, Boets & Wagemans, 2011; Kujala, Lepistö & Näätänen, 2013).

Secondly, speech perception is worth investigating because phonetics and phonology may be particularly predictive of autistics' language skills overall. Saul and Norbury (2020) evaluated Yoder, Watson, and Lambert's (2015) findings that "parental responsiveness, child response to joint attention, child communicative intent, and consonant inventory" were strong predictors of language development in autistic children. Of these, Saul and Norbury found only *consonant inventory* to significantly correlate with expressive language outcome. They found that replacing *consonant inventory* with the broader *phonetic repertoire* produced an increased model fit for predicting expressive language growth. They concluded the ability to map sounds to letters could be a “protective” factor — i.e. it could be related to less severe language deficits overall. Surprisingly, these findings could suggest that phonological skills are more predictive of overall language skills in ASD than the pragmatic factors identified by Yoder *et al*. This is intriguing given that pragmatic deficits are the unifying language trait in ASD.

The present study investigates the categorical perception of English stop consonants in individuals with a high Autism Spectrum Quotient (AQ; Baron-Cohen *et al.*, 2001). The study finds that high-AQ individuals show equal influence of sentence context and equal categorical perception compared to the low- and moderate-AQ groups.

1.1 Autism Spectrum Quotient (AQ)

According to recent studies, traits associated with ASD are not limited to clinically diagnosed autistics (Constantino & Todd, 2003; Folstein & Rosen-Sheidley, 2001; Lundström *et al.*, 2012). In light of this, some researchers have investigated autistic-like traits for documenting individual differences in phonological processing (e.g. Huang, 2007; Stewart & Ota, 2008; Turnbull, 2015; Yu, 2010). They did so using the Autism Spectrum Quotient (AQ; Baron-Cohen *et al.,* 2001): a short, self-administered test for measuring the degree of autistic-like traits expressed in adults with average IQ. It is comprised of 50 items, with 10 questions on five measures: “social skills, communication, attention to detail, attention switching, and imagination.”

Though not a diagnostic measure, the AQ has been used clinically for screening purposes; traits measured by the AQ show strong heritability and cross-cultural stability (Yu & Zellou, 2019). Woodbury-Smith *et al.* (2005) find a cut-off score of 26 as having 95% accuracy for identifying autistic (>26) and non-autistic (<26) individuals. For these reasons, the AQ was used in place of a formal diagnostic verification of autism (e.g. ADOS), which was not possible due to the limited scope of this study.

1.2 Modes and levels of processing

Atypical cognitive processing has been comprehensively shown in autistic individuals. Studies have shown enhanced perception in both visual and (Baron-Cohen & Hammer, 1997; Plaisted *et al.* 1999) auditory domains (Bonnel *et al.*, 2003; Jones *et al.*, 2009), including speech (Heaton, 2005; Heaton *et al.*, 2008). However, higher-order tasks in the same domains, such as understanding intonation, or interpreting emotion from faces and speech, is typically challenging (Ashwin *et al.*, 2007; Boraston *et a*l., 2007; Kujala *et al.*, 2005; McCann & Peppé 2003).

These contradictory findings have been understood as the result of how autistics coordinate the levels of processing (Kern *et al.*, 2006): low levels, which deal with stimuli's physical properties, and high levels, which link percepts to relevant mental representations, such as sentence context or phonemic categories (Stewart *et al.,* 2018). Generally, theoretical accounts of atypical processing in ASD assume good low-level processing, possibly due to reduced or disrupted high-level processing. The following is a summary of those accounts:

1. **Weak Central Coherence:** low and high levels of processing are weakly integrated (Frith, 2003; Happé, 2005). Local stimuli details are well-processed, sometimes to the detriment of contextualised meaning. Conversely, individuals with typical central coherence might prioritise contextualised meaning, sometimes to the detriment of memory for detail (Happé & Frith, 2006).
2. **Enhanced Perceptual Functioning:** increased low-level functioning (Mottron & Burack, 2001; Mottron *et al.*, 2006; Meilleur *et al.,* 2015). Prior knowledge is only used in perceptual tasks when it facilitates performance.
3. **Bayesian accounts:** a greater reliance on bottom-up processing due to inflexible updating of prior knowledge, a reduced role of prior knowledge, or overly precise bottom-up processing (Brock, 2012; Lawson *et al.*, 2014; Pellicano & Burr, 2012; Van de Cruys *et al.,* 2014).
4. **Reduced synaptic pruning:** In speech processing, synaptic pruning allows infants to specialise their perception to native phonemic contrasts (Gopnik *et al.*, 1999) by removing unnecessary neuronal structures (Chechik *et al.*, 1999). Evidence for reduced synaptic pruning is found in accelerated brain growth in autistic children (6-24 months), with growth decelerating earlier than in neurotypicals (Courchesne, 2004). Without synaptic pruning, autistic individuals may retain the ability to distinguish non-native contrasts (within-category differences) into adulthood (DePape *et al.*, 2012; Happé & Frith, 2006).

In this paper, we investigate high-level processing in terms of phonemic categories and sentence context, as well as low-level processing of consonants' Voice Onset Time (VOT). If these theories are correct, we might expect speech perception in high-AQ individuals to show atypical cognition through: (i) a reduced influence of sentence context; (ii) a reduced effect of prior phonemic knowledge — i.e. less categorical perception.

1.3 Voice Onset Time (VOT) and Categorical Perception

In this study, we investigate whether high-AQ individuals show a reduced effect of higher-level representations (sentence context) on the perception and categorisation of low-level stimuli (gradient phonetic information). In neurotypicals, a well-studied phonetic cue in speech perception is Voice Onset Time (VOT): the period between a stop's release and the onset of voicing (Lin & Wang, 2001; Lisker & Abramson, 1964).

Although it varies gradiently, listeners can categorise speech sounds based on VOT. For example, English /d/ and /t/ share the same place and manner of articulation, but are primarily distinguished by their place on the VOT continuum (Fernandez & Cairns, 2010). Other gradient acoustic cues determine voicing, (White *et al.*, 2006; Constantino *et al.,* 2007; Shultz *et al.,* 2012; Stewart *et al.*, 2018; You *et al.*, 2017), but VOT has produced the clearest categorisation in TD adults (Kuhl *et al*., 2008).

The categorisation of stimuli varying continuously along a physical dimension is called *categorical perception*. In speech perception, listeners will show a step-like category boundary: here, listeners might categorise percepts along a VOT continuum as voiced — almost 100% of the time — up to a certain VOT, after which they will categorise the segments as *voiceless* almost 100% of the time (Lisker & Abramson, 1964). This can be visualised as an S-curve (Figure 1).

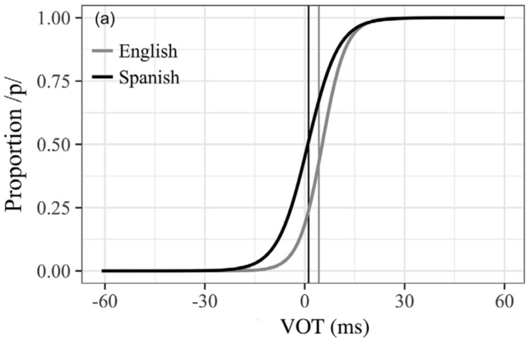

Figure 1. *Example of VOT categorical perception (from Casillas & Simonet, 2018; p.57)*

The steep change in listeners' identification of percepts is the *category boundary*: the VOT at which listeners categorise the stimuli at chance level (here, 50% voiced/voiceless). In English, the VOT boundary for stops is roughly 13-16ms for labials, 24-28ms for coronals, and 27-31ms for dorsals (Nakai & Scobbie, 2016); voiced stops are typically short-lag VOT (they have do not have negative VOT). In this study, we explore whether these category boundaries shift depending on sentence context, and how much.

