University of Southampton

Faculty of Environmental and Life Sciences

School of Psychology

Sex/Gender Differences in Speech Disfluency in Children with Autism Spectrum Condition: On the Distribution of 'Um' and 'Uh '

by

Bonnie Wong

Thesis for the degree of Doctorate in Educational Psychology

June 2019

University of Southampton

Abstract

Faculty of Environmental and Life Sciences

School of Psychology

Thesis for the degree of Doctorate in Educational Psychology

Sex/Gender Differences in Speech Disfluency in Children with Autism Spectrum Condition: On the Distribution of ‘Um’ and ‘Uh ’

by

Bonnie Wong

The first chapter presents a systematic review on the literature quantifying patterns of disfluency in autistic populations. Disfluencies are hypothesised to be an interruption of a speech plan with approximately 6% of spontaneous speech being disfluent. Around half of the autistic population display quantitatively and qualitatively different speech patterns compared with the neurotypical population. Therefore, these different speech patterns could represent a behavioural marker for autism. Twelve studies were screened and met the following inclusion criteria: empirical studies quantifying types of disfluencies in autism, with a sample size greater than five, and the inclusion of a neurotypical comparison group. The results indicated that filled pauses were produced significantly less frequently by autistic individuals relative to neurotypical controls. Autistic individuals produced significantly higher amounts of unfilled pauses, repetitions, and repairs than neurotypical groups. The pattern of results suggest that there may be meaningful differences in the production of disfluencies between autistic and neurotypical individuals. However, there are a number of methodological and conceptual issues (unaccounted moderating and mediating variables, varied methodologies, heterogeneity of autism) to overcome before meaningful conclusions can be drawn and generalised.

The purpose of the empirical study was to address the methodological issues that fewer pauses found previously in autistic groups may simply reflect sex differences rather than a phenotypic expression of autism due to sex-imbalanced groups. We included nearly equal numbers of males and females, to compare the use of filled pauses (Um, Uh) between autistic and neurotypical participants, and between sex, as well as the interaction between the two. In addition, filled pause use was also examined to investigate the linguistic camouflaging hypothesis. Eighty-four participants aged 8-13 from mainstream schools completed interviews (structured, unstructured) as well as tests measuring their verbal and nonverbal IQ. Neither significant main effects nor interactions were found for all indices for the structured task. For the unstructured task, no interactions were found. However, autistic participants produced a significantly higher proportion of Um to total words than neurotypical participants. Males produced a significantly higher proportion of Uhs than females, and females produced a higher Um ratio than males. Differences in filled pause production was affected by the type of interviews. The linguistic camouflaging hypothesis was disconfirmed. Future research should include larger samples (e.g., more females in the autistic group, children with high traits), different task types, and listeners' perceptions in order to further understand filled pauses in autism.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name:	Bonnie Wong

Title of thesis:	Sex/Gender Differences in Disfluency Use in Children with Autism Spectrum Condition: On the Distribution of 'Um' and 'Uh '

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signature:		Date:	24/07/2019

Acknowledgements

I would like to extend my thanks to my supervisors - Dr. Julie Hadwin and Dr. Hanna Kovshoff at the University of Southampton - their research knowledge and logical way of thinking have been of great value to me.

I would also like to express my sincere gratitude to my personal tutor, Dr. Sarah Wright, who have provided support and guidance through periods of personal difficulties. She has modelled psychology in practice in a way that I can only aspire to.

I also owe a deep thank you to my remarkable placement and research partner, Henry Wood, who has shared the ups and downs of the DEdPsych journey.

Last but not least, um… I would not be who I am as a person without the unconditional support and nurture of my parents and my sister:

虎父果然無犬女, 父母德高, 子女良教, 養育之恩, 無以為報

Definitions and Abbreviations

The terminology used follows the definitions recommended by the World Health Organisation, such that 'sex' refers to the biological differences between males and females in growth whilst gender refers to a socio- cultural construct such as roles, attributes, and opportunities that one society considers appropriate for males and females (Manandhar, Hawkes, Buse, Nosrati, & Magar, 2018). Gender is a highly nuanced social construct and should be approached as such in studies, researchers have not yet been able to separate the effects of sex and gender and therefore 'sex difference' is used throughout this thesis to reflect this.

In line with preferences expressed by self-advocates within the autistic community, identity first language (e.g. autistic participants) will be used throughout this paper (Dunn & Andrews, 2015; Kenny et al., 2016).

ADOS	Autism Diagnostic Observation Schedule
AS	Asperger Syndrome
AXIS	Appraisal tool for Cross-Sectional Studies
ASD	Autism Spectrum Disorder
BPVS	British Picture Vocabulary Scale
CCC	Children's Communication Checklist
CELF	Clinical Evaluation of Language Fundamentals
DSM-5	Diagnostic and Statistical Manual of Mental Disorders 5
HFA	High-functioning autism
ICD-11	International statistical classification of diseases and related health problems 11^{th} Revision

IQ	Intelligence Quotient
NT	Neurotypical
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WISC-III	Wechsler Intelligence Scale for Children

Chapter 1 Is Speech Disfluency a Marker for Autism? A Systematic Review.

1.1 Introduction

Autism Spectrum Disorder or Autism Spectrum Condition, (hereafter 'autism'), is a neurodevelopmental condition with heterogeneous phenotypes. It is characterised by early onset difficulties in social interaction and communication, and the presence of repetitive and restricted behaviours and/or interests (American Psychiatric Association, 2013). It is estimated to affect approximately 1-2% of the population and has predominantly been diagnosed in males, with male to female ratio of 4 to 1 (Baird et al., 2006). To date, it has not been possible to identify reliable biomarkers in the identification and diagnosis of autism. As a result, autism is diagnosed at the behavioural level (Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2015).

1.1.1 Language Impairment in Autism

One way in which social interaction and communication difficulties can manifest in individuals with autism is through language impairment. Indeed, there is a high prevalence of language impairment in individuals diagnosed with autism (Eigsti, De Marchena, Schuh, & Kelley, 2011). For example, Kjelgaard and Tager-Flusberg (2001) found that 76% of the autistic participants (aged 4 to 14) performed significantly below age expectations in standardised language measures such as Clinical Evaluation of Language Fundamentals (CELF; Semel, Wiig, & Secord, 2000). Similarly, a population-based cohort study carried out by Loucas et al. (2008) found that 56.9% of children (aged 9 to 14) with autism (and at least low-average non-verbal IQ measured by the Wechsler Intelligence Scale for Children) also met criteria for language impairment as measured by a battery of tests such as British Picture Vocabulary Scale 2nd Edition (Dunn et al., 1997), CELF, and Children's Communication Checklist (CCC; Bishop, 1998). The participants in the latter study were drawn from a total population cohort of 56946 children born between July 1990 and December 1991 in South East Thames (Baird et al., 2006) and should therefore be relatively representative of the extent to which language impairment is co-morbid with autism.

Language is hypothesised to be a uniquely human cognitive faculty that evolved for survival purposes partly through facilitating social bonding and social interaction (Dunbar, 2003; Pinker, 2013). For example, ancient monkeys and apes spent approximately 43% of daytime on social grooming to increase the social bonds amongst their groups. However, Dunbar (2009) argued that language evolved so that modern humans require less time (approximately 20%) to bond. In other words, the use of language allows time to be used more efficiently as humans can bond more quickly and with more than one human at a time. Despite humans' seemingly innate language ability, language development is highly variable in autistic individuals, and deficits and delays are relatively common (Whitehouse, Barry, & Bishop, 2007). Although longitudinal studies have found that autistic individuals' language abilities often improve over time, a recent meta- analysis found that children and young people between the ages of 6 and 19 with autism have significant deficits in receptive and expressive language skills; both skills were approximately 1.5 standard deviations behind their neurotypical peers (Kwok, Brown, Smyth, & Oram Cardy, 2015). In addition, pragmatic language impairments – socially oriented elements of language skill impairments – are thought to be highly prevalent in autistic individuals (Eigsti et al., 2011; Magiati, Tay, & Howlin, 2014) and thus form part of the diagnostic criteria as per the Diagnostic and Statistical Manual of Mental Disorders 5 (DSM-5; American Psychiatric Association, 2013) and the International statistical classification of diseases and related health problems 11th Revision (ICD11; World Health Organization, 2018).

1.1.2 Pragmatic Language

Although atypicalities in pragmatic language have been found amongst autistic individuals in previous research, pragmatic language is a broad concept with little consensus on how it should be defined or quantified by researchers (Young, Diehl, Morris, Hyman, & Bennetto, 2005). Prutting and Kirchner (1987) offered one of the first definitions of pragmatic language, namely the ability to use language in specific contexts and for specific purposes, which concerns both verbal and non-verbal aspects of language. Verbal aspects of pragmatic language include linguistic functions (e.g., register - altering one's speech content depending upon the listener, topic maintenance, and appropriate turn taking) and vocal production (e.g., pitch, prosody, volume). On the other hand, non-verbal aspects of pragmatic language include non-linguistic functions such as eye contact, body language and gestures (Young et al., 2005). Modern definitions of pragmatic language have expanded beyond communicative functions to reflect the notion that pragmatic language, social skills and emotional understanding are interlinked (Parsons et al., 2017). In this review, Prutting and Kirchners' (1987) definition is followed, and one quantifiable feature of pragmatic language – disfluencies – are examined in autistic populations.

1.1.3 Disfluencies

Disfluencies are hypothesised to be an interruption of a speech plan - instead of a deviation from this plan (Postma, Kolk, & Povel, 1990) - and may reflect difficulties with planning and delivering speech (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Scott, 2015). It is estimated that approximately 6% of utterances in typical spontaneous speech are disfluent (Fox Tree, 1995). Due to their relatively high prevalence in spontaneous speech, there is an increasing body of research focusing on disfluencies. The main types of disfluencies that have been investigated in research are pauses (an interval of pause during speech production), repetitions (repeating one or more words during speech production without correction of previous utterance), and repairs (repeating one or more words during speech production by correcting previous utterance). Bock (1996) reviewed the literature and methodologies of existing experimental paradigms of language production, and proposed that different disfluencies represent different underlying language production difficulties or are reflections of different strategies for correcting speech problems. Fraundorf & Watson (2014) tested this hypothesis in an experimental study, in which the participants' speech in a storytelling paradigm was coded and analysed for three disfluency types: repetition, filled pauses, and unfilled pauses. They found that repetitions were more likely to be produced when there was recently articulated material available to repeat. In other words, repetitions were more likely to be used than filled or unfilled pauses when participants encountered production difficulties and utterances of speech were already being produced. In addition, filled pauses (e.g. use of 'um', 'uh', 'you know') and unfilled pauses (unfilled gaps in speech) were used when speakers were between units of speech but articulation of the utterances had not yet begun. Specifically, they found that speakers most commonly used filled pauses when they were planning how to convey their next message, reflecting difficulties on a message (i.e., pre-verbal)level . However, unfilled pauses occurred when the speakers had difficulties not only on a message-level, but also on grammatical (i.e. grammatical encoding processes such as assigning grammatical roles, creating a linear sequence) and/or phonological levels (i.e. phonological encoding processes such as retrieving the sounds and organising them into a phonetic plan).

Despite this, it is important to note that the presence of disfluencies may not always reflect planning problems; fillers like 'um' may be used deliberately by speakers to indicate an upcoming pause or a 'dispreferred' response (Schegloff, 2010, p.141). For example,

Speaker A: It's sunny outside, shall we go for a walk?

Speaker B: Um…, maybe later when I finish writing this part of my thesis.

In this example, speaker B's response was dispreferred in the sense that it contradicts speaker A's original suggestion of going for a walk.

Traditionally, disfluencies were conceptualised by researchers as a hindrance during the speech comprehension process to the listener (e.g. Martin & Strange, 1968). However, recent findings from psycholinguistic research regarding the effects of disfluencies on a listener's comprehension showed that disfluencies can be both beneficial and detrimental to speech comprehension (Barr & Seyfeddinipur, 2010). For example, some findings suggest that the presence of repair (a type of disfluency used to resolve and/or highlight instances of miscommunication in previous utterances) benefits listeners' comprehension, such that a speaker's repair helps to achieve and maintain shared understanding amongst participants (Hellermann, 2009; Purver, Hough, & Howes, 2018). Repair occurs during conversations when some aspect of talk is acknowledged by one or more participants as erroneous for some reason (e.g. mispronunciation, mishearing, insufficient information provided). The speech can be remedied either through self-initiated repair (repair initiated by the speaker who produced the erroneous utterance) or other-initiated repair (erroneous utterance noted by others and that person initiates repair):

Speaker: My child is 5 years old, uh, Archie, he is 5.

In this utterance, the self- initiated repair of 'Archie, he is 5' enables the speaker to provide additional information about their child's sex and thus achieve and may maintain shared understanding of the speaker's child during the conversation. In addition, some findings suggest that the presence of disfluency benefits listeners' comprehension – a speaker's disfluency helps to introduce novel and/or low frequency information to a conversation (Owens, Thacker, & Graham, 2018):

Speaker: We have a table… [short pause] um…, a marble top table.

In this utterance, the short pause enables the speaker to signal that they are unlikely to refer to a high frequency (i.e. common or familiar in everyday discourse) object, such as a wooden table.

In contrast, some findings suggest that disfluencies impede listener's comprehension. For example, recent studies used event-related potentials (ERPs) to measure participants' neural responses to disfluency (e.g. repetition of words) in speech (Collard, Corley, MacGregor, & Donaldson, 2008; MacGregor, Corley, & Donaldson, 2009). A P300 wave elicited in the participants suggested that they were sensitive to the disfluencies they encountered and found it more difficult to resume linguistic processing following disruption to speech.

Speech disfluencies also affect how speakers are perceived by the listeners socially, and elevated rates of disfluency have been associated with negative social perceptions of the speaker (Panico, Healey, Brouwer, & Susca, 2005). For example, previous research found that the manner in which speech is delivered affects a listener's implicit judgement. Specifically, disfluent utterances bias listeners toward perceiving the speaker to be untruthful (e.g. King, Loy, & Corley, 2018; Loy, Rohde, & Corley, 2017). Furthermore, elevated rates of disfluency use can also lead listeners to infer that the speaker lacks knowledge about the conversational topic (e.g. Arnold, Fagnano, & Tanenhaus, 2003; Birch, Vauthier, & Bloom, 2008; Orena & White, 2015).

Overall, previous research has shown that males are more likely to produce disfluent speech than females (Bortfeld et al., 2001). Moreover, a number of studies have found that language production shows reliable age-related declines; specifically, older adults are more likely than young adults and children to produce disfluencies (e.g., unfilled pauses mid-sentence) during naturalistic speech (Shafto & Tyler, 2014). However, the predictive value of disfluencies have been found to be moderated by perceived characteristics of the speaker (Kidd, White, & Aslin, 2011). For example, children only used disfluencies as a referent (i.e. prediction about a speaker's intended referent) when they also perceived the speaker to be knowledgeable (Orena & White, 2015). In general, disfluencies occur at a rate of approximately 6 per 100 fluent words in typical spontaneous speech (Bortfeld et al., 2001; Clark & Fox Tree, 2002; Shriberg, 1996). In addition, they tend to occur at a higher rate in longer utterances (Shriberg, 1994) and in more complex utterances (Lickley, 2001). There is considerable individual variations in disfluency rates, but it is unlikely that any one individual never produces any disfluencies (Lickley, 2017).

1.1.4 Patterns of Disfluency Observed in Autism

Although 70-80% of autistic individuals develop functional spoken language, research suggests that approximately half of the autistic population display quantitatively (e.g., higher amount of filled pauses) and qualitatively (e.g., atypical tone) different speech patterns compared with the neurotypical population (Fusaroli, Lambrechts, Bang, Bowler, & Gaigg, 2017; Shriberg et al., 2001). There is a growing body of literature studying the types and rates of disfluency production in autistic individuals. Shriberg et al. (2001) first studied speech and prosody characteristics of adolescents and adults (average age of 20.7 years) with high-functioning autism (HFA) and Asperger Syndrome (AS), compared with neurotypical controls. Greater amounts of disfluencies in parsing (e.g. syllable repetition, single-word revision) and rate (e.g. articulation rate, pause time) were found in individuals with HFA and AS. Several studies have since reported differences in patterns of disfluencies between autistic individuals and neurotypical controls (e.g. Engelhardt, Alfridijanta, McMullon, & Corley, 2017; Scott, Tetnowski, Flaitz, & Yaruss, 2014; Suh et al., 2014).

1.1.5 Cognitive Features of Disfluency in Autism

Disfluencies (in both autistic and neurotypical individuals) are highly associated with various language formulation processes (Yaruss, Newman, & Flora, 1999). For example, Perkins, Kent, and Curlee (1991) hypothesised that disfluencies occur when various aspects of the language formulation processes are not sequenced correctly prior to speech production. On the other hand, Postma and Kolk's (1993) Covert Repair Hypothesis asserts that speech disfluencies are by-products of normal repair processes when speakers detect errors. Currently the mechanism by which disfluencies manifest in individuals with autism is unclear, however, examining existing cognitive theories in relation to pragmatic language may help further understand this phenomenon.