Categorical perception is robustly demonstrated in neurotypicals (Liberman *et al.*, 1957, 1967). For instance, English speakers struggle to distinguish non-native phonemic contrasts (Francis & Nusbaum, 2002; Werker & Tees, 1994). Such findings are thought to reflect internal phonological categories in listeners' linguistic system (Kuhl *et al.,* 2008); this prior knowledge facilitates perception through top-down, predictive processing of the phonological categories matching the auditory input (Kuhl *et al.,* 2008). Hence, the categorical perception paradigm enables simultaneous investigation of bottom-up *and* top-down processing. Given that categorical speech perception is well-established in neurotypicals, *reduced* categorical perception would provide strong evidence for atypical cognitive processing in ASD. Some evidence of this has been observed in visual perception (Soulières *et al.*, 2007); if this reflects reduced categorical perception in ASD more broadly, we expect similar results in speech perception tasks.

As of writing this paper, findings on speech perception in ASD are mixed. Although some have found increased sensitivity to acoustic differences (DePape *et al*., 2012; Happé & Frith, 2006), few have focused specifically on categorical perception (You *et al*., 2017; Constantino *et al*., 2007; White *et al*., 2006; Stewart *et al*., 2018). Of these, only You *et al*.'s identification-task study found prior phonemic knowledge to be less precise; the others found that discrimination of both between-category and within-category differences in ASD fell within the normal range. However, after indexing the difference between perception of end-point versus boundary phonemes, Stewart *et al*. (2018) found the autistic group showed a smaller categorical-discrimination index than TD controls. Although this did not reach statistical significance, they found a trend and moderate effect size. Moreover, the ASD group showed no noticeable peak in the discrimination curve at the category boundary. They concluded that their findings aligned with DePape *et al.* (2012) and Haesen *et al.* (2011), who suggest autistic individuals prioritise low-level perceptual details.

More research is needed on speech perception in ASD, because reduced categorical perception could explain some language-related difficulties reported in autistics (Kjelgaard & Tager-Flusberg, 2001). More gradient perception could lead to mishearing minimal pairs close to the boundary (e.g. *beer/pier*), or being distracted by acoustic within-category differences. Reduced categorical perception provides a testable connection between auditory perception and language skills in ASD, which so far is not fully established (cf. Jones *et al*., 2009; Jarvinen-Pasley *et al.*, 2008).

1.4 Sentence Context

Besides reduced categorical perception, atypical cognition in ASD could be demonstrated by a reduced effect of sentence context. This is because cognitive theories of autism explicitly predict that context is poorly processed (Happé & Frith, 2006). In this study, we use sentence context to investigate the effect of top-down processing, as sentence comprehension requires higher-order representations of complete ideas, as well as lexical knowledge of multiple words (Angosto, 2013).

In many models of speech comprehension, sentence context is thought to directly influence auditory word recognition (Cole & Jakimik, 1978; Connine, 1987; Jesse, 2019; Marslen-Wilson, 1975). This is well-established in neurotypicals. Studies have found listeners are more likely to categorise steps in phonetic continua to match semantic context (Abada *et al.,* 2008; Borsky *et al.,* 1998; Miller *et al.*, 1984; Schertz & Hawthorne, 2018). For example, Miller *et al.* (1984) tested a VOT continuum (*bath/path*) in two types of sentences, biased towards either the voiced or voiceless option. They found listeners tended to perceive ambiguous tokens as the contextually appropriate option, shifting the category boundary to fit the context.

Sentence context can also facilitate word-identification. Studies show that sentence context increases the accuracy of word-identification compared to words in isolation (Miller *et al.,* 1951; Pollack & Pickett, 1964). Additionally, less acoustic information is necessary for word-recognition when sentence context is available (Salasoo & Pijoni, 1985). Based on the previous studies, we expect low-AQ individuals to rely more on context than acoustic details. This could be shown through a shift in category boundaries depending on context: compared to a neutral sentence, the category boundary might have a greater VOT when the context favours the voiced sound, but have a shorter VOT when it favours the voiceless. If high-level processing is reduced in ASD, high-AQ individuals may show less sensitivity to context.

Importantly, higher-order prior knowledge must be available early enough before the target to guide phonetic retuning (i.e. shifting category boundaries). This is clear from studies which did not include sentence context, but focused on lexical knowledge (Clarke-Davidson *et al.,* 2008; Drouin *et al.,* 2016; Eisner & McQueen, 2005; van der Zande *et al.,* 2013). For example, when using longer words, Jesse and McQueen (2011) did not observe phonetic retuning in the onsets, as the words only became lexically unique later. They did, however, observe phonetic retuning when the same ambiguous segment was spliced into the word-final position. In this study, we replicate studies like Miller *et al.* (1984), where context is given through semantically related content words, appearing separately from and before the target. Thus, we avoid the issue of context appearing insufficiently early.

To my knowledge, no studies have investigated the effect of sentential context on categorical perception in ASD. Given that sentence comprehension is a form of higher-level processing, investigating both sentence context and categorical perception offers a clear contrast between individuals' potentially enhanced low-level processing and their high-level processing. It could also offer a more direct link between speech perception and autistics' more well-documented pragmatic deficits, which heavily involve context (APA, 2013).

1.5 Aims

This study aimed to explore the effects of sentence context on the categorisation of the VOT of English stop consonants. Based on the previous studies, it was predicted that:

1. All groups will shift the category boundary depending on sentence context, but
2. Compared to the other groups, high-AQ participants will have category boundaries in the biased contexts whose VOTs differ less from the neutral context.
3. High-AQ individuals may show more gradient perception at the category boundaries.

The VOT was manipulated of word-initial consonants in voiced/voiceless minimal pairs (e.g. *tie/dye*). Sentence context could be neutral or biased towards the voiced or voiceless option (Table 1), similar to Miller *et al.* (1984) and Schertz and Hawthorne (2018).

Table 1. *Sentences used*

	cage/gauge	**pier/beer**	**tie/dye**
Neutral	The bald janitor wiped the { cage | gauge }.	The interesting tourist enjoyed the { pier | beer }.	The annoying kid ruined the { tie | dye }.
Voiced bias	The mindful engineer adjusted the { *cage | gauge }.	The tired worker drank the { *pier | beer }.	The creative barber mixed the { *tie | dye }.
Voiceless bias	The sleazy ringmaster rattled the { cage | *gauge }.	The grouchy sailor scrubbed the { pier | *beer }.	The sleepy businessman loosened the { tie | *dye }.

2 Methods

2.1 Participants

The experiment recruited a total of 76 participants (age: 18-61; mean: 33.9; breakdown given in Table 2). Following the cutoff scores from Woodbury-Smith *et al.* (2005), the low-AQ group scored 0-25 out of 50, the moderate-AQ group 26-31, and the high-AQ group 32-50. The AQ groups were determined *post hoc* based on the AQ test participants completed. Participants were recruited on Amazon Mechanical Turk. They received a small amount of monetary compensation. Participants were excluded for not being native speakers of British English, which was determined via a demographic questionnaire (n=22).

Table 2. *Participant breakdown*

AQ	**Sex**			**Total**
	F	M	Other	
High AQ (32-50)	1	7	3	**11**
Moderate AQ (26-31)	3	8	1	**12**
Low AQ (0-25)	16	28	9	**53**
Total	**20**	**43**	**13**	**76**

2.2 Materials

Materials were adapted from Schertz and Hawthorne (2018). The stimuli consisted of sentences ending in target words beginning with stops drawn from a continuum from voiced to voiceless.