1.1.5.1 Social Inference Theory

Pragmatic inferences are required in social situations when one attempts to explain or predict others' thoughts, intentions, and behaviours. Theory of mind (the ability to attribute mental states to others and to have an understanding that others have beliefs, intentions and perspectives that are different from one's own) is considered to be an important component in pragmatic inference skills (Martin & McDonald, 2003). Different types of pragmatic inferences have been defined in the literature, and a number of different terms have been proposed to reflect the extent to which pragmatic inferences require theory of mind skills. For the purpose of this review, two terms – namely linguistic pragmatics and social pragmatics – proposed by Andrés-Roqueta and Katsos (2017) will be used. Linguistic pragmatics refer to the skills where structural language skills (e.g. grammar, syntax) and competence with pragmatic norms (i.e. interpretation from the context without the need to infer from others' mental states) are sufficient for individuals to perform successfully in the conversational context. In addition, social pragmatics encompasses both linguistic pragmatic skills and competence with theory of mind. Baron-Cohen (1995) originally suggested that autistic individuals have difficulties with theory of mind skills, and they tend to have weaker inference abilities than the neurotypical controls (see Loukusa & Moilanen (2009) for a review). Clark & Fox Tree (2002) argued that some disfluencies, such as filled pauses ('um' and 'uh'), are listener- oriented, meaning they are produced by the speakers as a signal of an imminent delay in speech. For example, uh is a signal of an upcoming short halt whilst um signals of an upcoming long delay. Due to difficulties with theory of mind, it is hypothesised that autistic individuals are less likely to attend to the needs of the listener during dialogues and therefore may produce fewer listener-oriented disfluencies (Andrés-Roqueta & Katsos, 2017; Engelhardt, Alfridijanta, McMullon, et al., 2017).

1.1.5.2 Executive Functioning

Executive functioning difficulties are highly prevalent in autistic individuals, which may partly explain differences in disfluency use. Executive function skills are a set of processes that are necessary for the cognitive control of behaviour: (1) holding and manipulating information in our brain over short periods of time, (2) selecting and monitoring behaviours that facilitate the attainment of chosen goals (Ozonoff, Pennington, & Rogers, 1991). Although the direction of causality is unclear in current literature, executive functions are positively correlated with pragmatic language skills (see Matthews, Biney, & Abbot-Smith, 2018 for a review). Previous research suggests that autistic individuals score lower than neurotypical controls on measures of working memory, cognitive flexibility, self- monitoring, and planning and organising skills (Kercood, Grskovic, Banda, & Begeske, 2014). Working memory refers to the skill of holding and manipulating information in one's mind, which is important during fluent conversations (Baddeley, 1992). Scott (2015) stated that the higher number of repairs produced during conversations in autistic individuals might be a result of difficulties with rapid retrieval of information. In addition, several studies found that increased cognitive load attenuates speakers' ability to plan and monitor their speech, resulting in an increase in their disfluency rates, and this may be particularly prevalent in the autistic population (e.g. Bortfeld et al., 2001; Metz & James, 2018).

1.1.6 Current Research Evidence of Disfluency in Autism

Appraising the research evidence on disfluencies is challenging due to the lack of formal definitions of different disfluency types and different methodologies employed in the studies. For example, one type of disfluency defined as the part of speech that is stopped in the middle of an utterance and revised has been called 'self- correction' in one study (e.g. Suh et al., 2014) whilst being defined as 'revision' in another study (MacFarlane et al., 2017), despite being of the same linguistic nature. Engelhardt et al. (2017) attempted to remedy this by grouping previous studies of disfluencies in autism into four categories; filled pauses, unfilled pauses, repetition, and repairs/revisions. Pauses are disfluencies that occur during natural conversation and can either be unfilled (silent pauses) or be filled with utterances such as 'um', 'uh', 'like', and 'I mean' (filled pauses). Filled pauses are short utterances widely used but often overlooked in pragmatic language which include discourse markers (e.g., you know, I mean) and filled pauses such as um, uh, like (Brennan & Williams, 1995). Repairs and revisions, also known as non-stuttering like disfluencies, often occur when one loses one's train of thought in conversation due to poor working memory skills (Clark & Fox Tree, 2002). Previous studies found that autistic individuals produced a greater amount of non- stuttering-like disfluencies than the neurotypical controls (e.g. Lake, Johanna K.; Humphreys, Karin R.; Cardy, 2010; Scott, 2015). To the author's knowledge, only Engelhardt et al. (2017) provided a summary of studies on disfluencies in autism. However, these findings are likely to be unrepresentative of the body of literature in this area as the authors did not carry out a systematic search on the current literature, nor did they provide inclusion and exclusion criteria for the studies. It should be noted, however, that the authors only set out to provide a summary of a number of studies, as opposed to, a systematic review on disfluencies.

1.1.7 The Present Study

To date, it has not been possible to identify reliable biomarkers for the identification and diagnosis of autism. As a result, autism is diagnosed at the behavioural level (Lai et al., 2015). As discussed, there is a high prevalence of language impairment in autistic individuals as well as significant overlaps in the behavioural presentation of individuals with autism and language impairments (Bishop, 2010), which makes it particularly difficult to differentiate individuals with autism and language impairments. Preliminary evidence suggests that disfluency use may be able to distinguish autistic individuals from those with language impairments, as well as neurotypical individuals. If this were the case, this may shed light on different underlying cognitive mechanisms (e.g., working memory, lexical processing) and thus inform interventions to support children and young people. For example, given the same behavioural presentation of pragmatic language difficulties, autistic children and young people may benefit from interventions that focus on social cognition whilst those who have language impairments may benefit from interventions that focus on grammatical skills (Creemers & Schaeffer, 2016; Luyster, Seery, Talbott, & Tager-Flusberg, 2011). In addition, disfluencies could potentially be used as relatively efficient and objective markers for autism and thus facilitate early identification for children with social communication difficulties, in order to benefit from early interventions to achieve best possible outcomes (Fricke et al., 2017; Pickles, Anderson, & Lord, 2014).

Accordingly, the present systematic review aims to answer the following research questions, following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines:

Is speech disfluency a marker for autism? In other words, is it possible to differentiate individuals with and without autism based upon disfluency use?

1.2 Method

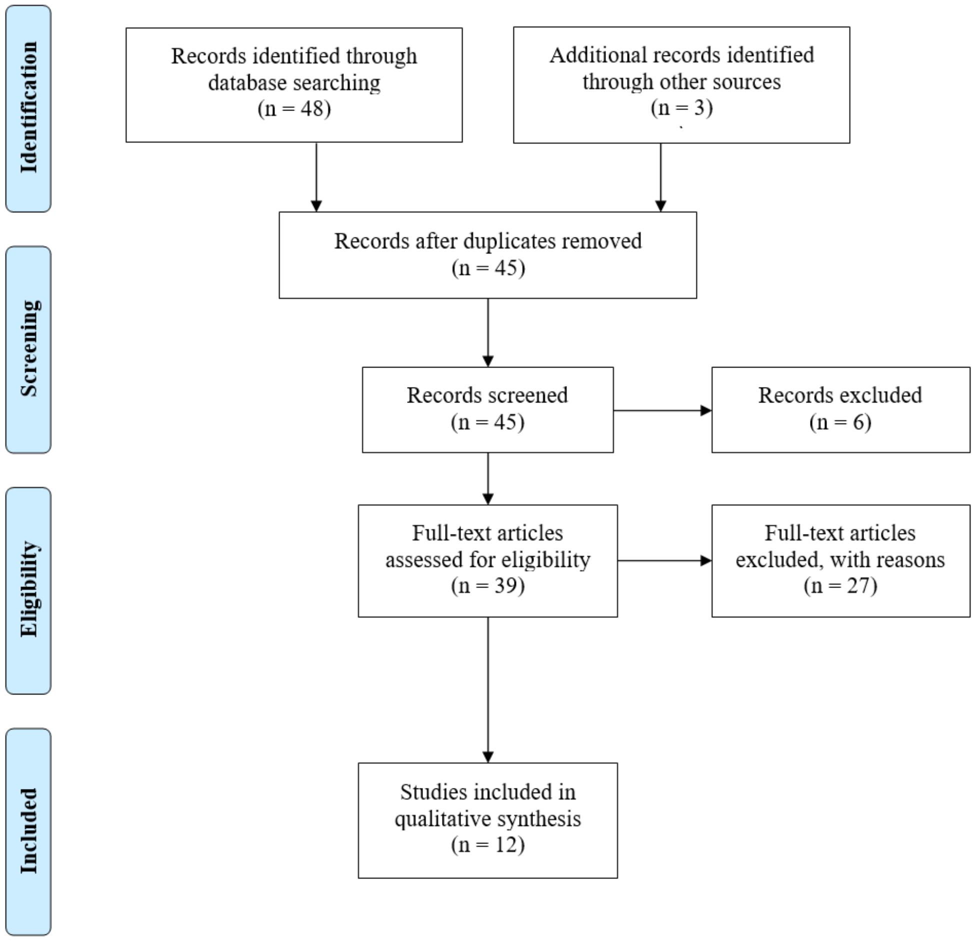
1.2.1 Search Strategy

This systematic review was conducted following the PRISMA guidelines (Moher, Liberati, Tetzlaff & Altman, 2009). A systematic literature search was undertaken using four databases, including PsychINFO, MEDLINE, PsychARTICLES, and CINAHL Plus. The initial search was performed on 16th September, 2018, updated on 10th January, 2019, and then 18th April, 2019 (Appendix A for details). The search terms used were ("filled pause*" or disfluenc* or "conversational filler* or "unfilled pause* or dysfluenc*) and (“autis* or asd or "autism spectrum disorder" or asperger* or "asperger's syndrome" or autistic disorder or asc or "autism spectrum condition" or "autistic spectrum condition "). There were no restrictions on the publication date or publication type. Additional search for articles and grey literature (e.g. unpublished dissertations, theses, and conference papers) was carried out using search engines, such as Google and Google Scholar. The reference lists in the papers identified from the above processes were screened to identify additional relevant articles. Specifically, any papers that appeared relevant (e.g. those that contained titles with disfluency) were screened for eligibility by reading the abstract. References were managed using EndNote reference manager software to remove duplicates. Experts (*N* = 2) in the field were also contacted and asked if they knew of any unpublished dataset appropriate for inclusion in the review.

1.2.2 Search Results

The initial systematic search yielded 48 papers that focused on the exploration of disfluencies in the autistic population (see Figure 1). A further three studies were identified from the reference lists in the papers. After removing duplications, the titles and abstracts of these (*N* = 45) studies were screened using Covidence (https://covidence.org/), an online software for evidence synthesis. According to the inclusion and exclusion criteria, 14 studies were identified as eligible for the systematic review after accessing the full-text articles; one dissertation study was not included because the full text was not available and the author did not respond to a request for access. Further full-text inspection revealed that two conference papers reported the same experimental findings and therefore only one (Lunsford, Heeman, Black & van Santen, 2010) was included and Heeman, Lunsford, Selfridge, Black & van Santen (2010) was excluded. This resulted in 12 included studies in the systematic review. Reasons for exclusion for each of the N = 27 papers is shown in appendix B.

Figure 1 PRISMA Flow Diagram



1.2.3 Reliability of Included Papers

Full-text of 12 papers were considered according to the following inclusion criteria: empirical study, quantification of disfluency types, samples including at least five autistic individuals, inclusion of at least one neurotypical comparison group, studies published in English or Chinese. There were no restrictions on the publication date or publication type. Qualitative studies, treatment studies and intervention studies were excluded as this review focuses on the group differences in disfluency production, instead of the effects of treatment and intervention associated with disfluency. Reliability was undertaken through a second researcher coding 50% of papers for each stage. Disagreements of eligibility were reconciled between researchers, and members of the supervisory team were consulted when required. Six or fewer discrepancies were found at each stage . All included studies were published in English.

1.2.4 Data Extraction

Data were extracted from N = 12 papers (see Table 1). This included sample size of the autism group and comparison group(s), age (range, mean and standard deviation if available) of the participants, disfluency types defined by Shriberg (1995): filled pauses, unfilled pauses, repetition, repair/ revision, results of the disfluency types (mean and standard deviation if available), category of tasks (monologue and dialogue). Corresponding authors of the articles were contacted if the data were not readily available.

Table 1. Sample Characteristics of Studies Included in the Review

Table 1. Sample characteristics of studies included in the review

Authors (date)	ASD Group				Comparison Group			
	Males (*n*)	Females (*n*)	Age		Males (*n*)	Females (*n*)	Age	
			M	*SD*			*M*	*SD*
Thurber and Tager-Flusberg, 1993	9	1	12.08		4	6	7.75	
					6*	4*	11.25	
Shriberg et al., 2001	15 HFA		21.6	10.8	53		26.25	12.7
	15 AS		20.7	10.9				
Lunsford et al., 2010	26 in total				22 in total			
Lake et al., 2011	13		27	10.9	13		Matched	
Suh et al., 2014	14 HFA	1 HFA	12.9	1.6	13	2	13	1.6
					12 OO	3 OO	12.4	1.8
Irvine et al., 2016	21	3	12.83	2.4	14	2	13;4	1.8
					19 OO	5 OO	13.6	4.3
Gorman et al., 2016	45	5	6.6	1.2	31	12	6.2	1.2
					11 SLI	7 SLI	7.1	1.1
de Marchena & Eigsti, 2016	17	1	14.1	1.5	15	3	15.4	1.5
Engelhardt et al., 2017	9 HFA	4	26.33	10.97	10	16	21.5	4.54
							19.7	0.84
MacFalane et al., 2017	47	6	6.7	1.1	32	13	6.8	1
					12 SLI	8 SLI	7.1	1
Parish- Morris et al., 2017	49	16	9.96	2.05	8	9	11.32	2.21
McGregor & Hadden., 2018	29	2	11		16	16	10.75	

Note : * Learning Disability; HFA: High Functioning Autism; AS: Asperger Syndrome; OO: Optimal Outcome; SLI: Specific Language Impairment
Engelhardt et al. (2017): 9 males matched and 1 unmatched on IQ; 4 females matched and 13 unmatched on IQ

Autism Group

The autism groups across the N=12 papers consisted of at least five participants with a diagnosis of Autism Spectrum Disorder or any subdomain of autism (e.g. Asperger's Syndrome, High-functioning Autism).

Comparison Group

The comparison group across papers consisted of at least five participants without a diagnosis of autism.

Filled pauses

A set of pauses during speech production that contain fillers such as 'um' and 'uh' (e.g. 'he uh likes it"; 'uh' is categorised as a filled pause), or discourse markers such as 'you know' and 'I mean '.

Unfilled pauses

An interval of pause during speech production that are not filled by fillers or discourse markers (e.g. ‘he [silent pause] likes it’; where the silent pause is categorised as an unfilled pause).

Repetition

Repeating one or more words during speech production without correction of previous utterance (e.g. ‘he liked liked it; where ‘liked’ is categorised as repetition).

Revision/ Repair

Repeating one or more words during speech production with correction of previous utterance (e.g. ‘she, he liked it’; where ‘she’ is categorised as revision/ repair).

Following methodology employed by Engelhardt, Alfridijanta, McMullon & Corley (2017), the experimental tasks employed in the research were categorised into two groups, namely monologue and dialogue, to facilitate comparison between studies.

Monologue studies

Monologue studies, as opposed to dialogue studies, are designed to elicit and control for specific verbal responses. For example, participants are required to produce a grammatical utterance by describing an event or a picture during speech production tasks.

Dialogue studies

Dialogue studies are designed to elicit verbal responses in a more interactive and naturalistic manner. For example, participants are required to engage in a conversation with researchers in a reciprocal manner.

1.2.5 Quality Assessment

Following data extraction, the unique studies were assessed systematically to judge the quality, such as reliability and validity, using the Appraisal tool for Cross-Sectional Studies (Downes, Brennan, Williams & Dean, 2016). A second researcher assessed 50% of the studies which were chosen randomly, to compare for consistency. The appraisal tool comprises 20 questions that are presented in the order that they should generally appear in a published article (i.e. Introduction, Methods, Results, Discussion and Other).

1.3 Results

1.3.1 Quality Assessment

The scores obtained from Appraisal tool for Cross-Sectional Studies ranged from 14 to 15 out of a possible score of 20 (see Appendix C) , indicating satisfactory methodological quality in the studies. Inter-rater agreement was 91%, which is excellent. The most common risks of bias stemmed from a relatively small autism group sample size across all of the studies (*M*= 28.50, *SD* = 17.29, range = 10-65) and the samples not necessarily being representative of the target population as they tended to be drawn from convenience samples. For example, autistic participants were recruited through local clinics and hospitals where they previously received their diagnoses. Although this is a relatively easy and cost-effect way to recruit participants, it is likely to lead to non-representative or biased samples. In addition, none of the articles reported the non-response rate, and it was therefore not possible to determine the levels of selection or non-response bias.

1.3.2 Sample Characteristics in the Studies

All the studies included an experimental group (consisting of participants with a diagnosis of autism), and a comparison group (consisting of neurotypical participants). One study (Thurber & Tager-Flusberg, 1993) included an additional comparison group (‘mildly mentally retarded’). Two studies (Irvine, Christina A; Eigsti, Inge-Marie; Fein, 2015; Suh et al., 2014) included an additional comparison group of ‘Optimal Outcome’ participants. Participants in the optimal outcome group previously met diagnostic criteria prior to age 5 years, but their symptom severity had improved such that they no longer met the criteria as per the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2002). Two studies (Gorman et al., 2016; MacFarlane et al., 2017) included an additional group of participants with Specific Language Impairment, a neurodevelopmental disorder characterised by language difficulties or delays in the absence of cognitive or developmental impairments (Bishop, 1992).

Most (*N* = 10) of the studies matched the autism group and comparison group (s) on age and verbal IQ (see Table 1); nine studies included participants aged below 18 while two studies included adult participants. One study included both adult and child participants with an age range of 10-29 years (Shriberg et al., 2001). Although the neurotypical and autism groups were matched on age, they did not test whether age moderated any of the effects found. The majority of the studies (*N* = 9) also included both male and female participants; nine studies included participants of both sexes. Two studies (Lake, Johanna K.; Humphreys, Karin R.; Cardy, 2010; Shriberg et al., 2001) included only male participants while one study (Lunsford, Heeman, Black, & van Santen, 2010) did not report sex/gender data. There was relatively low number of female autistic participants in all of the studies, ranging from one to six across *N* = 10 studies, compared to number of male autistic participants, ranging from 9-49 across *N* = 12 studies.

Table 2 Key Findings of Studies Included in the Review

Table 2 Key findings of studies included in the review

Authors (date)	Key Findings				Task
	Filled Pauses	Unfilled Pauses	Repetitions	Repairs/ Revision	
Thurber and Tager-Flusberg, 1993	N/A	ASD< NT	ns	ns	Monologue (Story-telling)
Shriberg et al., 2001	N/A	ASD>NT	ASD>NT	ASD>NT	Dialogue (ADOS interview)
Lunsford et al., 2010	ASD< NT				Dialogue (ADOS interview)
Lake et al., 2011	ASD<NT	ASD>NT	ASD> NT	ASD<NT	Dialogue
Suh et al., 2014	ns		HFA> NT OO = NT OO = HFA	HFA>NT OO>NT OO = HFA	
Irvine et al., 2016	ASD< NT	N/A	N/A	N/A	Monologue (Painting descriptions)
Gorman et al., 2016	Um ratio: ASD< NT	N/A	N/A	N/A	Monologue and Dialogue (ADOS tasks)
de Marchena & Eigsti, 2016	ASD> NT (Combined)			ASD> NT	Monologue (Narrative Task)
Engelhardt et al., 2017	N/A	ASD> NT	ns	ASD> NT	Monologue (Memorise and repeat a sentence)
MacFalane et al., 2017	ASD< NT		ns	ns	Monologue and Dialogue (ADOS tasks)
Parish- Morris et al., 2017	ASD< NT	N/A	N/A	N/A	Monologue and Dialogue (ADOS tasks)
McGregor & Hadden., 2018	ASD< NT	N/A	N/A	N/A	Monologue and Dialogue (ADOS tasks)

Note: NT: Neurotypical; HFA: High Functioning Autism; AS: Asperger Syndrome; OO: Optimal Outcome; SLI: Specific Language Impairment

Filled Pauses

Nine studies examined differences in filled pause use between autistic and neurotypical participants (see table 2). Overall, all but one study (Suh et al., 2014) found that autistic participants produced significantly fewer filled pauses (fillers and discourse markers) than neurotypical participants. Despite not reaching significance levels, Suh et al. (2014) also found that autistic participants used fewer filled pauses than neurotypical participants, which represented a small to medium effect ($\eta^2 = .07$).

Five studies investigated differential use of 'ums' and 'uhs'. Four (Gorman, Olson, Hill, Lunsford, Heeman, & van Santen, 2016; MacFarlane et al., 2017; McGregor & Hadden, 2018; Parish-Morris et al., 2017) found that autistic participants had a significantly smaller 'um' ratio (um/(um+uh)), which was mainly being driven by lower levels of 'um' in autistic participants. For the remaining study, Irvine et al (2016) found that autistic participants used 'um' significantly less frequently than the neurotypical and an 'optimal outcome' groups, and all groups used similar levels of 'uh'.

In summary, these results suggest that overall individuals with autism used fewer filled pauses than neurotypical individuals. In addition, autistic individuals had a smaller 'um' ratio, mainly being driven by higher levels of 'um' in neurotypical participants, which may make them sound less linguistically sophisticated, .

Unfilled pauses

In total, four studies examined differences in unfilled pauses between autistic and neurotypical participants (see table 2). Three studies (Engelhardt et al., 2017; Lake et al, 2011; Shriberg et al., 2001) found that the autistic group produced significantly more unfilled pauses than neurotypical group whilst one study (Thurber and Tager-Flusberg, 1993) found the opposite pattern.

Repetition

In total, six studies examined differences in repetitions during speech productions between autistic and neurotypical participants (see table 2). Three studies (Lake et al., 2011; Shriberg et al., 2001; Suh et al., 2014) found that the autism group produced significantly more repetitions than the neurotypical group, whilst three studies (Engelhardt et al., 2017; MacFarlane et al., 2017; Thurber & Tager-Flusberg, 1993) reported no group differences. However, the paired comparison between the autism group and unmatched controls was marginally significant $t(24) = 1.92$, $p = .07$. (Engelhardt et al., 2017). Furthermore, MacFarlane et al. (2017) found that the autism group, produced more repetitions ($M = 27.97$) compared to the neurotypical ($M = 20$) and Specific Language Impairment group ($M = 19$). However, neither test statistics nor standard deviations were reported, therefore it was not possible to calculate effect sizes. Furthermore, Thurber & Tager-Flusberg (1993) found that autistic participants used non-significantly more repetitions than neurotypical participants, which represented a medium effect ($d = 0.49$).

Repairs/ Revisions

In total, six studies examined differences in repairs/ revisions between autistic and neurotypical participants (see table 2). Four studies (De Marchena & Eigsti, 2016; Engelhardt et al., 2017; Shriberg et al., 2001; Suh et al., 2014) found that autism groups produced significantly more repairs/ revision than the neurotypical group whilst two studies (MacFarlane et al., 2017; Thurber & Tager-Flusberg, 1993) reported no group differences. Although MacFarlane et al. (2017) found non-significant results, they found that the autism group produced fewer repetitions ($M = 14.79$) compared to the neurotypical ($M = 15.71$), but more than Specific Language Impairment group ($M = 11.22$). Again, neither the statistics nor the standard deviations were reported; therefore, it was not possible to calculate the effect size.

1.4 Discussion

The aim of this systematic review was to understand the pattern of disfluency use in speech in groups of autistic speakers relative to neurotypical controls, and to ascertain whether disfluency could be a marker for autism. To address this aim, the following four disfluencies were investigated: (1) Filled Pauses (2) Unfilled Pauses (3) Repetitions and (4) Repairs/Revisions. Seven out of eight studies found that filled pauses were produced significantly less frequently by autistic individuals relative to neurotypical controls. Three out of four studies found that significantly more unfilled pauses were produced by the autistic group relative to the neurotypical group. For repetitions, three studies found that autism group produced significantly higher amounts of repetitions than the neurotypical group, whilst three studies yielded non-significant results. One study combined filled, unfilled pauses and repairs, finding that overall, these disfluencies were used significantly more often in autistic participants' speech. For repairs, four studies found that autism group produced significantly more repairs/revisions than the neurotypical group whilst two studies yielded non-significant group results. In summary, the pattern of results suggest that there may be meaningful differences in the production of disfluencies between autistic and neurotypical individuals. Despite this, such differences are dependent upon the specific type of disfluencies in question.

Although there appeared to be consistent findings among the studies for each type of disfluency, there are a number of methodological and conceptual issues to overcome before meaningful conclusions can be drawn and generalised. These issues include (1) potentially unaccounted for and unmeasured moderating and mediating variables, (2) methodological differences between studies, and (3) the heterogeneity of autism. Firstly, there are a number of potentially moderating and mediating variables for which were unaccounted but could influence the production in disfluencies. To illustrate this, research suggests that sex may influence the extent to which individuals produce filled pauses in the general population (Acton, 2011). Specifically, numerous studies have robustly found that males produce fewer filled pauses than females (Bortfeld et al., 2001; Laserna et al., 2014; Tottie, 2011). Many of the reviewed studies included more males relative to females in the autism than the neurotypical group. As such, the findings of lowered filled pause rates in the autistic groups may simply reflect sex-imbalanced groups, as opposed to an actual difference between autistic and neurotypical groups per se. In addition, disfluencies form a part of pragmatic language; it is therefore plausible that linguistic ability could be a potential confounding variable. A number of the reviewed studies did not match the participants on their linguistic abilities and these differences were not controlled for in their statistical analyses. For example, Irvine et al. (2016) reported significant differences in participants' linguistic abilities (p = .001) where autistic participants scored significantly lower, as measured by the Comprehensive Evaluation of Language Fundamentals (CELF), than the comparison groups (neurotypical and 'Optimal Outcome'). Similarly, McGregor and Hadden (2018) reported significantly lower scores obtained in the CELF by autistic than neurotypical participants (p = .0005). In addition, Gorman et al. (2016) reported that autistic participants performed significantly lower in the Verbal IQ subtest using the Wechsler scales test (p< .001), and in the CELF (p = .006), than the comparison groups (specific language impairment and neurotypical). They also reported that half of the autistic participants (*n* = 25) could be identified as language impaired according to the CELF score. Further, although the majority of included studies (*N* = 10) matched the autism and comparison group on age, the different studies included participants from a wide age range (6-50 years), which could partly account for some of the discrepant findings. Indeed previous research suggest that older adults are more likely than young adults and children to produce disfluencies (e.g., unfilled pauses mid-sentence) during naturalistic speech (Bortfeld et al., 2001; Shafto & Tyler, 2014), suggesting there may be developmental differences in disfluency use. In summary, the differences in disfluency use may be driven by these factors; as opposed to disfluencies representing a phenotypic manifestation of pragmatic language difficulties in individuals on the autism spectrum per se. Future research should directly investigate the effects of sex and language ability on disfluent speech.

Secondly, the reviewed studies differed considerably in terms of the methodologies and statistical analyses employed. To facilitate comparisons between studies in this review, disfluencies were categorised into four different types. However, within each disfluency, researchers employed different experimental tasks. For example, for filled pauses, Irvine et al. (2016) asked participants to create a narrative to describe twelve individual still paintings, whereas Lunsford et al. (2010) analysed data from Autism Diagnostic Observation Schedule (ADOS) semi-structured interviews. Similar differences in task-type were present for studies investigating the other three types of disfluencies. It is likely that different tasks involve distinct social and cognitive demands, and thus reflect different underlying cognitive processes (e.g., working memory, attention, lexical processing). It is likely that inconsistent findings within and between disfluency types can at least partly be explained by the very different tasks that were employed (e.g., monologue or dialogue). In addition, a limitation of this systematic review is that only two studies (de Marchena & Eigsti, 2016; Parish-Morris et al., 2017) reported means and standard deviations, meaning effect sizes could not be calculated which would have allowed for a more precise estimation of group differences. Given the heterogeneous nature of autism (which will be discussed in detail below) and methodological differences between studies, it is important for future research to employ similar tasks, as well as build upon previous experimental paradigms, to investigate whether there are differences between autistic and neurotypical participants in their disfluency production.

Finally, despite sharing underlying core features (social communication difficulties, restrictive and repetitive behaviours, and sensory differences); there is a wide degree of heterogeneity within the autistic population. The concept of heterogeneity applies to both within and between individuals and at multiple levels of analysis (e.g., behavioural, and/or cognitive levels). In addition, heterogeneity may be found in individual developmental progression (Lombardo, Lai, & Baron-Cohen, 2019). In the reviewed studies, eight out of twelve studies had small sample sizes (n < 30 in the autism group), rendering it to be difficult – if not impossible – to control for the large heterogeneity amongst participants in terms of their clinical features (e.g. executive functioning skills, verbal IQ, 'ASD symptom severity', language ability). Further, three reviewed studies included participants with a wide age range - the age range was 10 to 49 in Shriberg et al.'s (2001) study, and 19-35 in Lake et al.'s (2010) study. Engelhardt et al. (2017) did not report the age range for the participants but the standard deviation of 10.7 in the autistic group would suggest that they had a wide age range. In addition, it is possible that differences in disfluency use arise during the developmental course. However, the reviewed studies included a mixture of child and adult participants, and did not investigate whether differences were found between participants of different ages. Further, it is important to note that all the autistic participants in the reviewed studies were recruited from clinics via convenience sampling, and therefore they arguably presented with a greater number and/or severity of autistic traits to have come to clinical attention. Future research should include participants who have high autistic traits but have not received a clinical diagnosis (Wiggins, Robins, Adamson, Bakeman, & Henrich, 2012). There is emerging evidence that females may 'camouflage' and mask their autistic features, making them less likely to have come to clinical attention (Constantino, 2017; Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2015) and therefore are under-represented in the clinical population. This is particularly important for at least one disfluency – filled pauses – as sex is a potentially significant confound. If no significant differences in types of disfluencies are found between these two groups (high traits without a clinical diagnosis and with diagnosis), but different than neurotypical controls, this would provide tentative evidence that certain types of disfluencies may represent a subtle yet phenotypic manifestation of autism. In addition, if differences are found between females with high traits and a clinical diagnosis, this may help researchers to better understand the camouflaging phenomenon.

1.4.1 Future Research

Future research should aim to overcome the limitations inherent in many of the reviewed studies. Specifically, future research should aim to account for variables that potentially moderate or mediate disfluency production, such as sex, age, language ability, IQ, working memory, and theory of mind skills. In addition, future research should account for the large amount of heterogeneity found in autistic individuals by matching the groups. Both of these points are in accordance with a recent review by Lombardo, Lai and Baron-Cohen (2019) who argued that autism research should make use of big data, which is both 'broad' (i.e., large *N*) and 'deep' (i.e., multiple levels of data should be collected from the same individual). In addition, future research should attempt to utilise similar experimental paradigms to determine whether results replicate, as well to ascertain if task-type affects disfluency production; this should also shed light on the underlying mechanisms involved in disfluency production.

This review raised the question of whether disfluencies can be used as a marker for autism. Taken together, this review reports preliminary evidence suggesting that there may be reliable differences in disfluency production between autistic and neurotypical groups, and therefore could potentially be taken into account when revising diagnostic assessments which use measures of pragmatic language (e.g., the ADOS). In essence, this means that there is a possibility that disfluencies could, at least in part, predict whether an individual is likely to have autism. This paradigm lends itself well to machine learning studies, where the focus would be on predicting, as opposed to explaining the casual mechanisms of autism (Yarkoni & Westfall, 2017). It is beyond the scope of this review to address machine learning approaches, however, machine learning approaches generally focus on learning statistical functions from multidimensional data set in order to make predictions about individuals. Machine learning should also help overcome the heterogeneity in autism, by predicting specific constructs (such as executive functioning and social reciprocity) – as opposed to overall diagnostic label – and/or biomarkers that transcend diagnostic boundaries (Dwyer, Falkai, & Koutsouleris, 2018).

1.4.2 Implications for Educational Psychology

Preliminary evidence suggests there may be differences between the way in which autistic and neurotypical individuals produce disfluencies and may therefore serve as a relatively objective way to identify individuals who have social communication difficulties. For example, educational psychologists could record conversations with children and young people and then analyse the data. In addition, disfluencies may impact upon the perceptions of listeners, for example by making the speaker sound less linguistically sophisticated. This could have potential negative implications, such as making it less likely for an individual to be accepted by the neurotypical peers. Therefore, it may be useful to teach individuals with social communication difficulties to use disfluency in a more neurotypical way, as this may help them achieve personally important outcomes, such as securing desired employment. Given that children and young people on the autism spectrum are at greater risk for social difficulties, social exclusion, and bullying at school, practitioners could be made aware of these differences with a view to normalising differences in speech production for all pupils (e.g. pupils with accents, EAL, speech production needs, etc).

Moreover, given the subtle nature of disfluency, it will also be important to raise awareness of how disfluency may affect perceptions of the speaker, in order to try to reduce any potential bias associated with this, as opposed to focusing on modifying individuals' communication style. In addition, Environmental adaptations should be made to ensure all individuals have equal opportunities, and EPs are well placed to be at the forefront of this. For example, research has found that individuals with autism are perceived negatively by their neurotypical peers interacting with them face-to-face; however, when the conversational content without any audio-visual cues was evaluated, the biases dissipated (Sasson et al., 2017). This suggests that the style, rather than the conversational content per se, may drive negative perceptions of peers. As such, it may be useful to allow autistic individuals to provide written answers to questions when being evaluated (e.g., oral examinations; job interviews), as opposed to this being done face-to-face. Any adaption, however, should be made at the request of the individual, to help them achieve their own personal aspirations, as opposed to being enforced upon them.

1.4.3 Conclusion

In conclusion, this systematic review has found that there may be meaningful differences in the production of disfluencies between autistic and neurotypical individuals. However, future research is needed to overcome the methodological limitations of the reviewed studies before firm conclusions can be drawn.

Chapter 2 Sex Differences in Filled Pauses in Children with Autism Spectrum Condition: On the Distribution of 'Um' and 'Uh'

2.1 Introduction

2.1.1 Background

Autism is a neurodevelopmental condition diagnosed at the behavioural level, which affects approximately 1-2% of the population. A wide degree of heterogeneity is found in autism on multi-levels - including behavioural (e.g., type and frequency of repetitive and stereotyped behaviours), and/or cognitive (e.g., varied executive functioning skills) levels (Lombardo et al., 2019). Autism has predominantly been diagnosed in males with a male to female ratio of 4 to 1 (Baird et al., 2006), however, a recent surge of research and literature suggest that females may 'camouflage' in social situations, resulting in under-representation of females in the clinical population (e.g. Constantino, 2017; Hull, Mandy, & Petrides, 2017). 'Camouflaging' is broadly defined as masking behaviours that might be deemed as socially unacceptable or engaging in social behaviours that are deemed to be more neurotypical (Attwood, 2007).

In the context of social communication, despite similar levels of underlying social communication difficulties, findings from research suggest that autistic females are more likely to engage in social mimicry (i.e., modifying one's outward social expression through imitating or copying others' social behaviours) than autistic males and thus autistic females may appear behaviourally more neurotypical (Lai et al., 2017). For example, Rynkiewicz et al. (2016) found that autistic girls performed better on a non-verbal mode of communication (as measured by use of gesture) than autistic boys. The authors concluded that the more 'vivid' gestures used by autistic girls results in them appearing behaviourally more neurotypical, thus demonstrating camouflaging. Further work has considered sex differences in social communication; specifically, studies have found sex differences in pragmatic language use in autistic population and autistic girls present more similarly to neurotypical girls than autistic boys do to neurotypical boys (e.g. Boorse et al., 2019; Grieve et al., 2016).

2.1.2 Filler Words

Filler words are short utterances widely used but often overlooked in pragmatic language which include discourse markers (e.g., you know, I mean) and filled pauses such as um, uh, like (Brennan & Williams, 1995). Discourse markers are often called other terms in the literature, such as pragmatic markers, discourse particles, discourse connectives, connecting adverbials (Heine, 2013). Discourse markers refer to short phrases, that are syntactically independent from their grammatical structure, have a non-restrictive meaning, and generally do not serve a grammatical purpose (Laserna et al., 2014). On the other hand, filled pauses are short utterances, as opposed to short phrases, commonly used in spontaneous speech. In this study, filled pauses – specifically Um and Uh – are examined.