There were three continua, drawn from minimal pairs: *beer/pier, tie/dye,* and *cage/gauge*. All minimal pairs had an average Zipf score of 4.05 (range: 3.49-4.72 or 1-100fpmw) (SUBTLEX; Van Heuven *et al.*, 2014). They corresponded to the three places of articulation for stops in English, as using only one may increase the salience of within-category differences. Following Sullivan (2019), only one minimal pair was used per place of articulation. This shortened the experiment so all participants could hear the same stimuli; any differences between AQ groups in the influence of sentence context will be due to AQ and not the specific words or sentences.

A norming study was conducted to ensure that the target words were semantically appropriate given the sentence context. This was done on Amazon Mechanical Turk with 27 participants (F: 5, M: 17; N/A: 5; age: 19-65; mean: 34.7). The study tested 12 minimal-pairs, each in three sentence contexts, per place of articulation (15 pairs x 3 PoA = 45). Participants were shown the incomplete sentences (e.g. *The creative barber mixed the* _____) and rated the two options (e.g. *dye/tie*) from 1-7 based on how well they fit the sentence: 7 meant "very well" and 1, "not well at all." The mean ratings for the chosen sentence/minimal-pair sets are given in Table 3; a negative difference in mean ratings indicates a preference for the voiceless option. The sets were chosen based on the following criteria: (i) the neutral context was judged to not favor either word (i.e. the difference in mean rating between the words was <1); (ii) the voiced- and voiceless-initial word was judged to be an excellent fit in their respective contexts — i.e. it had a mean rating of >6.5 out of 7 *and* there was a clear difference in acceptability of the two words in biased contexts (difference in mean rating >3).

A 9-step VOT continuum was created for each minimal pair (10-70ms; per Winn, 2020). VOT was measured from 0ms before the stop burst to the zero crossing before the first clear cycle of the subsequent vowel. In total, there were 27 word tokens (3 continua x 9 steps).

Table 3. *Mean ratings of words chosen*

Word	Sentence	Voiced	Voiceless	Difference
cage	Neutral	4.96	4.63	0.33
	Voiced bias	6.67	3.58	3.08
	Voiceless bias	2.11	6.75	-4.64
pier	Neutral	6.45	5.80	0.65
	Voiced bias	6.52	1.17	5.35
	Voiceless bias	3.21	6.54	-3.33
tie	Neutral	4.33	4.96	-0.63
	Voiced bias	6.88	2.13	4.75
	Voiceless bias	2.14	7.00	-4.86

The continua were created using Winn's (2020) Praat script. Recordings of both the voiced- and voiceless-initial words were used as the base stimuli. These were taken from recordings of the complete sentences, to ensure the stimuli's overall pitch contour was as natural as possible. The script followed the "progressive cutback and replacement" method: the voiced stop's onset was progressively removed and replaced with the equivalent amount of the voiceless onset (cf. McMurray *et al.,* 2002, 2003). This was done because adding aspiration would be more unnatural, because the aperiodic segments of voiceless stops arise from devoicing of the vowel onset: they are not pre-appended (Winn, 2020). The stop was then spliced with the recording of the voiced-initial word, starting from the onset of voicing. F0 manipulation was set to 1 step, with the minimum and maximum F0 onset set to 0 and 20 Hz, respectively.

Sentences could be neutral or biased, either towards the voiced- or voiceless-initial word (Table 1, repeated below). Crucially, the initial stop of the target word did not occur anywhere else in the sentence in onset position. This ensured no unambiguous tokens could be used for comparison within a trial. Sentences were recorded in full, but were cut in Praat (Weenik & Boersma, 2008) to exclude the final word. The edited recordings were concatenated with the tokens from the continua. In total, there were 81 items (3 contexts x 3 continua x 9 steps). All were normalised for intensity with a Praat script.

Table 1. *Sentences used*

	cage/gauge	**pier/beer**	**tie/dye**
Neutral	The bald janitor wiped the { cage | gauge }.	The interesting tourist enjoyed the { pier | beer }.	The annoying kid ruined the { tie | dye }.
Voiced bias	The mindful engineer adjusted the { *cage | gauge }.	The tired worker drank the { *pier | beer }.	The creative barber mixed the { *tie | dye }.
Voiceless bias	The sleazy ringmaster rattled the { cage | *gauge }.	The grouchy sailor scrubbed the { pier | *beer }.	The sleepy businessman loosened the { tie | *dye }.

2.3 Procedure

Participants completed a self-paced, forced choice word-identification task. They were instructed to use headphones or earbuds to listen to the stimuli. They were told to select the word based on what they heard the speaker say, regardless of which word fits best in the sentence.

For each trial, participants saw the sentence written without the final word (e.g. *The creative barber mixed the* _____). Below it was a button which, when pressed, played an audio of the complete (manipulated) sentence. The sentence could only be played once. After it was played, participants saw the question, “What word did you hear?” on the screen. Below it were two buttons, each with a word to complete the sentence. Once they had clicked on a word, the next trial began on a new slide. Participants could not go back to change their answers.

In total, there were 81 trials (3 contexts x 3 continua x 9 steps). The order was randomised for each participant, while following the restriction that participants would not hear targets from the same continuum twice in a row.

At the end of the experiment, participants completed the AQ self-assessment. Each item was shown individually on one page, in the order given in Baron-Cohen *et al.*'s (2001) original design. The items are presented as statements (e.g. "I prefer to do things with others rather than on my own"). Below the statements were four buttons: "definitely disagree," "slightly disagree," "slightly agree," and "definitely agree." After clicking on an answer, the next item appeared on a new slide. Participants could not go back to change their answers.

Finally, participants completed a survey about their demographics, language background, and personal or family history of autism. The web experiment was run on Experigen (Becker & Levine, 2013).

3 Results

Participants were excluded for not completing the task (n=14) or not showing an S-curve (n=11). Given that the S-curve is a key indicator of categorical perception (see Section 1.3), a lack of S-curve indicates participants were not paying attention or completed the task incorrectly. For instance, participants may have selected the semantically appropriate option every time regardless of VOT, or selected the words at random.

This section presents the results with respect to the predictions in Section 1.5: (i) the effect of context and place of articulation on the category boundaries across all AQ groups; (ii) whether AQ influences how far the boundary shifts in different contexts; (iii) whether AQ correlates with more or less categorical perception.

3.1 Sentence Context and Place of Articulation

As expected based on previous studies (e.g. Miller *et al.*, 1984; Schertz & Hawthorne, 2018), each place of articulation had a different category boundary, which shifted depending on sentence context. Figure 2 (overleaf) shows the proportion of voiced responses by place of articulation, across all contexts and AQ groups. The points are the aggregate means for each of the nine steps in the experiment. The points were fitted to the best-fitting logistic function curve (shown as dashed curves) using the *curve_fit* function from the SciPy Optimisation module (Virtanen *et al.*, 2020). The category boundaries were computed by calculating the derivative of the curve at every point; the maximum derivative was the steepest point, which was also the category boundary (i.e. the point where the curve crossed 0.5 on the *y*-axis), shown as dashed vertical lines. Unexpectedly, the category boundary for labials (blue) had the longest VOT (41ms), while the boundary for dorsals had the shortest (orange; 34ms); typically, the reverse is found (Nakai & Scobbie, 2016).