2.1.3 Filled Pauses – Um and Uh

Although not pronounced identically, it is important to acknowledge variations of Um and Uh in the literature; such as the British spelling of 'erm' and Dutch spelling of 'øhm' to Um and British spelling of 'er' and Dutch spelling of 'øh' to Uh as they are sometimes reported as such in the studies (e.g. Navarretta, 2015; Rayson, Leech, & Hodges, 2010). As the majority of previous literature written in English reported 'Um' and 'Uh', this paper therefore follows these terminologies. Historically, filled pauses were regarded as 'noises' and were not researched due to their perceived extralinguistic nature. More recently researchers recognise that filled pauses have communicative value in speech and therefore should be investigated in their own right (Kjellmer, 2003). Previous studies have demonstrated that, rather than being random 'noises' in spoken language, filled pauses are influenced by a variety of cognitive, demographic and contextual factors (Kjellmer, 2003; Laserna et al., 2014). For example, greater use of filled pauses is typically associated with females (Acton, 2011); such sex differences are discussed in detail below.

Research suggests that filled pauses play either an unintentional role (i.e., reflecting speech planning difficulties) or an intentional role (i.e., a communicative function to signal pauses) in speech. In support of the view that the presence of filled pauses reflect planning difficulties, studies found that they are more likely to occur when the topic of the speech is unfamiliar (e.g. Bortfeld et al., 2001; Merlo & Mansur, 2004) and the tasks became more cognitively demanding (James et al., 2018; Metz & James, 2018).

Participants were also more likely to produce filled pauses when speech material is still being planned on a conceptual level and thus they often occur at the beginning of longer phrases (e.g. Fraundorf & Watson, 2014; O'Connell & Kowal, 2005; Navarretta, 2015). Whilst there is evidence to suggest that filled pauses may reflect planning difficulties, further research evidence also suggests that filled pauses do not necessarily have communicative value. For example, Finlayson & Corley (2012) found that filled pause use in participants did not differ between monologue and dialogue tasks, suggesting filled pauses are not signals produced intentionally by speakers. In contrast to the view that filled pauses play an unintentional role, Clark & Fox Tree (2002) analysed data derived primarily from the London-Lund corpus (Svartvik & Quirk, 1980) consisting 170,000 words from 50 face-to-face conversations, and asserted that the use of Um signals a speaker's intention to initiate a major delay whilst Uh signals minor delay. In addition, Walker, Risko, & Kingstone (2014) found that participants produced a significantly greater amount of filled pauses when responding to a human partner than a computer partner, suggesting filled pauses may play a functional role in communication. Interestingly, they also found that the presence of a human mediated the effect of filled pause use. In other words, participants produced more filled pauses when responding to a computer partner with the presence of a human researcher. Overall, findings from current research suggest that whilst filled pause use may indicate speech planning difficulties, they are also at least in part a product of a social use of language.

2.1.4 Sex Difference in Filled Pause Use in General Population

A number of research studies provided evidence that males are more disfluent, as indicated by higher proportions of filled pauses to total words, than females (Bortfeld et al., 2001; Shriberg, 1996; Tottie, 2011). However, Laserna et al. (2014) found that males and females produced similar amount of filled pauses. As discussed previously, Clark and Fox Tree (2002) asserted that Um and Uh may be pragmatically distinct, a number of studies have subsequently investigated Um and Uh separately. Higher ratio of Um to Uh produced by females than males was consistently found across experimental studies (e.g. Bortfeld et al., 2001; Laserna et al., 2014), and corpus-based studies (e.g. Acton, 2011; Tottie, 2011) which included a large amount of data from different large datasets across different settings. For example, Acton (2011) analysed data from the Speed Dating Corpus, where audio recordings from speed-dating sessions held for graduate students at a private American university were transcribed to 992 written transcripts with over 750,000 words. In addition, Tottie (2011) analysed data derived primarily from the London-Lund corpus (Svartvik & Quirk, 1980) consisting 170,000 words from 50 face-to-face conversations. In summary, a higher Um to Uh ratio has been consistently found in females than males. In addition, age was also found to be positively correlated with disfluency (Shafto & Tyler, 2014).

2.1.5 Filled Pause Use in Autistic Population

At least half of the autistic population demonstrates quantitatively (e.g., higher amount of filled pauses) and qualitatively (e.g., atypical tone) different speech patterns compared with the neurotypical population (Fusaroli, Lambrechts, Bang, Bowler, & Gaigg, 2017; Shriberg et al., 2001). There is also emerging evidence that the proportion of filled pauses to total words are produced differently by autistic and neurotypical individuals (Lake, Humphreys, Cardy, 2010; MacFarlane et al., 2017). Specifically, a number of studies found that autistic participants produced significantly fewer Ums to total words than neurotypical participants (Gorman et al., 2016; Heeman, Lunsford, Selfridge, Black, & van Santen, 2010; Irvine, Eigsti, & Fein, 2015). In addition, Irvine et al. (2015) and Lunsford et al. (2010) found that autistic participants and neurotypical participants did not differ in their rate of Uh to total words. Further, a number of studies found that autistic participants produced significantly lower Um ratios (total Um relative to total amount of Um and Uh) than neurotypical participants (Gorman et al., 2016; Lunsford et al., 2010; McGregor & Hadden, 2018). Despite this, Suh et al. (2014) found no significant difference in either filler use between autistic and neurotypical groups. Parish-Morris et al. (2017) found that autistic girls had higher Um ratios than autistic boys, reflecting a lower level of Uh use in autistic females, and they argued this serves as a 'linguistic camouflage' (p.8) as greater amount of Um relative to Uh use are often associated with superficial linguistic competence. The current study aims to replicate and build upon that of Parish-Morris et al. (2017), exploring whether there is a sex by diagnosis interaction for filled pause use. In the present study, one quantifiable feature of pragmatic language – disfluency – is examined. Specifically, the use of Um and Uh (filled pauses in disfluency) are examined to investigate the camouflaging hypothesis.

2.1.6 Current Study

The current study employed two tasks (one structured and one unstructured) to explore the use of filled pauses (Um, Uh) in autistic children and adolescents, compared with a neurotypical group, and between sex. Task one (structured) involved children and adolescents describing their personal experience of a drawing task they had completed. They were asked to describe the task in terms of their own perspective, and then from the perspective of a friend or family member. In task two (unstructured), children and adolescents were asked to create a narrative about a favourite game or hobby. The two tasks were each used to generate three indices of filled pauses including (1) the number of Ums (as a proportion of total words), (2) the number of Uhs (as a proportion of total words) and (3) the Um ratio (Um/Um+Uh).

Although autistic individuals have previously been found to produce fewer filled pauses than neurotypical individuals (De Marchena & Eigsti, 2016; Irvine, Eigsti, Fein, 2015; Lunsford et al.,2012), this finding may simply reflect sex differences rather than a phenotypic expression of autism, particularly considering previous research has tended to include many more males in the autistic, but not neurotypical group. Due to this methodological limitation (e.g., sex-imbalanced groups) in previous research it was not possible to hypothesise the direction of the effect for autistic individuals with respect to the use of Ums and Uhs in speech. We did, however, anticipate that autistic participants would use filled pauses differently in both tasks compared to neurotypical participants due to differences in executive functioning and social inference skills (as outlined in Chapter 1). We hypothesised that this difference would be most evident when considering Ums in the context of the unstructured task, due to the increased demands placed on executive functioning skills. Building upon Parish-Morris et al.'s (2017) study on linguistic camouflaging, we predicted that autistic females would have more similar Um ratio to neurotypical females than autistic males to neurotypical males, thereby normalising the way they sound in speech. In accordance with previous research, we anticipated that females would produce more Ums than males (MacFarlane et al., 2017; McGregor & Hadden, 2018; Parish-Morris et al., 2017), and males would produce more Uhs (Parish-Morris et al., 2017).

2.2 Method

2.2.1 Power analysis

No previous study has tested for an interaction between sex and diagnosis in terms of filled pause production, making it difficult to estimate an expected effect size. Therefore a power analysis was therefore conducted based upon the effect size difference (f = 0.37) between autistic and neurotypical participants for Um to total words from Parish-Morris et al. (2017), using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). This analysis demonstrated that a sample of 84 participants was needed to achieve 80% power when testing for interaction effects using ANCOVA.

2.2.2 Design

A 2x2 between subject design was employed comprising the following groups: 2 Sex (girls; boys) and 2 Group (autism; neurotypical). The dependent variables were proportion of Um to total words (Um/total words), proportion of Uh to total words (Uh/total words), and Um ratio (Um/Um+Uh).

2.2.3 Participants

Eighty-four children and adolescents aged 8-13 participated in this study, they were recruited from ten mainstream primary and two mainstream secondary schools in the South of England. Originally, 16 Special Educational Needs Coordinators (SENCOs) and/or headteachers from mainstream primary schools and three from mainstream secondary schools were approached by one of two researchers. Four SENCOs did not agree to participate. The remaining 12 SENCOs (or headteachers) from participating schools wrote to parents of all eligible children (aged 8-14) or approached specific parents/carers of children with autism or high autistic traits with brief details about the research.

The participants were split into two groups: autism and neurotypical; the autism group included children with a clinical diagnosis and children who scored higher than the cut-off point of nine on the Social and Communication Disorders Checklist (SCDC; Skuse, Mandy, & Scourfield, 2005), indicating high autistic traits.

The autism group (n = 40) included children with a clinical diagnosis of Autism Spectrum Disorder (n = 16) and Asperger's Syndrome (n = 2) confirmed by paediatrician. The remaining participants (n = 28) in the autism group were either in the process of assessment for autism diagnosis or included adolescents where significant concerns had been raised by school or parent about possible autism. The neurotypical group (n = 44) included children with no diagnosis and low traits according to the SCDC. There was no difference between boys who had a clinical diagnosis and those who had high traits without a diagnosis in respect to the three outcome measures of total use of Um to total words (p = .460), total use of Uh to total words (p = .323), and Um ratio (p = .831). Similarly, there was also no difference between girls who had a clinical diagnosis and those who had high traits without a diagnosis in respect to the three outcome measures of total use of Um to total words (p = .839), total use of Uh to total words (p = .319), and Um ratio (p = .178).

The final sample in the autism group comprised 22 boys and 18 girls (see Table 1), whilst the neurotypical group comprised 22 boys and 22 girls. Preliminary tests were run to test for pre-existing differences in verbal IQ, nonverbal IQ, full-scale IQ and age, using 2 X 2 between-subject ANOVAs with the following groups: Sex (females, males) and Group (autism, neurotypical). All main effects and interactions were non-significant (all ps > .116), except for verbal IQ (p = .012), which reflected lower verbal IQ scores for autistic participants than neurotypical participants.

2.2.4 Measures

2.2.4.1 Autistic traits.

The SCDC is a 12-item parent report scale designed to measure autistic traits. Parents rated each item on the scale as 'not true', 'quite or somewhat true', or 'very or often true', based upon how often their child's behaviour has been noticed by the parent over the past six months. The score range is 0-24 and a score of 9 or above suggest that the children are at higher risk of being on the autism spectrum. The SCDC was standardised in a sample of children with autism (n = 208) who were assessed using the Autism Diagnostic Interview – Revised (Lord et al, 1994) and the 3di (Skuse et al, 2004).

Validity of the SCDC was assessed by a comparison of mean scores with the 'clinical control group' (*n* = 76), and 'normal control group' (*n* = 118). Clinical control group included children diagnosed with a range of clinical disorders such as attention-deficit hyperactivity disorder, obsessive-compulsive disorder, Tourette syndrome, and conduct disorder. Discriminant validity was confirmed through comparing children with autism and the control groups. The SCDC has been used widely in published research to measure autistic traits due to increasing evidence supporting the notion that autism is a dimensional condition. (e.g. Ring et al., 2008; Wiggins et al., 2012). The SCDC has very high internal consistency (0.93) and sensitivity (0.90) as a screening instrument for autistic traits in the general population (Skuse et al., 2005).

2.2.4.2 Intelligence Quotient.

The Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II; Weschler, 2011) was employed to estimate participants' general intelligence. The word definition subtest was used to estimate verbal IQ whilst the matrix subtest was used to estimate performance IQ. Both scales provide a standardised score (*M* =100, *SD* = 15). McCrimmon and Smith (2014) noted that the WASI-II has good-to-excellent internal consistency, good-to-excellent test-retest reliability for the composites, and excellent inter-rater reliability. In addition, they note strong factor validity and acceptable-to-excellent concurrent validity.

2.2.4.3 Filled Pauses.

Structured Task

Each participant completed a semi-structured interview, they were first asked about an Interactive Drawing Task (IDT; Backer van Ommeren, Koot, Scheeren, & Begeer, 2017) that they previously engaged with the researcher. The IDT involved participants taking turns with a researcher to contribute to a drawing. No verbal instructions were given except 'We are going to do a drawing together' at the beginning of the task. The aim of this task was to measure social reciprocity (collected data is used in another researcher's thesis and not reported here).

Second, the participant was asked, 'Tell me about a friend at school.' If the participant indicated that they did not want to speak about a friend, the participant spoke about a family member instead. Third, the participant was asked about how did they think their friend/ family member would think about the task.

Unstructured Task

The participants were then asked to speak about their favourite hobby or game, using prompts developed by Nippold et al. (2005) with a purpose to elicit language sample in research - in particular expository discourse (a discourse that explains or describes a topic). Each participant was required to speak for a maximum of five minutes in total.

2.2.4.4 Transcription and Reliability

In the first instance, each audio file was individually transcribed by one of four undergraduate voluntary research assistants (VRAs) or one of two researchers (BW or HW). The VRAs were asked to read relevant background literature and received trainings on transcribing and coding. In addition, all VRAs signed a confidentiality agreement (see Appendix D). All words and fillers were transcribed, but partial words were not included. Next, BW and HW transcribed an additional 27% (n =16) of the audio files in order to compare for reliability. Three variables (Um count, Uh count, total word count) were tallied across the transcriptions, a one-way random intraclass correlations (ICC) were used to calculate interrater reliability: ICC (1,2) = .97 for Um, ICC (1,2) = .93 for Uh, and ICC (1,2) = .95 for total words, suggesting excellent reliability.

2.2.5 Procedure

After obtaining ethical approval from the University of Southampton Ethics Committee, and Research Governance, head teachers and/or SENCOs of primary schools were approached with research information. Once written confirmation to take part in this study was obtained from the school, the school sent an information sheet detailing the purpose of the study, consent forms and SCDC (see Appendix E) to the parent/ carers of eligible children or to specific parents of children with autism or high autistic traits. Parents who expressed an interest in taking part in the study were sent a pack of paperwork: an information sheet (see Appendix F), a consent form (see Appendix G), and the SCDC. Parents sent completed consent form and SCDC to SENCOs and were collected by the researchers.

All participants took part in the study in a quiet room in their school with a window. Prior to commencement of the study, participants were given an adapted version of information sheet with simplified language to read (see Appendix H), explaining the purpose and the process of the study. For those who were not proficient readers, researchers read the information sheet aloud to them. Additional verbal assent was gained from children and they signed the consent forms (see Appendix I) if they were happy to take part. The participants were also reminded that they could terminate the study at any time if they wished, and the data collected during the study was confidential. Half of the participants (*n* =42) completed the research with BW and half completed with HW.

All participants completed the study in the following order: (1) Interactive Drawing Test, (2) Reading the Mind in the Eye Test (all participants completed this but data was not analysed in this study), (3) WASI-II. All participants then completed an interview in which they were asked about the IDT, what a friend or family member would think of the task, and their favourite game or hobby (for interview schedule, see Appendix I). The interviews were recorded on a digital voice recorder. Each session lasted approximately 45-60 minutes per participant. All participants were debriefed, and reminded that they could contact their parent or teacher if they changed their mind and wished for their data to be withdrawn from the study.

2.2.6 Statistical Analyses

In terms of basic task performance, we analysed structured and unstructured tasks separately. For each task type, we considered the three different indices related to participants' use of Ums and Uhs using a univariate ANCOVA with two Group (autistic, neurotypical) and two Sex (male, female) as predictors and verbal IQ as a covariate (as autistic participants had significantly lower verbal IQ than neurotypical participants). As this analysis will yield both main and interaction effects, this will allow us to disentangle the effects of sex and diagnosis on filled pauses.

2.3 Result

2.3.1 Distribution of Data

All distributions were skewed, except for Um ratio in the unstructured task for autistic males. However, ANOVA is relatively robust to violations of normality. For example, Blanca, Alarcón, Arnau, Bono, and Bendayan (2017) conducted a systematic investigation of the robustness of *F* tests to violations of normality, finding it to be robust to type 1 error in 100% of the case studies. Accordingly, we decided to proceed with the analysis without transforming the data.

2.3.2 Descriptive Statistics

Table 1 presents descriptive statistics for the main outcome measures in the tasks (the number of total words, use of Ums and Uhs in speech and their relative relationship to total words and with each other). A 2x2 ANOVA was conducted, with two Group (autistic, neurotypical) and two Sex (male, female) as predictors, in order to test whether there were any differences for total words produced. For the structured task, there was no significant main effect of group (p = .727), sex (p = .066), nor the interaction between the two (p = .611). For the unstructured task, there was no main effect of group (p = .083), a significant main effect of sex (p = .046), and no significant interaction between the two (p = .813). Accordingly, we controlled for total words produced in the following analyses.

Exploratory correlations were conducted between age and IQ and with all outcome measures. For the structured task, all correlations were non-significant (all *p*s > .22) except for age and proportion of Uh to total words, r = .29, p = .01, indicating that the number of Uhs increased with age. For the unstructured task, all correlations were non-significant (all *p*s > .081) except age and Uh to total words, r = .35, p < .01 and between IQ and Um (to uh) ratio, r = .31, p < .01.