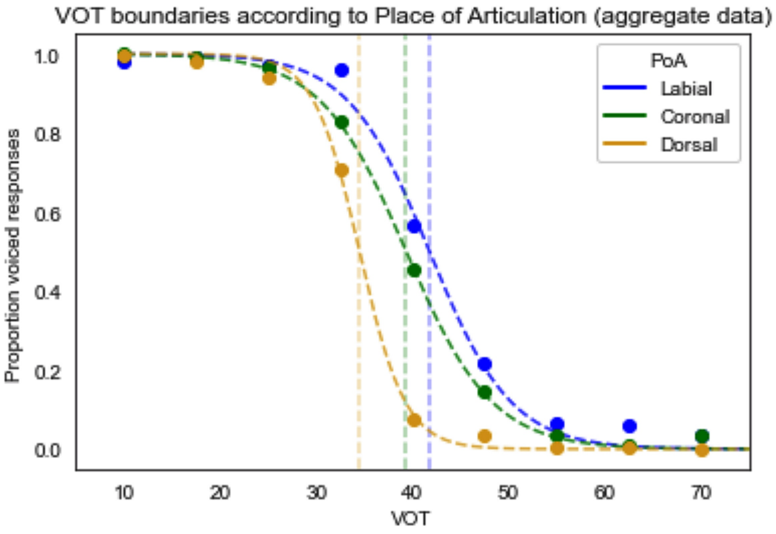

Figure 2. *Voiced response by place of articulation*

The category boundaries also clearly shifted depending on sentence context. Figure 3 (overleaf) shows the proportion of voiced responses in the three contexts, across the three places of articulation. In the neutral context (purple), the boundary is roughly 39ms. In the voiced bias context (red), the boundary shifted to a longer VOT of roughly 41ms. In the voiceless bias context (blue), the boundary had a shorter VOT of roughly 34ms.

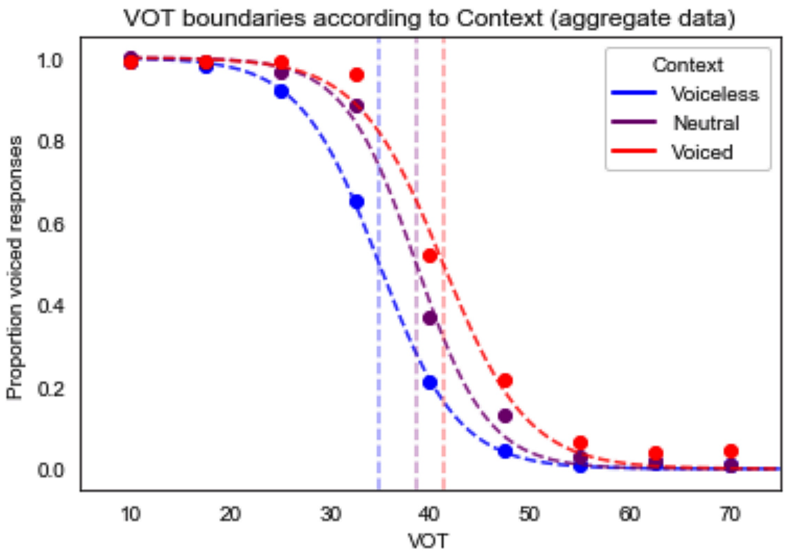

Figure 3. *Voiced response by sentence context*

Statistical analysis was done using mixed-effects logistic regression models (Jaeger, 2008) implemented in R (R Core Team, 2016) with the *lme4* package (Bates *et al.*, 2015). The models did not have a maximal random effects structure; they had random intercepts for participants, but a more complex random effects structure was not possible due to convergence issues. Only steps 3-7 were considered in the statistical analysis as this was the expected ambiguous range; the lowest and highest two steps were expected to be unambiguous.

The initial model predicted the odds of a voiced response and had fixed effects for Step[1] (3-7), Context (Neutral *vs* Voiced, or Voiceless), and Place of Articulation (Coronal *vs* Dorsal or Labial), with an interaction between context and place of articulation. It was then compared to a model without the interaction between context and place of articulation using a likelihood ratio test (with the *anova()* function; Baayen, 2008). The interaction did not significantly improve model fit (at an alpha level .05), so it was removed from the model.

The final model's fixed effects are summarised in Table 4 (overleaf). The baseline (intercept) condition is the Neutral context and Coronal place of articulation. The table reports the *p*-value based on the Wald *z* from the model summary*.* Since the model predicts a voiced response, a positive Estimate means 'more likely to be a voiced response'. A negative Estimate means 'less likely to be a voiced response.'

---

[1] Since the steps fell at regular intervals along the VOT scale, it does not matter whether the model considers step number or the actual VOT value — the outcome would be the same.

Table 4. *Summary of the fixed effects for the final model*

Predictor	Estimate	Standard error	Wald *z*	p-Value (Wald *z*)
Intercept	11.2	0.474	23.7	< .001
Step	-2.25	0.089	-25.5	< .001
Context				
Type = *Voiced bias*	0.862	0.162	5.31	< .001
Type = *Voiceless bias*	-1.24	0.166	-7.44	< .001
Place of Articulation				
Type = *Dorsal*	-1.53	0.169	-9.03	< .001
Type = *Labial*	0.811	0.162	5.00	< .001

Unsurprisingly, there was a significant main effect of Step, indicating that, as VOT increased, participants were increasingly unlikely to identify the word as voiced-initial. Secondly, there was a significant main effect of both the voiced and voiceless bias contexts: both shifted relative to the neutral context. The main effect of Place of Articulation was also significant. However, given that there was no significant interaction between Context and Place of Articulation, the results suggest that context shifted the category boundary by roughly the same amount for all places of articulation.

3.2 AQ and Sentence Context

All AQ groups shifted their category boundaries depending on context, and had very similar results (Figure 4, overleaf). In the low-AQ group (left plot), the category boundary was around 49ms in the neutral context (purple). In the voiceless bias context (blue), it shifted to have a shorter VOT of around 35ms; in the voiced bias context (red), around 42ms. In the mid-AQ group (middle plot), the category boundary shifted from 39ms in the neutral context (purple) to around 34ms and 41ms in the voiceless (blue) and voiced (red) bias contexts, respectively. In the high-AQ group (right plot), the boundary VOTs were around 33ms, 38ms, and 40ms in the voiceless bias (blue), neutral (purple), and voiced bias contexts (red), respectively.

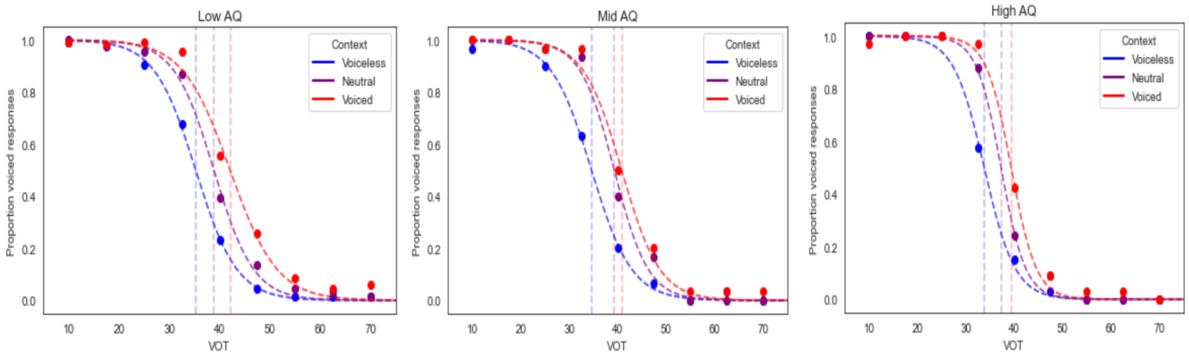

Figure 4. *Voiced response by sentence context Voiced response by sentence context in the all AQ groups. Left to right: low, mid, high AQ.*

Statistical analysis of AQ’s effect was done with a plain logistic regression model, rather than a mixed-effects model, due to convergence issues. The model is summarised in Table 5 (overleaf). The model predicted the odds of a voiced response and had the predictor variables of Step, AQ group (High *vs* Low or Mid), Context (Neutral *vs* Voiced, or Voiceless), and AQ group by Context interaction. Although there was a significant effect of Step, there was no significant effect of AQ Group, and no significant interaction between AQ Group and Context.

Table 5. *Summary of the plain logistic regression model for AQ*

Predictor	Estimate	Standard error	*t*-value	p-Value (> | *t* |)
Intercept	1.72	0.035	49.3	< .001
Step	-0.255	0.005	-55.6	< .001
AQ group				
Low	0.038	0.030	1.28	0.201
Mid	0.051	0.038	1.34	0.180
Context				
Voiced	0.061	0.037	1.64	0.102
Voiceless	-0.091	0.037	-2.45	0.014
AQ Group : Context				
Low:Voiced	0.028	0.042	0.674	0.500
Mid:Voiced	-0.021	0.054	-0.384	0.701
Low:Voiceless	-0.015	0.042	-0.359	0.720
Mid:Voiceless	-0.042	0.053	-0.790	0.430

3.3 AQ and Categorical Perception

This section explores whether the high-AQ groups showed less (or more) categorical responding, which would result in a flatter (or steeper) curve. The results show that all AQ groups had very similar identification curves. Figure 5 shows the proportion of voiced responses by AQ group, across all contexts and places of articulation. Both the moderate- and low-AQ groups (pink and purple lines, overlapping) had the category boundary at roughly 39ms VOT, while the low-AQ group's (yellow) boundary was around 37ms.

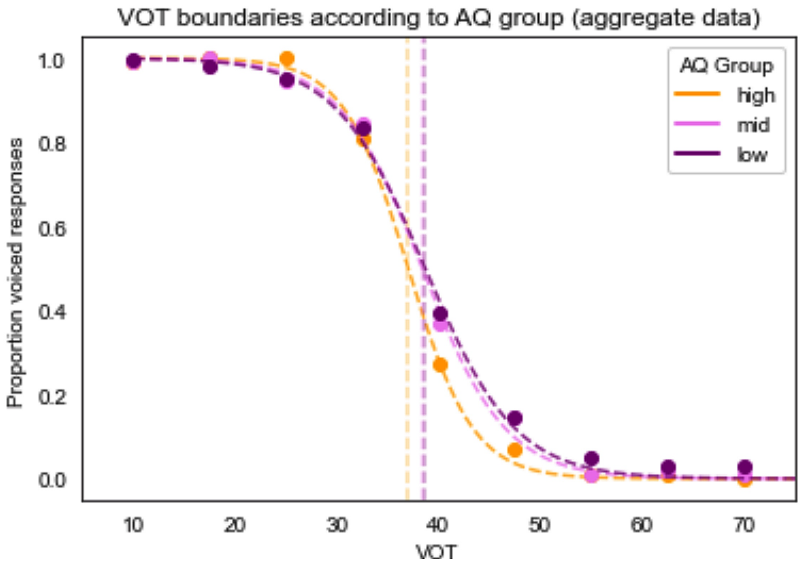

Figure 5. *Voiced response by AQ group*

To measure how categorical the responses were, the derivative of the curves was calculated at each point. The maximum derivative indicated the steepest points of each curve. When comparing AQ groups, all three had similar maximum derivatives (Figure 6, overleaf). Additionally, the maximum derivative of each participant was calculated (Figure 7, overleaf). The derivative only had a 0.03 correlation coefficient with AQ score; thus, AQ did not affect the degree to which responses were categorical or gradient.

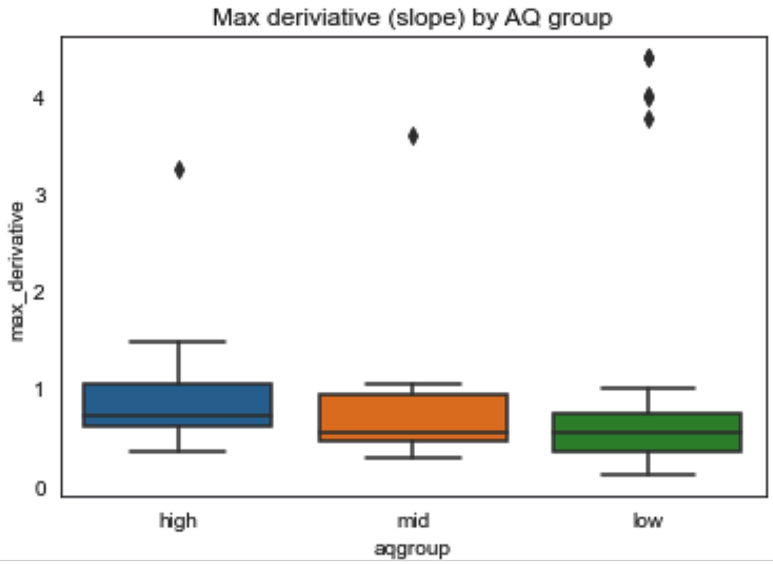

Figure 6. *Maximum derivative by AQ group*

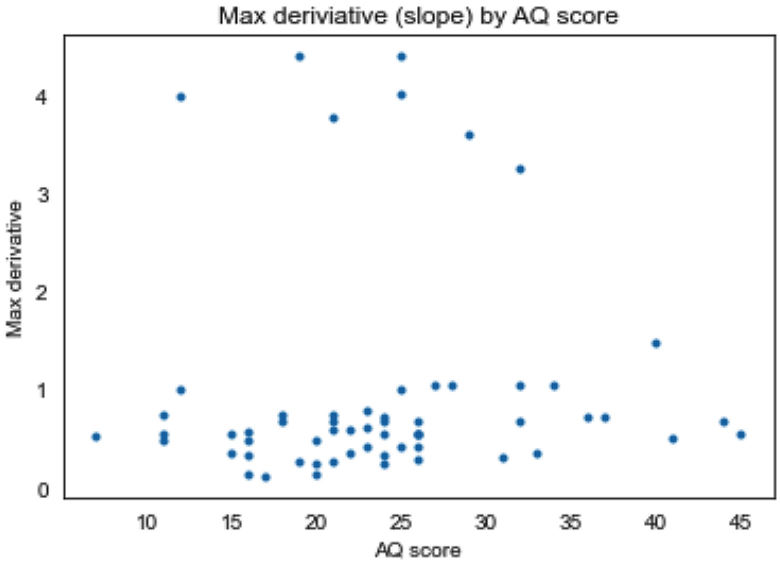

Figure 7. *Maximum derivative by individual AQ score*

4 Discussion

The findings are in line with the prediction that sentence context affects the categorical perception of VOT, but they do not support a model of speech perception in autism where top-down sentence processing is reduced compared to bottom-up perception of phonetic details.

Firstly, the study found that both the voiced and voiceless bias contexts had a significant main effect. When the stimuli was ambiguous, listeners were more likely to choose the semantically appropriate option. These results support the findings of Miller *et al*. (1984) and Schertz and Hawthorne (2018), who found that listeners take in both types of information — the lower-level phonetic details of the percept *and* the higher-level sentence context — and that listeners rely on sentence context to facilitate perception, especially when the stimuli is unfamiliar or ambiguous.

Secondly, the study found that all AQ groups shifted the category boundaries to a similar extent for the two bias contexts. According to different models of speech perception, sentence context either directly influences phonetic retuning (Mirman, 2008; Mirman *et al.,* 2006) or works at the decision level to guide phonetic retuning offline (Norris & McQueen, 2008; Norris, *et al.*, 2016). Because the study was an offline task (i.e. it did not observe the intermediate stages of processing through eye-tracking or electroencephalography/EEG), it cannot identify the exact level of sentence context's influence in speech perception. However, these accounts generally agree that sentence context is a higher-level representation than phonetic information. That sentence context significantly affected the high-AQ group, equally to the other groups, indicates that top-down processing is not different from neurotypical processing, at least in this kind of task.