Table 3 Descriptives for autistic and neurotypical participants

Table 3. Descriptives for autistic and neurotypical boys and girls.

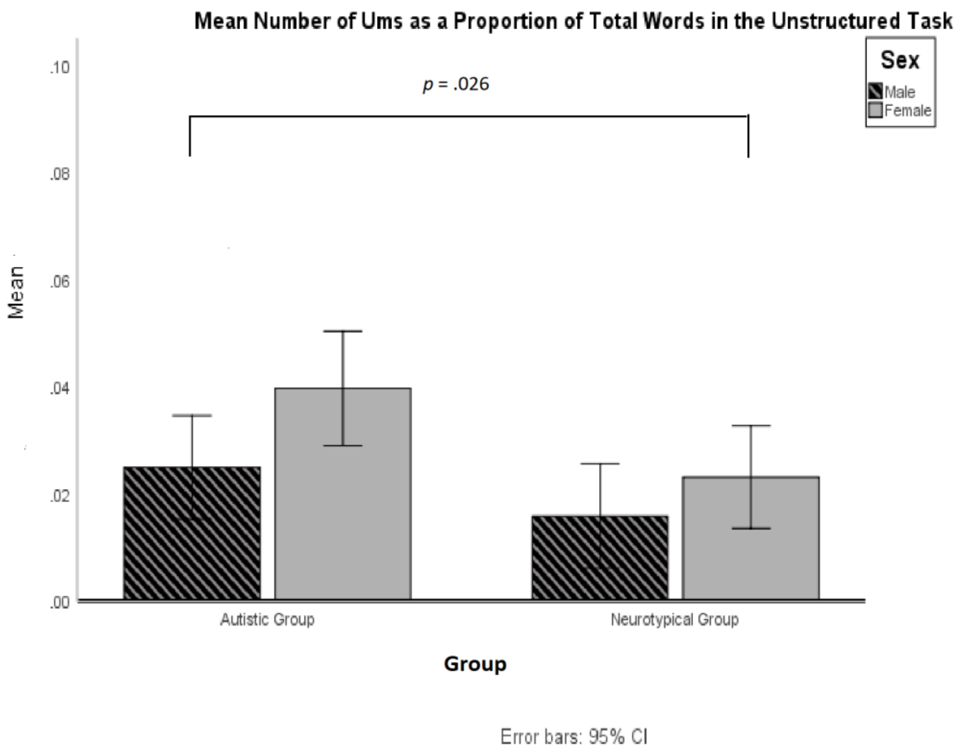
	Autistic (n = 40)				Neurotypical (n = 44)			
	Boys (n = 22)		Girls (n = 18)		Boys (n = 22)		Girls (n = 22)	
	M (SD)	Range	*M (SD)*	Range	*M (SD)*	Range	*M (SD)*	Range
Structured Task								
Total Words	217.046 (133.02)	51-681	163.778 (94.27)	57-418	197.772 (109.69)	63-542	167.364 (54.25)	94-284
Total Um	7.636 (7.62)	0-28	6.056 (6.27)	0-26	5.546 (4.47)	0-17	5.273 (4.87)	1-22
Total Uh	2.762 (5.12)	0-22	1.529 (1.84)	0-6	2.818 (3.33)	0-14	1.136 (1.91)	0-8
Total Um to Total Words	0.039 (0.04)	0-0.14	0.039 (0.03)	0-0.09	0.030 (0.02)	0-0.10	0.032 (0.03)	0.004- 0.13
Total Uh to Total Words	0.008 (0.02)	0-0.10	0.006 (0.01)	0-0.03	0.007 (0.01)	0-0.04	0.004 (0.01)	0-0.23
Um Ratio	0.691 (0.38)		0.731 (0.32)		0.658 (0.36)		0.819 (0.26)	
Unstructured Task								
Total Words	289.046 (138.14)	110-670	215.778 (88.70)	90-438	370.682 (312.99)	127-1670	277.955 (97.30)	85-528
Total Um	6.864 (7.01)	0-21	7.222 (5.02)	0-16	6.046 (6.09)	0-28	5.955 (5.19)	0-18
Total Uh	3.455 (5.07)	0-22	0.889 (1.88)	0-8	3.273 (4.90)	0-16	0.909 (1.34)	0-4
Total Um to Total Words	0.024 (0.02)	0-0.08	0.039 (0.03)	0-0.11	0.017 (0.01)	0-0.04	0.024 (0.02)	0-0.07
Total Uh to Total Words	0.013 (0.02)	0-0.11	0.006 (0.01)	0-0.04	0.011 (0.02)	0-0.06	0.003 (0.004)	0-0.02
Um Ratio	0.569 (0.37)		0.809 (0.34)		0.653 (0.38)		0.724 (0.39)	

2.3.3 Um to Total Words:

For the structured task, there was no significant main effect of group or sex, and the interaction between these variables was not significant (all *F*s < 2 and *p*s > .1).

For the unstructured task, there was a significant main effect of group, $F(1,78) = 5.16$, $p = .026$, $\eta_p^2 = .044$, and a non-significant main effect of sex, $F(1,78) = 3.56$, $p = .063$, $\eta_p^2 = .044$. However, there was no significant interaction between sex and group, $F(1,78) = 0.62$, $p = .432$. As can be seen from Figure 2, the significant main effects indicated that autistic participants ($M = 0.306$, $SD = 0.28$) produced significantly more Um to total words than neurotypical participants ($M = 0.206$, $SD = 0.18$).

Figure 2 Mean Number of Ums as a Proportion of Total Words in the Unstructured Task



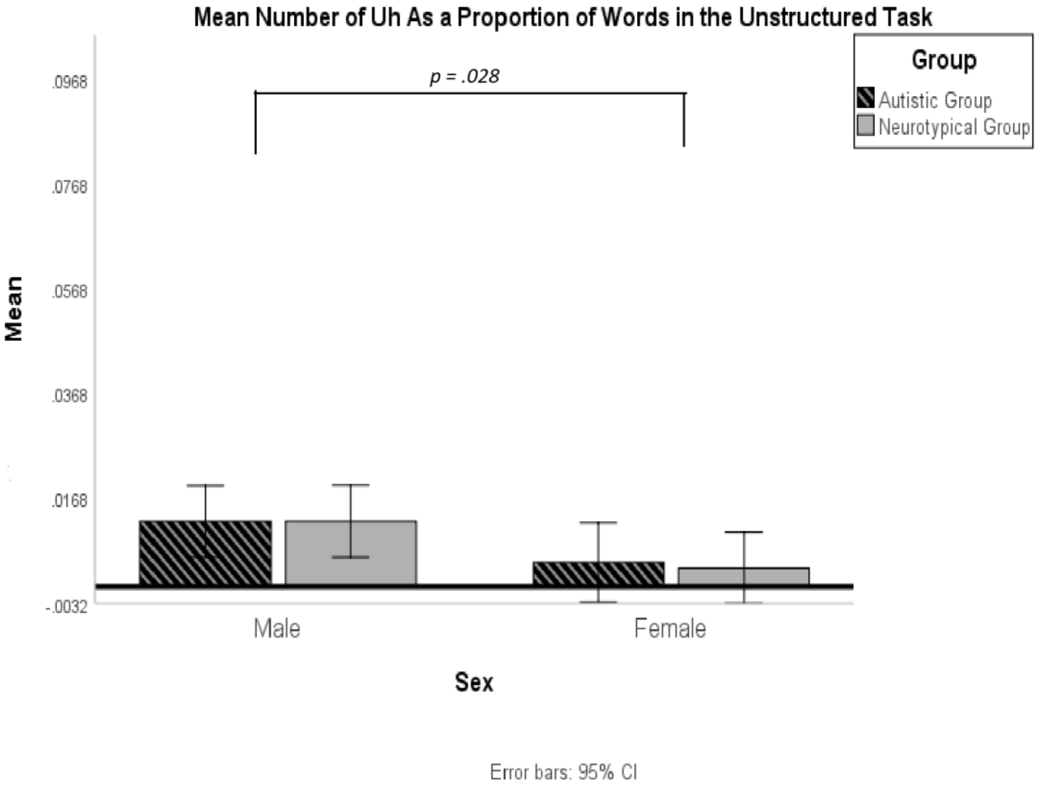
Note: The absolute maximum (i.e.1) is not represented in this figure.

2.3.4 Uh to total words:

For the structured task, there was no significant main effect of group or sex, and the interaction between these variables was not significant (all *F*s < 0.68 and *p*s > .41).

For the unstructured task, there was no significant main effect of group, $F(1,77) < 0.01$, $p = .967$. However, there was a significant main effect of sex, $F(1,77) = 5.01$, $p = .028$, $\eta_p^2 = .061$. There was no significant interaction between sex and group $F(1,77) = 0.14$, $p = .709$. As can be seen from Figure 3, the significant main effect indicated that males ($M = 0.012$, $SD = 0.21$) produced significantly more Uhs to total words than females ($M = 0.004$, $SD = 0.01$).

Figure 3 Mean Number of Uh As a Proportion of Words in the Unstructured Task



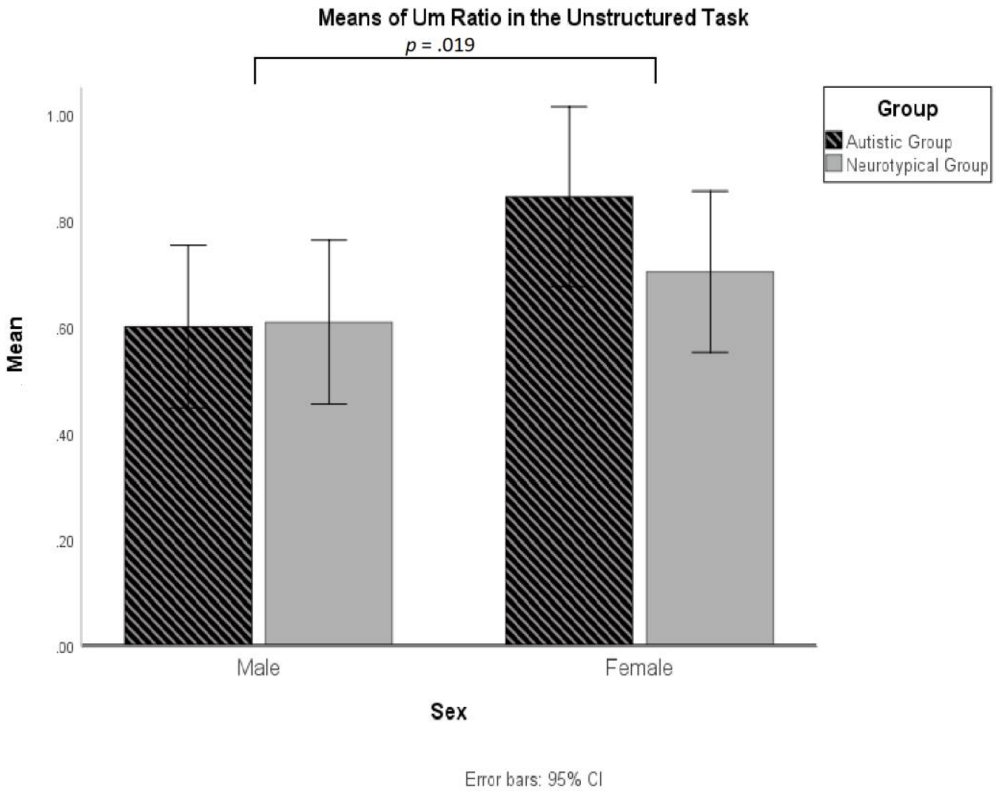
Note: The absolute maximum (i.e.1) is not represented in this figure.

2.3.5 Um Ratio:

For the structured task, there was no significant main effect of group or sex, and the interaction between these variables was also not significant (all *F*s < 2.5 and *p*s > .1).

For the unstructured task, there was no significant main effect of group, $F(1,78) = 0.41$, $p = .525$. However, there was a significant main effect of sex, $F(1,78) = 5.73$, $p = .019$, $\eta_p^2 = .068$. There was no significant interaction between sex and group $F(1,78) = 0.60$, $p = .440$. As can be seen from Figure 4, the significant main effect indicated that females ($M = 0.762$, $SD = 0.37$) produced significantly higher Um ratio than males ($M = 0.611$, $SD = 0.37$).

Figure 4 Means of Um Ratio in the Unstructured Task



2.3.6 Camouflaging Hypothesis

In order to test whether autistic females had more similar Um ratio to neurotypical females than autistic males to neurotypical males, we ran two independent samples t-tests to compare autistic and neurotypical females, and autistic and neurotypical males.

For the structured task, there was a non-significant difference between the Um ratio of autistic and neurotypical females, $t(38) = -0.95$, $p = .346$, which represented a small effect ($d = -0.30$). Similarly, there was a non-significant difference between the Um ratio of autistic and neurotypical males, $t(42) = 0.20$, $p = .839$, which represented a very small effect ($d = 0.06$).

For the unstructured task, there was a non-significant difference between the Um ratio of autistic and neurotypical females, $t(38) = 0.73$, $p = .468$, which represented a small effect ($d = 0.24$). Similarly, there was a non-significant difference between the Um ratio of autistic and neurotypical males, $t(42) = -0.73$, $p = .467$, which represented a small effect ($d = -0.22$).

2.4 Discussion

The current study investigated whether filled pause use differed between autistic and neurotypical children, and whether autistic female participants were more likely to engage in linguistic camouflage (as evidenced by more similar Um ratio use to neurotypical girls). Although differences in filled pause use between neurotypical and autistic children have been found in previous research (e.g., Gorman et al., 2016; McGregor & Hadden, 2018; Parish-Morris et al., 2017), it was unclear whether these differences were a result of phenotypic manifestation of autism or, alternatively, were simply driven by normative sex differences. The current study included a sufficient number of participants to explore interactions between sex and diagnosis to overcome the methodological limitations of previous research. Across two tasks (structured and unstructured), the proportional use of Uh increased with age. In addition, IQ was found to be positively correlated with Um ratio in the unstructured task.

For the structured task, neither significant main effects of sex or group, nor interactions were found for all indices (i.e., proportion of Um to total words, proportion of Uh to total words, Um ratio). For the unstructured task, we found only a main effect of group for Um to total words, meaning autistic participants produced a higher proportion of Um compared with neurotypical participants. In addition, after controlling for age, we found only a main effect of sex for Uh to total words, meaning males produced a higher proportion of Uh compared with females. In addition, we found a main effect of sex for Um ratio, such that females produced a significantly higher Um ratio than males (after controlling for IQ). No analysis showed a significant interaction between sex and diagnosis.

2.4.1 Group Difference in Um to Total Words

Autistic participants produced a higher proportion of Um to total words than neurotypical participants in the unstructured (but not structured) task. Considering that communication difficulties are a core aspect of autism (Demetriou et al., 2018), the open-ended and relatively unstructured nature of the task may have been particularly difficult for the autistic participants. Filled pauses, such as Um and Uh, are hypothesised by some researchers to be unintentional and reflect speech planning difficulties in both autistic and neurotypical populations (Clark & Fox Tree, 2002). This finding, therefore, suggests that the higher proportion of Ums for autistic participants may reflect planning difficulties at least in some contexts, which is in line with previous studies (e.g. Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Fraundorf & Watson, 2014; Metz & James, 2018). This finding, however, contrasts with previous research which studied filled pause use in the autistic population and has found lower rates of Um in autistic, than neurotypical participants (e.g., Gorman et al., 2016; McGregor & Hadden, 2018; Parish-Morris et al., 2017). The aforementioned studies, however, all included a higher proportion of males to females in the autistic group, meaning lower levels of Um production may have been driven by sex imbalanced groups. This seems particularly likely considering higher Um use in females than males has been robustly found in previous research. Indeed, we included nearly equal numbers of males and females in the autistic group (ratio: 1.2 to 1) and found higher levels of Um for these participants relative to neurotypical peers.

2.4.2 Sex Difference in Uh to Total Words

Exploratory analysis in the current study revealed a significantly positive correlation between age and higher Uh usage, which has also been found in previous research (e.g., Acton, 2011; Tottie, 2011). Accordingly, we included age as a covariate. The results showed that males produced more Uhs to total words than females in the unstructured task. Consistent with previous research, we found no significant main effect of group (Gorman et al., 2016; McGregor & Hadden, 2018). Overall, the pattern of results suggests that elevated use of Uh is robustly found in males and is not affected by autism. We have found that Um and Uh are differentially affected by sex and diagnosis, suggesting they are likely to be pragmatically distinct (Clark & Fox Tree, 2002). Future research should, therefore, analyse them separately as opposed to combining them into a singular category of filled pause.

2.4.3 Sex Difference in Um Ratio

Consistent with a number of previous research studies, we found a higher Um ratio in females than males (e.g. Acton, 2011; Laserna et al., 2014; Tottie, 2011) in the unstructured task. However, there was no interaction between sex and diagnosis, suggesting differences between autistic males and females were being driven primarily by their sex (as opposed to specific manifestation of autism). In addition, Parish-Morris et al. (2017) did not test for interaction effects, meaning their results may have also simply been driven by sex. Parish-Morris et al. (2017) argued that autistic girls may appear more neurotypical than autistic boys superficially due to elevated Um ratio, which may make their social difficulties harder to detect by parents and teachers. Only two previous studies examined specifically the relationship between filled pauses and parental perception of autistic symptoms and the results were mixed. Irvine et al. (2015) found that elevated rate of Um was correlated with higher autistic symptoms reported by parents in the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003), whilst Parish-Morris et al. (2017) found no relationship between rate of Um and SCQ scores.