Thirdly, the study did not find any statistically significant difference in the categorical perception of the AQ groups. This does not support the findings of Depape *et al.* (2012), who found that autistic individuals' discrimination of within- and between-category differences were more similarly accurate compared to TD controls, whose accuracy was lower in the within-category condition. It also does not support the findings from Haesen *et al.* (2011), who posit that autistic individuals focus on low-level perceptual details. The findings in this study are, however, in line with many studies that focused specifically on categorical perception in autism, which did not find phonemic knowledge to be less categorical (Chiodo *et al.*, 2019; Constantino *et al.*, 2007; White *et al.*, 2006; Stewart *et al.*, 2018).

The findings are not in line with the theoretical explanations of Weak Central Coherence (Happé & Frith, 2006), nor is it in line with Bayesian accounts that predict a reduced role of prior knowledge (e.g. Pellicano & Burr, 2012). This would have been demonstrated if context did not change the category boundaries for the high-AQ group, or if the boundary had shifted less, compared to the low-AQ group; this was not the case for this study. Additionally, the mentioned accounts predict enhanced bottom-up processing. For example, some Bayesian accounts argue for a reduced role of prior knowledge. This implies that knowledge of phonemic categories would also be reduced, leading to more gradient perception. Other Bayesian accounts argue that autistics show greater or more irregular precision of bottom-up processing (Brock, 2012;

Lawson *et al.,* 2014). This could have been demonstrated if the ambiguity of the intermediate stimuli (steps 3-7) was reflected in the high-AQ group's responses, such as by having mixed responses, resulting in a flatter curve, compared to the other groups; this was not the case for the results of this study.

A third Bayesian view argues that, for autistics, prior knowledge normally influences performance in perceptual tasks, except when the structure of perceptual information changes, at which point updating prior knowledge is challenging (Van de Cruys *et al*., 2014). One example could be asking participants to split the same VOT continuum into three categories, when there were previously only two. Relatedly, visual studies (Sapey-Triomphe *et al.*, 2018; Soulières *et al.*, 2007) show that categorical perception can be developed through exposure, as they tested novel categories (e.g. ellipses, varying in length but not height). In this study, participants were very familiar with the categories, since the study tested English-speakers on native phonemic contrasts. It is possible that AQ would have a stronger effect if this study had tested foreign-language contrasts, rather than native ones.

Finally, the Enhanced Perceptual Model account (Mottron *et al.*, 2006) argues that top-down processing is only used by autistics in perceptual tasks when it facilitates performance. In this case, the ability to predict which words are likely to appear in sentences based on context would make speech recognition more efficient, and possibly less prone to error (Chow, 2020). This could explain why the findings show a normal influence of sentence context in the high-AQ group. Moreover, in speech production, speakers show a wide range of individual differences For example, many studies document individual variation in VOT production (Allen *et al*., 2003, Newman, 1997). Bottom-up processing of all between- and within-speaker variation may lead to perceptual overloading (Chiodo *et al.*, 2019); efficient speech perception relies on categorical phonemic knowledge. Thus, based on the Enhanced Perceptual Model, the finding that high-AQ individuals did not have reduced categorical perception is expected.

The study was unable to explicitly test the model's assumption that bottom-up processing is enhanced as this would have been better demonstrated through a discrimination task. In that case, the high-AQ group is expected to better identify within-category differences. As with the Bayesian accounts, the absence of more gradient perception in the high-AQ group could indicate that bottom-up processing is not significantly enhanced. However, it could be that bottom-up processing *is* enhanced, but that top-down processing is still heavily weighted in decision-making when it seems facilitative to do so. Thus, the results partially support the Enhanced Perceptual Model hypothesis.

One reason the study did not find a significant effect of AQ could be that the sentences were quite literal. Although, following Schertz and Hawthorne (2018), the sentence contexts were high-level enough to skew perception of ambiguous phonetic information, it is possible they were not high-level enough to produce a significant effect of AQ. This is plausible given that the evidence of weakly integrated top-down processing in autism comes from tasks like interpreting emotions from intonation or facial expressions (Ashwin *et al.*, 2007; Boraston *et al.*, 2007; Kujala *et al.,* 2005; McCann and Peppé 2003). Given that pragmatic deficits are the most well-documented feature of language in autism (e.g. metaphor comprehension, see Kalandadze *et al.*, 2019), it is possible that a task involving more pragmatic inferences would have produced a stronger effect of AQ.

One limitation of this study is that it did not test clinically diagnosed autistic participants — though, following the cut-off score (26/50) from Woodbury-Smith *et al.* (2005), there were presumably some autistic individuals in at least the high-AQ group. Although the AQ is used clinically as a screening test (Yu & Zellou, 2019), it was designed for adults with at least average IQ (Baron-Cohen *et al.*, 2001). Given that autistic individuals vary greatly in both verbal- and non-verbal IQ (Anderson, Liang & Lord, 2014), the findings may not generalise to all autistic people. In the future, this study should be repeated with clinically diagnosed autistic individuals.

A related limitation is that this study did not test participants' IQ. Previous studies that found enhanced bottom-up perception in autistics often relied on verbal IQ to match participants to controls (e.g. Lepistö *et al.*, 2009; Teder-Sälejärvi *et al.,* 2005); this might have led to underestimation of autistic individuals' competence in other domains (Dawson *et al.*, 2007; Soulières *et al.*, 2011). Chiodo *et al.* (2019) illustrate the importance of this consideration: when matching autistic individuals with and without speech-onset delay, those with the same verbal IQ could differ greatly in non-verbal IQ, and vice versa. In the visual domain, evidence of enhanced low-level processing, demonstrated through perceptual peaks, disappeared when autistic and control groups were matched on non-verbal IQ (see Barbeau *et al.*, 2013; Dawson *et al.,* 2007). Moreover, autistic individuals with average or above-average IQ often express autistic traits to a lesser degree (Anderson, *et al.*, 2014). In the future, this study should be repeated with autistic individuals with a range of verbal and non-verbal IQ, and have IQ-matched controls.

Finally, some studies have suggested that only autistic individuals with language delay may express atypical auditory perception skills (e.g. Bonnel & Hafter, 2006; Bonnel *et al.*, 2010; Wodka *et al.*, 2013). They argue that atypical speech perception can arise from poor language development and vice versa, which has been found in ASD (McBride-Chang, 1996; Morais *et al.,* 1986; Serniclaes, 2006). Atypical low-level perception, such as difficulties with ignoring irrelevant (within-category) phonetic details, may hinder identification of native-language characteristics more broadly (Flagg *et al.*, 2005; Teder-Sälejärvi *et al.,* 2005; Serniclaes *et al.*, 2001). Although this study could not investigate participants' history of speech-onset delay, participants did not show significant language impairment, as they were able to independently navigate the task with only written instructions. Future studies testing autistic individuals with a history of language delay or impairment might see a clearer interaction of AQ and sentence context or categorical perception.

5 Conclusion

To summarise, this study found that sentence context and place of articulation influenced the category boundaries of English stops. This was true of all AQ groups. It also found that high-AQ participants did not show a reduced effect of sentence context, nor was their categorical perception reduced. This is consistent with findings in the literature that language skills are greatly variable in ASD, with some individuals' skills being comparable to TD controls (Chiodo *et al.,* 2019; Constantino *et al.,* 2007; Loucas *et al.*, 2008; Stewart *et al.*, 2018; Tager-Flusberg, 2000; White *et al.,* 2006).