2.4.4 Linguistic Camouflage Hypothesis

Two t-tests were run in order to test the linguistic camouflaging hypothesis. In contrast to Parish-Morris et al. (2017), we did not find that autistic girls had more similar Um ratios to neurotypical girls, than autistic boys did to neurotypical boys, suggesting that this form of linguistic camouflaging in females was not present in our sample. The effect size demonstrated that autistic males used more similar Um ratio than neurtotypical males, therefore the camouflaging hypothesis was disconfirmed. Despite this, it is important to examine rates of Um and Uh to total words when investigating linguistic camouflage, and how these may impact upon others' perception. Given equal Um ratios, it is possible that children produce Um and Uh at a different rate (to total words) in the language sample, causing them to sound differently in general. In our sample, for example, autistic girls (*M* = 3.87%) produced a higher proportion of Ums than neurotypical girls (*M* = 2.38%) in the unstructured task, despite having similar Um ratios. Given some disfluencies (e.g., repairs, filled pauses) are associated with negative social perceptions (e.g. Loy, Rohde, & Corley, 2017; Panico, Healey, Brouwer, & Susca, 2005), it is possible that elevated Um use could negatively impact upon listeners' perception.

There is increasing evidence suggesting that autism is a dimensional condition; children who have high autistic traits (but with no diagnosis) are likely to have similar levels of underlying difficulties than children who receive a diagnosis (Ozonoff et al., 2014; Wainer, Ingersoll, & Hopwood, 2011). Despite this, children who have high traits are likely to present with fewer overt autistic symptoms (e.g. social communicative difficulties) and therefore may not come to clinical attention. In view of this, it is possible that children with high traits sound more neurotypical superficially – as indicated by higher Um ratio – than children with a diagnosis. To illustrate this, our study included participants with high traits whilst Parish-Morris et al. (2017) only included children who received a diagnosis from the clinic. This interpretation is tentatively supported by the fact that autistic girls in our study had higher Um ratios (85%) than the girls who received diagnoses of autism (75%) in Parish-Morris et al. (2017). Future research should include larger numbers of participants, in order to be able to compare children with high traits (without a diagnosis) and those with a clinical diagnosis, as well as neurotypical participants.

2.4.5 Differences between Task Types

Consistent with previous research, differences in filled pause production was affected by the type of narrative task employed in the current study. Specifically, this study showed that participants' performance (as measured by the indices) only differed in the unstructured task. Previous research such as Gorman et al. (2016) found a significant main effect of Autism Diagnostic Observation Schedule (ADOS) activity, where the highest proportion of Uh occurred during the 'conversation' activity, and the lowest proportion of Uh occurred in the 'tell a story from a book' activity. Similarly, MacFarlane et al. (2017) found a significant main effect of ADOS activity on filled pause production, where relatively more filled pauses were produced in the 'conversation' and 'description of a picture book' activities than 'play' and 'tell a story from a book' activity by all participants (autistic, specific language impairment and neurotypical). One reason that tasks types may moderate filled pause production is due to the relative demands they place on executive functions (e.g., planning, inhibition), which may explain elevated filled pause use in the conversation activities in the research above. However, the studies above did not report whether task types affected filled pause production differently for autistic and neurotypical participants. The current study extended previous research to include a structured and an unstructured task in order to investigate how executive functioning may moderate filled pause production.

2.4.6 Directions for Future Research

Our findings provide preliminary evidence that autistic and neurotypical children produce Ums differently, specifically we found that autistic children produced a greater proportion of Ums than neurotypical children after controlling for the effects of verbal IQ. Elevated Um rate may therefore serve as a relatively objective behavioural marker to differentiate autistic and neurotypical participants, at least in unstructured verbal tasks. It will be useful to ascertain whether this finding replicates in larger samples of children, which can also investigate potential developmental differences. We have also provided evidence that the type of narrative task moderates filled pause use, it would therefore be useful for future research to systematically manipulate specific aspects of task type (e.g., amount of prompts, types of question, length of speech) in order to investigate whether task type moderates the effects of filled pause production for autistic and neurotpyical participants. For example, previous studies found that filled pause rates have been shown to increase with task difficulty (e.g., Finlayson & Corley, 2012; Lay & Paivio, 1969),

A strength of this study is that we have recruited children with high autistic traits who have not received a clinical diagnosis (i.e. come to clinical attention) and are therefore more likely to be camouflaging. Indeed, we found elevated Um ratios in autistic girls relative to the ones in Parish-Morris et al. (2017). However, it would be useful for future research to compare males and females with high autistic traits (who have not received a diagnosis), with those who have received a diagnosis. This should help develop our understanding of the differences between children who do or do not receive a diagnosis, despite similar levels of autistic traits, and therefore shed light on the camouflaging hypothesis.

Previous studies have included children with specific language impairment (SLI) in the comparison group (Gorman et al., 2016; McGregor & Hadden, 2018) and found that children with SLI produced similar levels of filled pauses to neurotypical children, but significantly higher rate of filled pauses than autistic children. Despite this, the groups were imbalanced in terms of their sex ratio; meaning group differences (or lack thereof) may have been driven primarily by sex difference, which again limits the conclusion that can be drawn from the studies. Future research should include equal numbers of females and males in all groups in order to account for normative sex differences in filled pause production. Although we controlled for verbal IQ in the current study, it would be useful for future research to include a broader range of language measures. This would allow investigations of whether group differences are driven by underlying differences in other specific language domains (e.g. pragmatic ability), as well as, determine the extent to which these domains are associated with filled pause use.

Previous studies have found that disfluencies are associated with negative social perception (e.g., Finlayson & Corley, 2012). Although higher Um ratios may serve to superficially normalise the way in which autistic children sound, it is possible that excessive or limited Um usage will make them sound atypical, despite having high Um ratio. For example, it is possible that there is an optimal Um usage, as well as proportion of fillers. In addition, the use of Um and Uh may be perceived differently. It is therefore fundamentally important for future research to include measures of listeners' perception of linguistic competency in order to test these hypotheses. Further, we have found that Um and Uh are differentially affected by sex and diagnosis, suggesting they are likely to be pragmatically distinct (Clark & Fox Tree, 2002). Future research should, therefore, analyse them separately as opposed to combining them into a singular category of filled pause.

Gender is a highly nuanced social construct that is influenced by multiple variables (e.g., societal expectations, parental attitudes, self-concept, biological factors). To date, researchers have not yet been able to separate the effects of sex and gender, it is, therefore, important for future research to try to disentangle these effects (Lai et al., 2015). This is particularly important considering 'sex' has robustly been found to influence filled pause use but little is known about the effect of gender.

2.4.7 Implications for Educational Psychology

The current study highlighted sex difference in two indices of filled pause production (Uh to total words and Um ratio). It is therefore important for practitioners to take into account these sex differences when assessing individuals. In support of this, a recent meta-analysis found that autistic individuals display typical sex differences in core autistic traits such as social interaction and repetitive behaviours (Hull et al., 2017). Although it potentially is counter-intuitive, it would be useful to compare a child with possible social communication needs with children of the same sex, in order to ascertain their relative strengths and needs.

Filled pauses are subtle in nature and may not always be consciously perceived by the listener. For example, Kjellmer (2003) argued that if a person was asked to repeat an utterance that contains a filled pause (e.g., I am, um, happy.), it is unlikely that the response would contain the filled pause. As such, practitioners may not always be consciously aware of the way in which filled pauses influence their perceptions. As mentioned, higher Um ratios may be associated with greater linguistic competency, which we found in autistic girls relative to autistic boys. Consequently, autistic girls may be perceived as more socially competent which could make it more difficult to notice their difficulties. It may therefore be useful for practitioners to collect objective and detailed language samples, or alternatively pay closer attention to children and young people's social use of language.

In triangulation with other assessment methods, collecting detailed language samples, especially in unstructured context such as the playground, could yield potentially informative data. This is particularly important considering that both girls and boys on the autistic spectrum may camouflage their difficulties and appear more neurotypical behaviourally; consequently they may be missed by practitioners and not have the opportunity to benefit from early intervention (Pickles et al., 2014). Recently there is emerging evidence that there may be a female specific phenotype of autism (Wood & Wong, 2017). Despite this, current conceptualisation of autism have been developed predominately from male samples, such that our understanding of the condition may be male-biased (Kirkovski, Enticott, & Fitzgerald, 2013; Rynkiewicz et al., 2016), which may result in autistic females being under or misdiagnosed using traditional assessments (e.g., ADOS). Adults who received late diagnoses of autism often reported exhaustion when trying to 'fit-in' and making sense of their difficulties in childhood, which are associated with mental health difficulties such as depression, anxiety (Cage & Troxell-Whitman, 2019; Leedham, Thompson, Smith, & Freeth, 2019). Whilst the utility of a diagnostic label of autism is a contentious issue, it is important for Educational Psychologists to be aware of the potential female specific phenotype of autism as this would help inform psychological formulation.

Appendix A Search Strategy Protocol

This form is adapted from: Miller, S.A. (2001). PICO worksheet and search strategy. US National Center for Dental Hygiene Research.

PICo Worksheet and Search Strategy Protocol

1. Define your question using PICo by identifying: Population, Interest, and Context:

Population: Neurotypical children and children with a diagnosis of Autism Spectrum Disorder (ASD)

Interest: Disfluencies in spontaneous speech

Context: All

2. Type of study – methodology:

Quantitative, empirical studies

List specific qualitative search terms or filters:

(autism or asd or autism spectrum disorder or asperger* or autism spectrum condition or autisti* spectrum) AND ("filled pause*" or disfluenc* or "conversational filler* or "unfilled pause* or dysfluenc*)

3. Type of study – data collection:

List types of data collection instruments: Frequency Count

4. List main topics and alternate terms from your PICo question that can be used for your search:

- Discourse markers
- Fillers

5. Write out your search strategy:

6. List any limits that may apply to your search:

Age: none

Year(s) of publication: none

Language(s):. English and Chinese

	Inclusion criteria	Exclusion criteria
Participants	- Individuals with a diagnosis of ASD and neurotypical children	- Any participants with co-morbid diagnoses - Studies that include only one group of children - Any Group (n < 5)
Outcomes	- Measures of disfluencies	- Intervention studies - Treatment studies
Language/country	- English/ Chinese	None
Date		Studies before 2000
Type of research	- Quantitative studies	- Intervention studies - Single case studies

7. List the databases you will search:

PsychInfo, Medline, Psych Articles, CINAHL (Plus with Full Text), ERIC, Pro Quest Education Collection , Web of Science, PubMed, ProQuest Dissertations & Theses A&I, SCOPUS,

Appendix B List of Excluded Studies for Systematic Review

Record	Reason for exclusion
Disfluencies in Children With Language Impairment (2019) Bergström, A., Johansson, M., & Eklund, R. DOI: 10.1177/1048395019833703e	No autistic group
Assessing Bilingual Children: Are Their Disfluencies Indicative of Stuttering or the By-product of Navigating Two Languages? (2018) Byrd CT DOI: 10.1055/s-0038-1667161	No autistic group
The visual attention span deficit in Chinese children with reading fluency difficulty. (2018) Zhao, Jing.; Liu, Menglian.; Liu, Hanlong.; Huang, Chen. DOI: 10.1016/j.ridd.2017.12.017	Visual disfluency, not speech disfluency
Speech disfluencies in children with Down Syndrome. (2018) Eggers, Kurt.; Van Eerdenbrugh, Sabine DOI: 10.1016/j.jcomdis.2017.11.001	No autistic group
Automated screening for Fragile X premutation carriers based on linguistic and cognitive computational phenotypes. (2017) Movaghar A et al. DOI: 10.1038/s41598-017-02682-4	No autistic group

Disfluency characteristics of children with attention-deficit/hyperactivity disorder symptoms. Hyunkyung et al. (2017) DOI: 10.1016/j.jcomdis.2016.12.001	No autistic group
Assessing language disfluency in school-aged children with autism spectrum disorder in a virtual, public speaking task. Torabian, S. (2017)	Dissertation, contacted author but no response
Neurodevelopment for syntactic processing distinguishes childhood stuttering recovery versus persistence Usler, E. & Weber-Fox, C. (2015)	No autistic group
Lost in the literature, but not the caseload: Working with atypical disfluency from theory to practice. (2014) DOI: 10.1055/s-0034-1371757	Single case study
Oral electromyography activation patterns for speech are similar in preschoolers who do and do not stutter. Walsh, B. (2013) DOI: 10.1044/1092-4388(2013/12-0177)	No autistic group
Dysfluencies in the speech of adults with intellectual disabilities and reported speech difficulties. (2013). Coppens-Hofman et al DOI: 10.1016/j.jcomdis.2013.08.001	No autistic group
Language dysfluencies in females with the FMR1 premutation. (2013). Sterling et al DOI: 10.1016/j.bandc.2013.02.009	No autistic group
Are language production problems apparent in adults who no longer meet diagnostic criteria for attention-deficit/hyperactivity disorder? (2012) Engelhardt et al. DOI: 10.1080/02643294.2012.712957	No autistic group

Teasing out specific language impairment from an autism spectrum disorder. (2012) Tierney et al. DOI: 10.1097/DBP.0b013e31824ea235	Single case study
The uses of conversational speech in measuring language performance and predicting behavioural and emotional problems. (2012) https://search.proquest.com/docview/871183997	No autistic group
Language production strategies and disfluencies in multi-clause network descriptions: A study of adult attention-deficit/hyperactivity disorder. (2011) Engelhardt et al. http://dx.doi.org/10.1037/a0022436	No autistic group
Disfluency characteristics observed in young children with autism spectrum disorders: a preliminary report. (2010). Plexico LW; Cleary JE; McAlipine A; Plumb AM http://dx.doi.org/10.1044/ffd20.2.42	No neurotypical group
The role of inhibition in the production of disfluencies. (2010). Engelhardt et al. http://dx.doi.org/10.3758/MC.38.5.617	No autistic group
Disfluent utterances of Japanese youth with Down's syndrome.	No autistic nor neurotypical group
Early speech- and language-impaired children: Linguistic, literacy, and social outcomes. (2006). Glogowska et al. http://dx.doi.org/10.1017/S0012162206001046	No autistic group
Handwriting and Attention in Children and Adults with Attention Deficit Hyperactivity Disorder. (2004). Tucha and Langa.	Not speech disfluency
Conversational characteristics of children with Fragile X syndrome: Repetitive speech. (2001)	No autistic nor neurotypical group

http://dx.doi.org/10.1352/0895-8017(2001)106<0028:CCOCWF>2.0.CO;2	
Cluttered communication in a deafened adult with autistic features. (1996). Thacker & Austen http://dx.doi.org/10.1016/S0094-730X(96)00029-0	Single case – not autistic
Conversational skills of individuals with fragile-X syndrome: A comparison with autism and Down syndrome. (1991). Ferrier et al. http://dx.doi.org/10.1111/j.1469-8749.1991.tb14961.x	No neurotypical group
Autism and the fragile X syndrome. http://dx.doi.org/10.1097/00004703-198309000-00002	No neurotypical group
Reduction of inappropriate verbalizations in an emotionally disturbed adolescent. Herdman (1979).	Single case design
Language patterns of adolescent autistics. Simmons et al. (1975). http://dx.doi.org/10.1007/BF01540680	No neurotypical group

Appendix C Quality Assessment of Included Studies

Appraisal of Cross-sectional Studies Checklist	Thurber and Tager-Flusberg	Shriberg et al. 2001	Lunsford et al., 2010	Lake et al., 2011	Suh et al., 2014	Irvine et al., 2016	Gorman et al., 2016	de Marchena & Eigsti, 2016	Engelhardt et al., 2017	MacFalane et al., 2017	Parish- Morris et al., 2017	McGregor & Hadden., 2018
Introduction												
Were the aims/objectives of the study clear?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Methods												
Was the study design appropriate for the stated aim(s)?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the sample size justified?	N	N	N	N	N	N	N	N	N	N	N	N
Was the target/reference population clearly defined? (Is it clear w	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the sample frame taken from an appropriate population base	N	N	N	N	N	N	N	N	N	N	Y	N
Was the selection process likely to select subjects/participants tha	N	N	N	N	N	N	N	N	N	N	N	N
Were measures undertaken to address and categorise non-respond	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK
Were the risk factor and outcome variables measured appropriate	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the risk factor and outcome variables measured correctly u	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Is it clear what was used to determine statistical significance and	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the methods (including statistical methods) sufficiently des	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Results												
Were the basic data adequately described?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Does the response rate raise concerns about non-response bias?	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK
If appropriate, was information about non-responders described?	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK	DK
Were the results internally consistent?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the results presented for all the analyses described in the m	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Discussion												
Were the authors' discussions and conclusions justified by the res	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the limitations of the study discussed?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Other												
Were there any funding sources or conflicts of interest that may a	N	N	N	N	N	N	N	N	N	N	N	N
Was ethical approval or consent of participants attained?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	14	14	14	14	14	14	14	14	14	14	15	14

Appendix D Confidentiality Agreement

CONFIDENTIALITY AGREEMENT

Sex/Gender Differences in the Language Use in Children with Autism Spectrum Condition: On the Distribution of ‘Um’ and ‘Uh ’

I ________________________, agree to maintain full confidentiality in regards to any and all audio or video files and documentation received from the **Sex/Gender Differences in the Language Use in Children with Autism Spectrum Condition: On the Distribution of ‘Um’ and ‘Uh ’**study. Furthermore, I agree:

1. To hold in strictest confidence the identification and / or identifiable information of any individual that may be inadvertently revealed during the transcription, listening or viewing of audio/video-taped interviews, or in any associated documents.
2. To not make copies of any audio/video files or other study–related documents.
3. To ensure that study-related audio/video tapes, files and materials I am required to work with are stored in a safe, secure location when not in use.

I am aware that I can be held legally liable for any breach of this confidentiality agreement, and for any harm incurred by individuals if I disclose identifiable information contained in the audio/video-tapes and/or files to which I will have access.

Name (printed)

Signature

Date

Appendix E Social and Communication Disorders Checklist (Skuse et al, 2005)

Social and Communication Disorders Checklist (Skuse et al, 2005)

1. What is your child's name?

   ____

   (As soon as we receive this questionnaire, your child will be assigned a number and their name will be removed from this document)

2. Has your child ever received a diagnosis of Autism / Autistic Spectrum Disorder/ Asperger Syndrome?

   Yes / No

3. If yes, please specify which diagnosis and by whom this was diagnosed (e.g. paediatrician, psychiatrist).

   ____

Checklist

For each item, please mark the box that best describes your child's behaviour over the past 6 months.