The results do not support cognitive theories of ASD where low-level stimuli is enhanced at the expense of top-down processing, at least with respect to the phonemic categories and sentence contexts tested. Based on the findings for enhanced low-level perception in ASD, it could be that enhanced perception is restricted to the visual domain (Baron-Cohen & Hammer, 1997; Plaisted *et al.,* 1999; Sapey-Triomphe *et al.*, 2018; Soulières *et al.*, 2007), or to simple acoustic stimuli in the auditory domain (Lepistö *et al.*, 2009; Teder-Sälejärvi *et al.*, 2005).

References

Abada, S. H., Baum, S. R., & Titone, D. (2008). The effects of contextual strength on phonetic identification in younger and older listeners. *Experimental aging research*, *34*(3), 232-250.

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, *113*(1), 544-552.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *55*(5), 485-494.

Angosto, A., Sánchez, P., Álvarez, M., Cuevas, I., & León, J. A. (2013). Evidence for top-down processing in reading comprehension of children. *Psicología Educativa*, *19*(2), 83-88.

Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (2007). Differential activation of the amygdala and the 'social brain'during fearful face-processing in Asperger Syndrome. *Neuropsychologia*, *45*(1), 2-14.

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.

Barbeau, E. B., Soulières, I., Dawson, M., Zeffiro, T. A., & Mottron, L. (2013). The level and nature of autistic intelligence III: Inspection time. *Journal of abnormal psychology*, *122*(1), 295.

Baron-Cohen, S., & Hammer, J. (1997). Parents of children with Asperger syndrome: what is the cognitive phenotype?. *Journal of cognitive neuroscience*, *9*(4), 548-554.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of autism and developmental disorders*, *31*(1), 5-17.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Becker, Michael and Jonathan Levine (2013) Experigen – an online experiment platform. Available at http://becker.phonologist.org/experigen.

Belmonte, M. K., Saxena-Chandhok, T., Cherian, R., Muneer, R., George, L., & Karanth, P. (2013). Oral motor deficits in speech-impaired children with autism. *Frontiers in Integrative Neuroscience*, *7*, 47.

Bishop, J. (2012). WPP, No. 111: Focus, prosody, and individual differences in "autistic" traits: Evidence from cross-modal semantic priming.

Bonnel, A. M., & Hafter, E. (2006). Sensitivity to increments analyzed separately for each level in a roving paradigm. *The Journal of the Acoustical Society of America*, *120*(5), 3127-3127.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *Journal of cognitive neuroscience*, *15*(2), 226-235.

Boraston, Z., Blakemore, S. J., Chilvers, R., & Skuse, D. (2007). Impaired sadness recognition is linked to social interaction deficit in autism. *Neuropsychologia*, *45*(7), 1501-1510.

Borsky, S., Tuller, B., & Shapiro, L. P. (1998). “How to milk a coat:” The effects of semantic and acoustic information on phoneme categorization. *The Journal of the Acoustical Society of America*, *103*(5), 2670-2676.

Boucher, J. (2012). Research review: structural language in autistic spectrum disorder–characteristics and causes. *Journal of Child Psychology and Psychiatry*, *53*(3), 219-233.

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: comment on Pellicano and Burr. *Trends in cognitive sciences*, *16*(12), 573.

Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics*, *71*, 51-64.

Chechik, G., Meilijson, I., & Ruppin, E. (1999). Neuronal regulation: A mechanism for synaptic pruning during brain maturation. *Neural computation*, *11*(8), 2061-2080.

Chiodo, L., Mottron, L., & Majerus, S. (2019). Preservation of categorical perception for speech in autism with and without speech onset delay. *Autism Research*, *12*(11), 1609-1622.

Chow, W. Y., & Chen, D. (2020). Predicting (in) correctly: listeners rapidly use unexpected information to revise their predictions. *Language, Cognition and Neuroscience*, *35*(9), 1149-1161.

Clarke-Davidson, C. M., Luce, P. A., & Sawusch, J. R. (2008). Does perceptual learning in speech reflect changes in phonetic category representation or decision bias?. *Perception & psychophysics*, *70*(4), 604-618.

Connine, C. M. (1987). Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*, *26*(5), 527-538.

Cole, R. A., & Jakimik, J. (1978). Understanding speech: How words are heard. *Strategies of information processing*, 67-116.

Constantino, J. N., Yang, D., Gray, T. L., Gross, M. M., Abbacchi, A. M., Smith, S. C., ... & Kuhl, P. K. (2007). Clarifying the associations between language and social development in autism: A study of non-native phoneme recognition. *Journal of Autism and Developmental Disorders*, *37*(7), 1256-1263.

Courchesne E (2004) Brain development in autism: Early overgrowth followed by premature arrest of growth. Mental Retardation and Developmental Disability Research Review 10: 106–111

Dawson, M., Soulières, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological science*, *18*(8), 657–662. https://doi.org/10.1111/j.1467-9280.2007.01954.x

DePape, A. M. R., Hall, G. B., Tillmann, B., & Trainor, L. J. (2012). Auditory processing in high-functioning adolescents with autism spectrum disorder. *PloS one*, *7*(9), e44084.

Drouin, J. R., Theodore, R. M., & Myers, E. B. (2016). Lexically guided perceptual tuning of internal phonetic category structure. *The Journal of the Acoustical Society of America*, *140*(4), EL307-EL313.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & psychophysics*, *67*(2), 224-238.

Fernández, E. M., & Cairns, H. S. (2010). *Fundamentals of psycholinguistics*. John Wiley & Sons.

Flagg, E. J., Cardy, J. E. O., Roberts, W., & Roberts, T. P. (2005). Language lateralization development in children with autism: insights from the late field magnetoencephalogram. *Neuroscience letters*, *386*(2), 82-87.

Folstein, S. E., & Rosen-Sheidley, B. (2001). Genetics of autism: complex aetiology for a heterogeneous disorder. *Nature Reviews Genetics*, *2*(12), 943-955.

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human perception and performance*, *28*(2), 349.

Frith, U. (2003). *Autism: Explaining the enigma*. Blackwell Publishing.

Gopnik, A., Meltzoff, A. N., & Kuhl, P. K. (1999). *The scientist in the crib: Minds, brains, and how children learn*. William Morrow & Co.

Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. *Research in autism spectrum disorders*, *5*(2), 701-714.

Happé, F. (2005). The weak central coherence account of autism. *Handbook of autism and pervasive developmental disorders*, *1*, 640-649.

Happé, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of autism and developmental disorders*, *36*(1), 5-25.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of autism and developmental disorders*, *35*(6), 787.

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive neuropsychology*, *25*(6), 771-782.

Huang, H. C. (2007). Lexical context effects on speech perception in Chinese people with autistic traits.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434-446.

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

Jesse, A. (2019). Sentence context guides phonetic retuning to speaker idiosyncrasies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Jesse, A., & McQueen, J. M. (2011). Positional effects in the lexical retuning of speech perception. *Psychonomic bulletin & review*, *18*(5), 943-950.

Jones, C. R., Happé, F., Baird, G., Simonoff, E., Marsden, A. J., Tregay, J., ... & Charman, T. (2009). Auditory discrimination and auditory shensory behaviours in autism spectrum disorders. *Neuropsychologia*, *47*(13), 2850-2858.

Kalandadze, T., Bambini, V., & Næss, K. A. B. (2019). A systematic review and meta-analysis of studies on metaphor comprehension in individuals with autism spectrum disorder: Do task properties matter?. *Applied Psycholinguistics*, *40*(6), 1421-1454.

Kern, J. K., Trivedi, M. H., Garver, C. R., Grannemann, B. D., Andrews, A. A., Savla, J. S., ... & Schroeder, J. L. (2006). The pattern of sensory processing abnormalities in autism. *Autism*, *10*(5), 480-494.n

Key, A. P., Yoder, P. J., & Stone, W. L. (2016). Consonant differentiation mediates the discrepancy between non‐verbal and verbal abilities in children with ASD. *Journal of Intellectual Disability Research*, *60*(5), 478-490.

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

Kuhl, P., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *Annu. Rev. Neurosci.*, *31*, 511-534.

Kujala, T., Lepistö, T., & Näätänen, R. (2013). The neural basis of aberrant speech and audition in autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, *37*(4), 697-704.

Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in human neuroscience*, *8*, 302.

Lepistö, T., Kuitunen, A., Sussman, E., Saalasti, S., Jansson-Verkasalo, E., Nieminen-von Wendt, T., & Kujala, T. (2009). Auditory stream segregation in children with Asperger syndrome. *Biological psychology*, *82*(3), 301-307.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, *74*(6), 431.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental psychology*, *54*(5), 358.

Lin, C. Y., & Wang, H. C. (2011). Automatic estimation of voice onset time for word-initial stops by applying random forest to onset detection. *The Journal of the Acoustical Society of America*, *130*(1), 514-525.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, *20*(3), 384-422.

Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The Lancet*, *392*(10146), 508-520.

Lord, C., & Paul, R. (1997). Language and communication in autism. In D. J. Cohen, &F. R. Volkmar (Eds.), Handbook of autism and pervasive developmental disorders (2nd ed., pp. 195–225). New York: Wiley.

Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, *49*(11), 1184-1192.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, *10*(1), 29-63.

Matsuzaki, J., Kuschner, E. S., Blaskey, L., Bloy, L., Kim, M., Ku, M., ... & Roberts, T. P. (2019). Abnormal auditory mismatch fields are associated with communication impairment in both verbal and minimally verbal/nonverbal children who have autism spectrum disorder. *Autism Research*, *12*(8), 1225-1235.

McBride‐Chang, C. (1996). Models of speech perception and phonological processing in reading. *Child development*, *67*(4), 1836-1856.

McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders*, *38*(4), 325-350.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33-B42.

McMurray, B., Tanenhaus, M. K., Aslin, R. N., & Spivey, M. J. (2003). Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of Psycholinguistic Research*, *32*(1), 77-97.

Miller, J. L., Green, K., & Schermer, T. M. (1984). A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics*, *36*(4), 329-337.

Meilleur, A. A. S., Jelenic, P., & Mottron, L. (2015). Prevalence of clinically and empirically defined talents and strengths in autism. *Journal of autism and developmental disorders*, *45*(5), 1354-1367.

Mirman, D. (2008). Mechanisms of semantic ambiguity resolution: Insights from speech perception. *Research on Language and Computation*, *6*(3-4), 293-309.

Mirman, D., McClelland, J. L., & Holt, L. L. (2006). An interactive Hebbian account of lexically guided tuning of speech perception. *Psychonomic bulletin & review*, *13*(6), 958-965.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition*, *24*(1-2), 45-64.

Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism.

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *Journal of autism and developmental disorders*, *36*(1), 27-43.

Nakai, S., & Scobbie, J. M. (2016). The VOT category boundary in word-initial stops: Counter-evidence against rate normalization in English spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*.

Newman, R. (1997). *Individual differences and the link between speech perception and speech production* (Doctoral dissertation, University at Buffalo).

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review, 115,* 357–395. http://dx.doi.org/10.1037/0033-295X.115.2.357

Norris, D., McQueen, J. M., & Cutler, A. (2016). Prediction, Bayesian inference and feedback in speech recognition. *Language, cognition and neuroscience*, *31*(1), 4-18.

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in cognitive sciences*, *16*(10), 504-510.

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of child psychology and psychiatry*, *40*(5), 733-742.

Pollack, I., & Pickett, J. M. (1963). The intelligibility of excerpts from conversation. *Language and Speech*, *6*(3), 165-171.

R Core Team (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Retrieved from http://www.r-project.org/.

Sapey-Triomphe, L. A., Sonié, S., Hénaff, M. A., Mattout, J., & Schmitz, C. (2018). Adults with autism tend to undermine the hidden environmental structure: evidence from a visual associative learning task. *Journal of autism and developmental disorders*, *48*(9), 3061-3074.

Saul, J., & Norbury, C. (2020). Does phonetic repertoire in minimally verbal autistic preschoolers predict the severity of later expressive language impairment?. *Autism*, 1362361319898560.

Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of memory and language*, *24*(2), 210-231.

Schertz, J., & Hawthorne, K. (2018). The effect of sentential context on phonetic categorization is modulated by talker accent and exposure. *The Journal of the Acoustical Society of America*, *143*(3), EL231-EL236.

Serniclaes, W. (2006). Allophonic perception in developmental dyslexia: Origin, reliability and implications of the categorical perception deficit. *Written Language & Literacy*, *9*(1), 135-152.

Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Demonet, J. F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia.

Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *The Journal of the Acoustical Society of America*, *132*(2), EL95-EL101.

Soulières, I., Mottron, L., Saumier, D., & Larochelle, S. (2007). Atypical categorical perception in autism: Autonomy of discrimination?. *Journal of autism and developmental disorders*, *37*(3), 481-490.

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with “autistic” traits. *Cognition*, *109*(1), 157-162.

Stewart, M. E., Petrou, A. M., & Ota, M. (2018). Categorical speech perception in adults with autism spectrum conditions. *Journal of autism and developmental disorders*, *48*(1), 72-82.

Sullivan, L. (2019). *The Effect of Cognitive Processing Style on the Perceptual Compensation of Stop Voicing for Place of Articulation*. Presentation, University of Toronto.

Tager-Flusberg, H. (2000). Understanding the language and communicative impairments in autism. In *International review of research in mental retardation* (Vol. 23, pp. 185-205). Academic Press.

Teder-Sälejärvi, W. A., Pierce, K. L., Courchesne, E., & Hillyard, S. A. (2005). Auditory spatial localization and attention deficits in autistic adults. *Cognitive Brain Research*, *23*(2-3), 221-234.

Turnbull, R. (2015). *Assessing the listener-oriented account of predictability-based phonetic reduction* (Doctoral dissertation, The Ohio State University).

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: predictive coding in autism. *Psychological review*, *121*(4), 649.

Van der Zande, P., Jesse, A., & Cutler, A. (2013). Lexically guided retuning of visual phonetic categories. *The Journal of the Acoustical Society of America*, *134*(1), 562-571.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, *67*(6), 1176-1190.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., … SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Weenink, P. B. D., & Boersma, P. (2008). Praat: doing phonetics by computer.

Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual review of psychology*, *50*(1), 509-535.

White, S., Frith, U., Milne, E., Rosen, S., Swettenham, J., & Ramus, F. (2006). A double dissociation between sensorimotor impairments and reading disability: A comparison of autistic and dyslexic children. *Cognitive Neuropsychology*, *23*(5), 748-761.

Winn, M. B. (2020). Manipulation of voice onset time in speech stimuli: A tutorial and flexible Praat script. *The Journal of the Acoustical Society of America*, *147*(2), 852-866.

Wodka, E. L., Mathy, P., & Kalb, L. (2013). Predictors of phrase and fluent speech in children with autism and severe language delay. *Pediatrics*, *131*(4), e1128-e1134.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of autism and developmental disorders*, *35*(3), 331-335.

Yoder, P., Watson, L. R., & Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. *Journal of autism and developmental disorders*, *45*(5), 1254-1270.

You, R. S., Serniclaes, W., Rider, D., & Chabane, N. (2017). On the nature of the speech perception deficits in children with autism spectrum disorders. *Research in developmental disabilities*, *61*, 158-171.

Yu, A. C., & Zellou, G. (2019). Individual differences in language processing: Phonology. *Annual Review of Linguistics*, *5*, 131-150.