	Not true	Quite or sometimes true	Very or often true
1. Not aware of other people's feelings	☐	☐	☐
2. Does not realise when others are upset or angry	☐	☐	☐
3. Does not notice the effect of his/her behaviour on other members of the family	☐	☐	☐
4. Behaviour often disrupts family life	☐	☐	☐
5. Very demanding of other people's time	☐	☐	☐
6. Difficult to reason with when upset	☐	☐	☐
7. Does not seem to understand social skills, e.g. persistently interrupts conversations	☐	☐	☐
8. Does not pick up on body language	☐	☐	☐
9. Does not appear to understand how to behave when out (e.g. in shops, or other people's homes)	☐	☐	☐
10. Does not realise if s/he offends people with her/his behaviour	☐	☐	☐
11. Does not respond when told to do something	☐	☐	☐
12. Cannot follow a command unless it is carefully worded	☐	☐	☐

Do you have any other comments or concerns? (If yes, please describe.)

Thank you for your participation!

Appendix F Parent Information Sheet

Parent Information Sheet

Study Title: An Exploration of Gender Differences During Social Interaction in Children and Adolescents

Researcher: Henry Wood, Bonnie Wong

ERGO number: 32187

You are being invited to take part in the above research study. To help you decide whether you would like to take part or not, it is important that you understand why the research is being done and what it will involve. Please read the information below carefully and ask questions if anything is not clear or you would like more information before you decide to take part in this research. You may like to discuss it with others but it is up to you to decide whether or not to take part. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

We are Trainee Educational Psychologists studying for the Doctorate in Educational Psychology at the University of Southampton.

We are interested in gender differences in social interaction in children and adolescents aged 8 – 13 years of age, including those young people who have a diagnosis of autism spectrum disorder (ASD). According to the American Psychiatric Association (APA), Autism spectrum disorder (ASD), also known as autism spectrum condition, is a neurodevelopmental syndrome characterised by 'persistent impairment in reciprocal social communication and social interaction, and restricted, repetitive patterns of behaviour, interests or activities'.

We want to explore the 'camouflaging' phenomenon, defined as a strategy that is suggested to be commonly used by girls diagnosed with ASD to "mask" difficulties in social interactions that can be experienced in young people with ASD.

UNIVERSITY OF Southampton

Why have I been asked to participate?

You have been asked to participate because your child is aged between 8 and 13 years of age.

What will happen to me if I take part?

You will be asked to complete a questionnaire about your child(ren), which is attached alongside the consent form, consists of 12 items that consider autism traits in children and young people who do and do not have a diagnosis of ASD. In addition, we may ask you to complete a questionnaire that measures different social skills in your child, including e.g., communication, cooperation and empathy.

If you agree for your child to take part, they will come to a quiet room with the researcher in the school and will be asked to complete a pen and paper based task where they will asked to guess the emotional state (e.g., happy) of a picture of an individual when only the eye region is visible. They will also be asked to complete a test that measures their verbal ability and non-verbal ability and a drawing task with the researcher. Finally, they will be interviewed about the tasks they completed with the researcher, and asked to talk about a game or hobby they enjoy.

The drawing task and subsequent interview will be video recorded.

Are there any benefits in my taking part?

Your participation will help develop our understanding of the social challenges in ASD. We hope that this project will help to identify aspects of social situations that young people diagnosed with ASD find difficult.

As a thank you for your valuable contribution and time, we are able to offer you a £5 voucher for participation. Additionally, we are able to offer you an opportunity to attend a workshop on Autism led by us.

Will my participation be confidential?

Your participation and the information we collect about you during the course of the research will be kept strictly confidential.

Only members of the research team and responsible members of the University of Southampton may be given access to data about you for monitoring purposes and/or to carry out an audit of the study to ensure that the research is complying with applicable regulations. Individuals from regulatory authorities (people who check that we are carrying out the study correctly) may require access to your data. All of these people have a duty to keep your information, as a research participant, strictly confidential.

All of the data we collect as part of this project will be pseudonymised and treated confidentially. Only the research team has access to such data. All participants will be assigned a number, which will be used for data analysis. A spreadsheet that links the participants' names to the number will be kept on a password protected memory stick that only the researchers have access to. Voluntary Research Assistant (VRA) s (Charlotte Burham, Natalia Nadolska, Lara Rosa, Sarah Davil) are used to assist with transcription of the interview data in an anonymised format. They have signed a confidentiality agreement in order to protect participants' data.

This will allow the researchers to delete participants' data after the experiment has taken place if requested. Consent forms will be stored in a locked cabinet at one of the researchers' house and destroyed one year after data collection. All data will be stored in line with the Data Protection Act (2018).

Do I have to take part?No, it is entirely up to you to decide whether or not to take part. If you decide you want to take part, you will need to sign a consent form to show you have agreed to take part.

What happens if I change my mind?

You have the right to change your mind and withdraw at any time without giving a reason and without your participant rights being affected, up to the point of data analysis. If you withdraw, your data will be destroyed.

What will happen to the results of the research?

Your personal details will remain strictly confidential. Research findings made available in any reports or publications will not include information that can directly identify you without your specific consent.

The results of the research will be written up as part of our theses, which may be published in the future (e.g. academic journals). A copy of the results will be available upon request. The research data will be stored for a minimum of 10 years, as per University of Southampton policy.

Where can I get more information?

If you have any further questions please contact Henry Wood at hwd1n15@soton.ac.uk or Bonnie Wong at b.wong@soton.ac.uk

What happens if there is a problem?

If you have a concern about any aspect of this study, you should speak to the researchers (Henry Wood at hwd1n15@soton.ac.uk or Bonnie Wong at b.wong@soton.ac.uk) who will do their best to answer your questions.

If you remain unhappy or have a complaint about any aspect of this study, please contact the University of Southampton Research Integrity and Governance Manager (023 8059 5058, rgoinfo@soton.ac.uk).

Data Protection Privacy Notice

The University of Southampton conducts research to the highest standards of research integrity. As a publicly-funded organisation, the University has to ensure that it is in the public interest when we use personally-identifiable information about people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use information about you in the ways needed, and for the purposes specified, to conduct and complete the research project. Under data protection law, 'Personal data' means any information that relates to and is capable of identifying a living individual. The University's data protection policy governing the use of personal data by the University can be found on its website (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

This Participant Information Sheet tells you what data will be collected for this project and whether this includes any personal data. Please ask the research team if you have any questions or are unclear what data is being collected about you.

Our privacy notice for research participants provides more information on how the University of Southampton collects and uses your personal data when you take part in one of our research projects and can be found at http://www.southampton.ac.uk/assets/sharepoint/intranet/ls/Public/Research%20and%20Integrity%20Privacy%20Notice/Privacy%20Notice%20for%20Research%20Participants.pdf

Any personal data we collect in this study will be used only for the purposes of carrying out our research and will be handled according to the University's policies in line with data protection law. If any personal data is used from which you can be identified directly, it will not be disclosed to anyone else without your consent unless the University of Southampton is required by law to disclose it.

Data protection law requires us to have a valid legal reason ('lawful basis') to process and use your Personal data. The lawful basis for processing personal information in this research study is for the performance of a task carried out in the public interest. Personal data collected for research will not be used for any other purpose.

For the purposes of data protection law, the University of Southampton is the 'Data Controller' for this study, which means that we are responsible for looking after your information and using it properly. The University of Southampton will keep identifiable information about you for 10 years after the study has finished after which time any link between you and your information will be removed.

To safeguard your rights, we will use the minimum personal data necessary to achieve our research study objectives. Your data protection rights – such as to access, change, or transfer such information – may be limited, however, in order for the research output to be reliable and accurate. The University will not do anything with your personal data that you would not reasonably expect.

If you have any questions about how your personal data is used, or wish to exercise any of your rights, please consult the University's data protection webpage (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page) where you can make a request using our online form. If you need further assistance, please contact the University's Data Protection Officer (data.protection@soton.ac.uk).

Thank you for taking the time to read the information sheet and considering taking part in the research.

[4th March 2019] [Version number 5] [Ethics/IRAS number 32187]

Appendix G Parent Consent Form

UNIVERSITY OF Southampton

CONSENT FORM

Study title: An Exploration of Gender Differences During Social Interaction in Children and Adolescents

Researcher name: Henry Wood, Bonnie Wong

ERGO number: 32187

Participant Identification Number:

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (4th March, 2019, Version 5) and have had the opportunity to ask questions about the study.	
I agree to take part in this research project and agree for my data to be used for the purpose of this study.	
I understand my participation is voluntary and I may withdraw until the point of data analysis for any reason without my participation rights being affected.	
I understand my child's participation is voluntary and he/she may withdraw up until the point of data analysis for any reason without his/her rights being affected.	
I understand that my child will be video recorded whilst completing a drawing task and the subsequent interview.	
I understand that the information collected about my child may be anonymised and used in future ethically approved research studies.	
I understand that information collected about me during my participation in this study will be stored on a password protected computer and that this information will only be used for the purpose of ethically approved research studies.	

Name of participant (print name)……………………………………………………………………………

Signature of participant…………………………………………………………………………………….

Date……………………………………………………………………………………….. ………………….

Name of researcher (print name)…BONNIE WONG AND HENRY WOOD…………………………

Signature of researcher …………………………………………………………………………………….

Date………..

[4th March 2019] [Version Number 4] [Ethics/IRAS reference 32187]

Appendix H Child Information Sheet

UNIVERSITY OF Southampton

Invite for participation – Child/Young person

Hello,

We are researchers from the University of Southampton.

We are interested in how children socialise and make friends. Your parents and/or Head Teacher have said they are happy for you to take part in our study. However, we need to make sure you are happy to take part. If you are not, then you can stop at any time.

If you agree to take part, we will complete some questionnaires with you, as well as complete a drawing task with you. We will then ask you some questions about what you have done. The whole process will take around 45 minutes. If you need a break, then just let us know.

Although this is going to be recorded on the video camera, everything you do will be confidential. You can choose not to take part, but if you do it will help us to understand how children socialise and make friends.

Henry and Bonnie

[May 20, 2018] [Version 1] [Ethics/IRAS number 32187]

Appendix I Child Consent Form

UNIVERSITY OF Southampton

CONSENT FORM – Child/Young person

Dear Henry and Bonnie

I have read, or been read, the information letter which explains the research that you would like to do. I understand what you would like me to do and am happy to continue. I understand that I can stop at any time.

My details are:

Name:

Name of school:

Signature: Date:

[May 20, 2018] [Version 1] [Ethics/IRAS number 32187]

Appendix J Interview Schedule/ Prompts

Structured Task

1. **Tell me about the drawing task that you did with the other researcher (Bonnie/ Henry)?**
    a. Do you remember drawing something with Bonnie/ Henry?
    b. When did you do the drawing?
    c. What did you draw?
    d. How did you feel about the task?
        i. Did you enjoy it or not?
        ii. Why?

2. **Tell me how your friend/ family member would think of the task.**
    a. Can you think of a friend/ family member?
    b. How are they normally like with drawing?
    c. Which part of the task do you think they would like the most and why?
    d. Which part of the task do you think they would like the least and why?

Unstructured Task

The following script will be used (adapted from Nippold et al 2005).

3. **I am hoping to learn what people know about certain topics. There are no penalties for incorrect answers.**
    a. What is your favourite game or hobby?
    b. Why is [e.g. chess] your favourite game?
    c. I am not too familiar with the game of [chess], so I would like you to tell me all about it. For example, tell me what the goals are, and how many people may play a game. Also tell me about the rules that players need to follow. Tell me everything you can think of about the game of [chess] so that someone who has never played before will know how to play.
    d. Now I would like you to tell me what a player should do in order to win the game of (chess). In other words, what are some key strategies that every good player should know?

Reference:

Nippold, M., Hesketh, L., Duthie, J. and Mansfield, T. (2005). Conversational Versus Expository Discourse. *Journal of Speech Language and Hearing Research*, 48(5), 1048-1064.

List of References

Acton, E. K. (2011). On Gender Differences in the Distribution of um and uh. In *NWAV 39* (Vol. 17).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Andrés-Roqueta, C., & Katsos, N. (2017). The Contribution of Grammar, Vocabulary and Theory of Mind in Pragmatic Language Competence in Children with Autistic Spectrum Disorders. *Frontiers in Psychology*, *8*(6), 2008–2012. https://doi.org/10.3389/fpsyg.2017.00996

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. Journal of Psycholinguistic Research, 32(1), 25–36. https://doi.org/10.1023/A:1021980931292

Attwood T (2007) The Complete Guide to Asperger's Syndrome. London: Jessica Kingsley Publishers.

Backer van Ommeren, T., Koot, H. M., Scheeren, A. M., & Begeer, S. (2017). Sex differences in the reciprocal behaviour of children with autism. *Autism*, *21*(6), 795–803. https://doi.org/10.1177/1362361316669622

Baddeley, A. D. (1992). Working memory. *Science*, 255, 556–559.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215. https://doi.org/10.1016/S0140-6736(06)69041-7

Baron-Cohen, S. (1995).*Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: The MIT Press.

Barr, D. J., & Seyfeddinipur, M. (2010). The role of fillers in listener attributions for speaker disfluency. *Language and Cognitive Processes*, *25*(4), 441–455. https://doi.org/10.1080/01690960903047122

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, 107(3), 1018–1034. https://doi.org/10.1016/j.cognition.2007.12.008

Bishop, D. V. (1992). The underlying nature of specific language impairment. *Journal of Child*

Psychology and Psychiatry, and Allied Disciplines, *33*(July 1991), 3–66. https://doi.org/10.1111/j.1469-7610.1992.tb00858.x

Bishop, D.V.M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*, 879–891. Retrieved from https://onlinelibrary.wiley.com/journal/14697610

Bishop, D. V. M. (2010). Overlaps between autism and language impairment: Phenomimicry or shared etiology? Behavior Genetics, 40(5), 618–629. https://doi.org/10.1007/s10519-010-9381-x

Blanca, M.J., Alarcón, R., Arnau, J., Bono, R., and Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29*(3), 552-557. https://doi.org/10.7334/psicothema2016.383

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review*, *3*(4), 395–421.

Boorse, J., Cola, M., Plate, S., Yankowitz, L., Pandey, J., Schultz, R. T., & Parish-Morris, J. (2019). Linguistic markers of autism in girls : evidence of a " blended phenotype " during storytelling. *Molecular Autism*, 1–12. https://doi.org/10.1186/s13229-019-0268-2

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech*, *44*(2), 123–147. https://doi.org/10.1177/00238309010440020101

Brennan, Susan E.; Williams, M. (1995). The Feeling of Another's Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *Journal of Memo*, *34*, 383–398.

Cage, E., & Troxell-Whitman, Z. (2019). Understanding the Reasons, Contexts and Costs of Camouflaging for Autistic Adults. *Journal Of Autism And Developmental Disorders*, *49*(5), 1899-1911. https://doi.org/10.1007/s10803-018-03878-x

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12062148

Collard, P., Corley, M., MacGregor, L. J., & Donaldson, D. I. (2008). Attention Orienting Effects of Hesitations in Speech: Evidence From ERPs. *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(3), 696–702. https://doi.org/10.1037/0278-7393.34.3.696

Constantino, J. N. (2017). Taking stock of critical clues to understanding sex differences in the prevalence and recurrence of autism. *Autism*, *21*(6), 769–771. https://doi.org/10.1177/1362361317704414

Creemers, A., & Schaeffer, J. (2016). *Specific Language Impairment and High Functioning Autism: Evidence for Distinct Etiologies and for Modularity of Grammar and Pragmatics*. In K. Perkins, Laurel; Dudley, Rachel; Gerard, Juliana; Hitczenko (Ed.) (pp. 1–12). Somerville, MA, USA: Cascadilla Proceedings Project.

De Marchena, A., & Eigsti, I. M. (2016). The art of common ground: Emergence of a complex pragmatic language skill in adolescents with autism spectrum disorders. *Journal of Child Language*, *43*(1), 43–80. https://doi.org/10.1017/S0305000915000070

Demetriou, E. A., Lampit, A., Quintana, D. S., Naismith, S. L., Song, Y. J. C., Pye, J. E., … Guastella, A. J. (2018). Autism spectrum disorders: A meta-analysis of executive function. *Molecular Psychiatry*, *23*(5), 1198–1204. https://doi.org/10.1038/mp.2017.75

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., … Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS ONE*, *10*(9), 1–15. https://doi.org/10.1371/journal.pone.0136100

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, *6*(12), 1–7. https://doi.org/10.1136/bmjopen-2016-011458

Dunbar, R. I. M. (2003). The Social Brain: Mind, Language, and Society in Evolutionary Perspective. *Annual Review of Anthropology, 32*(1), 163–181. https://doi.org/10.1146/annurev.anthro.32.061002.093158

Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. Annals of Human Biology, 36(5), 562–572. https://doi.org/10.1080/03014460902960289

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). *British Picture Vocabulary Scale (2nd edn)*. NFER Nelson.

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology*, *14*, 91–118.

Eigsti, I. M., De Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders, 5*(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Engelhardt, P. E., Alfridijanta, O., McMullon, M. E. G., & Corley, M. (2017). Speaker-Versus Listener-Oriented Disfluency: A Re-examination of Arguments and Assumptions from Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *47*(9), 2885–2898. https://doi.org/10.1007/s10803-017-3215-0

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.

Finlayson, I. R., & Corley, M. (2012). Disfluency in dialogue: An intentional signal from the speaker? *Psychonomic Bulletin and Review*, *19*(5), 921–928. https://doi.org/10.3758/s13423-012-0279-x

Fraundorf, S. H., & Watson, D. G. (2014). Alice's adventures in um-derland: Psycholinguistic sources of variation in disfluency production. *Language, Cognition and Neuroscience*, *29*(9), 1083–1096. https://doi.org/10.1080/01690965.2013.832785

Fricke, S., Burgoyne, K., Bowyer-Crane, C., Kyriacou, M., Zosimidou, A., Maxwell, L., … Hulme, C. (2017). The efficacy of early language intervention in mainstream school settings: a randomized controlled trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(10), 1141–1151. https://doi.org/10.1111/jcpp.12737

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). "Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis." *Autism Research*, *10*(3), 384–407. https://doi.org/10.1002/aur.1678

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

Grieve, J., Liberman, M., Wieling, M., Coleman, J., Fruehwald, J., & Bouma, G. (2016). Variation and Change in the Use of Hesitation Markers in Germanic Languages. *Language Dynamics and Change*, *6*(2), 199–234. https://doi.org/10.1163/22105832-00602001

Heeman, P. A., Lunsford, R., Selfridge, E., Black, L., & van Santen, J. (2010). Autism and interactional aspects of dialogue. *Sigdial*, 249–252. Retrieved from http://dl.acm.org/citation.cfm?id=1944506.1944551

Heine, B. (2013). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, *51*(6), 1205–1247. https://doi.org/10.1515/ling-2013-0048

Hellermann, J. (2009). Looking for evidence of language learning in practices for repair: A case study of self-initiated self-repair by an adult learner of english. *Scandinavian Journal of Educational Research*, *53*(2), 113–132. https://doi.org/10.1080/00313830902757550

Hull, L., Mandy, W., & Petrides, K. V. (2017). Behavioural and cognitive sex/gender differences in autism spectrum condition and typically developing males and females. *Autism*, *21*(6), 706–727. https://doi.org/10.1177/1362361316669087

Irvine, Christina A; Eigsti, Inge-Marie; Fein, D. A. (2015). Uh, Um, and Autism Filler Disfluencies as Pragmatic Markers in Adolescents with Optimal Outcomes from Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 1061–1070. https://doi.org/10.1007/s10803-015-2651-y

James, L. E., Chambers, B. N., Placzek, C. L., James, L. E., Chambers, B. N., & Placzek, C. L. (2018). How scenes containing visual errors affect speech fluency in young and older adults young and older adults. *Aging, Neuropsychology, and Cognition*, *25*(4), 520–534. https://doi.org/10.1080/13825585.2017.1337061

Kercood, S., Grskovic, J. A., Banda, D., & Begeske, J. (2014). Working memory and autism: A review of literature. *Research in Autism Spectrum Disorders*, *8*(10), 1316–1332. https://doi.org/10.1016/j.rasd.2014.06.011

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

Kidd, C., White, K. S., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science*, *14*(4), 925–934. https://doi.org/10.1111/j.1467-7687.2011.01049.x

King, J. P. J., Loy, J. E., & Corley, M. (2018). Contextual Effects on Online Pragmatic Inferences of Deception. *Discourse Processes, 55*(2), 123–135. https://doi.org/10.1080/0163853X.2017.1330041

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An Investigation of Language Impairment in Autism: Implications for Genetic Subgroups. *Language and Cognitive Processes*, *16*(2–3), 287–308. https://doi.org/10.1080/01690960042000058

Kjellmer, G. (2003). Hesitation. In Defence of ER and ERM. *English Studies*, *84*(2), 170–198. https://doi.org/10.1076/enst.84.2.170.14903

Kwok, E. Y. L., Brown, H. M., Smyth, R. E., & Oram Cardy, J. (2015). Meta-analysis of receptive and expressive language skills in autism spectrum disorder. R*esearch in Autism Spectrum Disorders, 9*, 202–222. https://doi.org/10.1016/j.rasd.2014.10.008

Lai, M. C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Academy of Child and Adolescent Psychiatry*, *54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Larkin, F., Hobson, J. A., Hobson, R. P., & Tolmie, A. (2017). Collaborative competence in dialogue: Pragmatic language impairment as a window onto the psychopathology of autism. *Research in Autism Spectrum Disorders*, *43–44*(April), 27–39. https://doi.org/10.1016/j.rasd.2017.09.004

Lai, M. C., Lombardo, M. V., Ruigrok, A. N. V., Chakrabarti, B., Auyeung, B., Szatmari, P., … Baron-Cohen, S. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, *21*(6), 690–702. https://doi.org/10.1177/1362361316671012

Lake, Johanna K.; Humphreys, Karin R.; Cardy, S. (2010). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. *Psychonomic Bulletin & Review*, *18*(1), 135–140. https://doi.org/10.3758/s13423-010-0037-x

Laserna, C. M., Seih, Y.-T., & Pennebaker, J. W. (2014). Um . . . Who Like Says You Know . *Journal of Language and Social Psychology*, *33*(3), 328–338. https://doi.org/10.1177/0261927x14526993

Leedham, A., Thompson, A., Smith, R., & Freeth, M. (2019). ‘I was exhausted trying to figure it out’: The experiences of females receiving an autism diagnosis in middle to late adulthood. *Autism*, 136236131985344. https://doi.org/10.1177/1362361319853442

Lickley, R. (2017). Disfluency in typical and stuttered speech. *Fattori Sociali e Biologici Nella Variazione Fonetica - Social and Biological Factors in Speech Variation*, 373–387. https://doi.org/10.17469/O2103AISV000019

Lickley, R. J. (2001). ISCA Archive Dialogue moves and disfluency rates. *Proceedings of DISS’01*, 93–96.

Lombardo, M. V, Lai, M.-C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular Psychiatry*. https://doi.org/10.1038/s41380-018-0321-0

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 developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659-685.

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 and Allied Disciplines*, *49*(11), 1184–1192. https://doi.org/10.1111/j.1469-7610.2008.01951.x

Loukusa, S., & Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders*, *3*(4), 890–904. https://doi.org/10.1016/j.rasd.2009.05.002

Loy, J. E., Rohde, H., & Corley, M. (2017). Effects of Disfluency in Online Interpretation of Deception. *Cognitive Science, 41*, 1434–1456. https://doi.org/10.1111/cogs.12378

Lunsford, R., Heeman, P. A., Black, L., & van Santen, J. P. H. (2010). Autism and the use of fillers: Differences between `um' and `uh'. *DiSS-LPSS*, (September), 107–110.

Luyster, R. J., Seery, A., Talbott, M. R., & Tager-Flusberg, H. (2011). Identifying early-risk markers and developmental trajectories for language impairment in neurodevelopmental disorders. *Developmental Disabilities Research Reviews*, *17*(2), 151–159. https://doi.org/10.1002/ddrr.1109

MacFarlane, H., Gorman, K., Ingham, R., Hill, A. P., Papadakis, K., Kiss, G., & Van Santen, J. (2017). Quantitative analysis of disfluency in children with autism spectrum disorder or language impairment. *PLoS ONE*, *12*(3), 1–20. https://doi.org/10.1371/journal.pone.0173936

MacGregor, L. J., Corley, M., & Donaldson, D. I. (2009). Not all disfluencies are are equal: The effects of disfluent repetitions on language comprehension. *Brain and Language*, *111*(1), 36–45. https://doi.org/10.1016/j.bandl.2009.07.003

Magiati, I., Tay, X. W., & Howlin, P. (2014). Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. *Clinical Psychology Review, 34*(1), 78–86. https://doi.org/10.1016/j.cpr.2013.11.002

Manandhar, M., Hawkes, S., Buse, K., Nosrati, E., & Magar, V. (2018). Gender, health and the 2030 agenda for sustainable development. *Bulletin of the World Health Organization*, *96*(9), 644–653. https://doi.org/10.2471/BLT.18.211607

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language*, *85*(3), 451–466. https://doi.org/10.1016/S0093-934X(03)00070-1

Martin, J. G., & Strange, W. (1968). The perception of hesitation in spontaneous speech. *Perception and Psychophysics, 3,* 427-438.

Matthews, D., Biney, H., & Abbot-Smith, K. (2018). Individual Differences in Children's Pragmatic Ability: A Review of Associations with Formal Language, Social Cognition, and Executive Functions. *Language Learning and Development*, *14*(3), 186–223. https://doi.org/10.1080/15475441.2018.1455584

McCrimmon, A., & Smith, A. (2012). Review of the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II). *Journal Of Psychoeducational Assessment*, *31*(3), 337-341. doi: 10.1177/0734282912467756

McGregor, K. K., & Hadden, R. R. (2018). Brief Report: "Um" Fillers Distinguish Children With and Without ASD. *Journal of Autism and Developmental Disorders*, *0*(0), 0. https://doi.org/10.1007/s10803-018-3736-1

Merlo, S., & Mansur, L. L. (2004). Descriptive discourse: Topic familiarity and disfluencies. *Journal of Communication Disorders*, *37*(6), 489–503. https://doi.org/10.1016/j.jcomdis.2004.03.002

Metz, M. J., & James, L. E. (2018). Specific effects of the Trier Social Stress Test on speech fluency in young and older adults. *Aging, Neuropsychology, and Cognition*, *00*(00), 1–19. https://doi.org/10.1080/13825585.2018.1503639

Navarretta, C. (2015). The Functions of Fillers , Filled Pauses and Co-Occurring Gestures in Danish Dyadic Conversations. *3rd European Symposium on Multimodal Communicaiton*, 55–61.

Nilsson, K. K., & de López, K. J. (2016). Theory of Mind in Children With Specific Language Impairment: A Systematic Review and Meta-Analysis. *Child Development*, *87*(1), 143–153. https://doi.org/10.1111/cdev.12462

O'Connell, D. C., & Kowal, S. (2005). Uh and um revisited: Are they interjections for signaling delay? *Journal of Psycholinguistic Research*, *34*(6), 555–576. https://doi.org/10.1007/s10936-005-9164-3

Orena, A. J., & White, K. S. (2015). I Forget What That's Called! Children's Online Processing of Disfluencies Depends on Speaker Knowledge. *Child Development, 86*(6), 1701–1709. https://doi.org/10.1111/cdev.12421

Owens, S. J., Thacker, J. M., & Graham, S. A. (2018). Disfluencies signal reference to novel objects for adults but not children. *Journal of Child Language*, *45*(03), 581–609. https://doi.org/10.1017/S0305000917000368

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 32*(7), 1081–105. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1787138

Ozonoff, S., Young, G. S., Belding, A., Hill, M., Hill, A., Hutman, T., … Iosif, A. M. (2014). The broader autism phenotype in infancy: When does it emerge? *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(4), 398-407.e2. https://doi.org/10.1016/j.jaac.2013.12.020

Panico, J., Healey, E. C., Brouwer, K., & Susca, M. (2005). Listener perceptions of stuttering across two presentation modes: A quantitative and qualitative approach. *Journal of Fluency Disorders, 30*(1), 65–85. https://doi.org/10.1016/j.jfludis.2005.01.003

Parish-morris, J., Liberman, M. Y., Cieri, C., Herrington, J. D., Yerys, B. E., Bateman, L., … Schultz, R. T. (2017). Linguistic camouflage in girls with autism spectrum disorder, 1–12. https://doi.org/10.1186/s13229-017-0164-6

Parish-Morris, J., Liberman, M. Y., Cieri, C., Herrington, J. D., Yerys, B. E., Bateman, L., … Schultz, R. T. (2017). Linguistic camouflage in girls with autism spectrum disorder. *Molecular Autism*, *8*(1), 1–12. https://doi.org/10.1186/s13229-017-0164-6

Parsons, L., Cordier, R., Munro, N., Joosten, A., & Speyer, R. (2017). A systematic review of pragmatic language interventions for children with autism spectrum disorder. *PLoS ONE, 12* (4), e0172242. https://doi.org/10.1371/journal.pone.0172242

Perkins, W.H., Kent, R.D., & Curlee, R.F. (1991). A theory of neuropsycholinguistic function in stuttering. *Journal of Speech and Hearing Research*, 34, 734–752.

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

Pinker, S. (2013). *Language, Cognition, and Human Nature: Selected Articles* . New York, NY: Oxford University Press.

Postma, A., & Kolk, H. (1993). The covert repair hypothesis: Prearticulatory repair processes in normal and stuttered disfluencies. *Journal of Speech and Hearing Research, 36*, 472–487.

Postma, A., Kolk, H., & Povel, D. J. (1990). On the relation among speech errors, disfluencies, and self-repairs. *Language and Speech, 33*(1), 19–29. https://doi.org/10.1177/002383099003300102

Rayson, P., Leech, G. N., & Hodges, M. (2010). Social Differentiation in the Use of English Vocabulary: Some Analyses of the Conversational Component of the British National Corpus. *International Journal of Corpus Linguistics*, *2*(1), 133–152. https://doi.org/10.1075/ijcl.2.1.07ray

Ring, H., Woodbury-Smith, M., Watson, P., Wheelwright, S., & Baron-Cohen, S. (2008). Clinical heterogeneity among people with high functioning autism spectrum conditions: Evidence favouring a continuous severity gradient. *Behavioral and Brain Functions*, *4*, 1–6. https://doi.org/10.1186/1744-9081-4-11

Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire (SCQ).* Los Angeles: Western Psychological Services.

Rynkiewicz, A., Schuller, B., Marchi, E., Piana, S., Camurri, A., Lassalle, A., & Baron-Cohen, S. (2016). An investigation of the “female camouflage effect” in autism using a computerized ADOS-2 and a test of sex/gender differences. *Molecular Autism*, *7*(1), 1–8. https://doi.org/10.1186/s13229-016-0073-0

Sasson, N. J., Faso, D. J., Nugent, J., Lovell, S., Kennedy, D. P., & Grossman, R. B. (2017). Neurotypical Peers are Less Willing to Interact with Those with Autism based on Thin Slice Judgments. *Nature Publishing Group*, (February), 1–10. https://doi.org/10.1038/srep40700

Schegloff, E. A. (2010). Some Other “Uh(m)”s. *Discourse Processes*, *47*(2), 130–174. https://doi.org/10.1080/01638530903223380

Scott, K. S. (2015). Dysfluency in Autism Spectrum Disorders. *Procedia - Social and Behavioral Sciences, 193*, 239–245. https://doi.org/10.1016/j.sbspro.2015.03.266

Semel, E., Wiig, E.H., & Secord, W. (2000). *Clinical Evaluation of Language Fundamentals – Third Edition UK* (CELF-3 UK). London: The Psychological Corporation.

Shafto, M. A., & Tyler, L. K. (2014). Language in the aging brain: The network dynamics of cognitiv e decline and preservation. *Science*, *346*(6209), 583–588. https://doi.org/10.1126/science.1254404

Shriberg, E. (1994). *Preliminaries to a theory of speech disfluencies*. *Ph.D. dissertation*. University of California, Berkeley.

Shriberg, Elizabeth. (1996). Disfluencies in Switchboard. *Proceedings of International Conference on Spoken Language Processing*, 11–14.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and Prosody Characteristics of Adolescents and Adults With High-Functioning Autism and Asperger Syndrome. *Journal of Speech Language and Hearing Research*, *44*(5), 1097. https://doi.org/10.1044/1092-4388(2001/087)

Skuse, D. H., Mandy, W. P. L., & Scourfield, J. (2005). Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *British Journal of Psychiatry*, *187*(06), 568–572. https://doi.org/10.1192/bjp.187.6.568

Skuse, D.,Warrington, R., Bishop, D., et al (2004). The developmental, dimensional and diagnostic interview (3di): a novel computerized assessment for autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 548-558. https://doi.org/10.1097/00004583-200405000-00008

Stefanatos, G. A., & Baron, I. S. (2011). The ontogenesis of language impairment in autism: A Neuropsychological perspective. *Neuropsychology Review*, *21*(3), 252–270. https://doi.org/10.1007/s11065-011-9178-6

Suh, J., Eigsti, I. M., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an Autism Spectrum Disorder (ASD). *Journal of Autism and Developmental Disorders*, *44*(7), 1681–1694. https://doi.org/10.1007/s10803-014-2042-9

Svartvik, J., & Quirk, R. (1980). *A corpus of English conversation*. Lund: Gleerup.

Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. *Journal of Autism and Developmental Disorders*, *23*(2), 309–322. https://doi.org/10.1007/BF01046222

Tottie, G. (2011). Uh and Um as sociolinguistic markers in British Englishany. *International Journal of Corpus Linguistics*, *16*(2), 173–197. https://doi.org/10.1075/ijcl.16.2.02tot

Walker, E. J., Risko, E. F., & Kingstone, A. (2014). Fillers as Signals: Evidence From a Question-Answering Paradigm. *Discourse Processes*. Taylor & Francis. https://doi.org/10.1080/0163853X.2013.862478

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The structure and nature of the broader autism phenotype in a non-clinical sample. *Journal of Psychopathology and Behavioral Assessment*, *33*(4), 459–469. https://doi.org/10.1007/s10862-011-9259-0

Wechsler D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition* (WASI-II). San Antonio, TX: NCS Pearson.

Whitehouse, A. J. O., Barry, J. G., & Bishop, D. J. M. (2007). The broader language phenotype of autism: A comparison with specific language impairment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 48*(8), 822–830. https://doi.org/10.1111/j.1469-7610.2007.01765.x

World Health Organization. (2018). *International statistical classification of diseases and related health problems* (11th Revision). Retrieved from https://icd.who.int/browse11/l-m/en

Wiggins, L. D., Robins, D. L., Adamson, L. B., Bakeman, R., & Henrich, C. C. (2012). Support for a dimensional view of autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders*, *42*(2), 191–200. https://doi.org/10.1007/s10803-011-1230-0

Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Yaruss, J. S., Newman, R. M., & Flora, T. (1999). Language and disfluency in nonstuttering children's conversational speech. *Journal of Fluency Disorders*, *24*(3), 185–207. https://doi.org/10.1016/S0094-730X(99)00009-1

Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The Use of Two Language Tests to Identify Pragmatic Language Problems in Children With Autism Spectrum Disorders. *Language Speech and Hearing Services in Schools, 36*(1), 62. https://doi.org/10.1044/0161-1461(2005/006)