From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome

Coralie Chevallier[1], Ira Noveck[1], Francesca Happé[2], Deirdre Wilson[3, 4]

Accepted by *Research in Autism Spectrum Disorders.*

[1]Laboratoire Langage, Cerveau et Cognition, CNRS - Université de Lyon

[2]Institute of Psychiatry, King's College London

[3]Department of Phonetics and Linguistics, University College London

[4]CSMN, University of Oslo, Norway

Abstract

We report findings concerning the understanding of prosody in Asperger Syndrome (AS), a topic which has attracted little attention and led to contradictory results. Ability to understand grammatical prosody was tested in three novel experiments. Experiment 1 assessed the interpretation of word stress, Experiment 2 focused on grammatical pauses, and Experiment 3 tested the discrimination of the question contour. Acoustic tasks were also used to assess the perception of pitch, duration, intensity and prosodic contours. AS participants performed as well as typically developing controls in all our tasks. This provides support in favour of the view that grammatical prosody is spared in Asperger Syndrome.

Keywords: Asperger Syndrome; Prosody; Grammar; Language.

It is commonly reported that Autism Spectrum Disorders (ASDs) are linked with language deficits and in a way that varies with the intensity of the disorder (Pry, Petersen, & Baghdadli, 2005). For instance, in classical autism, approximately 50% of the population never manage to develop a functional language (Bailey, Phillips, & Rutter, 1996; Rapin, 1991). In contrast, individuals with milder forms of autism do acquire language but with some delay (e.g., children with High Functioning Autism, HFA). Finally, individuals with Asperger Syndrome (AS) – the least severe form of the disorder – speak fluently and go through the typical stages of language acquisition (APA, 1994). Regardless of the severity of the disorder, most researchers agree that the language deficit in AS and HFA mainly affects pragmatics while sparing formal aspects of language (e.g., Attwood, 2006; Landa, 2000). This hypothesis has been mainly supported by data indicating that syntax is generally spared (e.g., Ghaziuddin et al., 2000) while the use and understanding of various phenomena relying on pragmatic enrichment are not (e.g., non literal language, U. Frith & Happé, 1994; Happé, 1993; Martin & McDonald, 2004); (or humour, Lyons & Fitzgerald, 2004; Werth, Perkins, & Boucher, 2001). Prosody is a topic that has obvious relevance to the above dissociation because it serves both grammatical and pragmatic functions, which could be dissociated among listeners with an ASD.

“Prosody” relates to a variety of phenomena (including intonation, stress, rhythm, etc.) serving grammatical as well as pragmatic functions through variations in pitch, intensity and duration (Cruttenden, 1997; Wilson & Wharton, 2006). At the grammatical level, stress and rhythm help the hearer chunk the speech flow into words and clauses (e.g., “Dragonfly and carrot” has a different rhythm and stress pattern than “Dragon, fly, and carrot”), intonation patterns help to identify sentence type (e.g., a final rising contour signals a question and a falling contour signals a declarative), and word stress enables one to differentiate between nouns and verbs (e.g., PREsent vs. preSENT) (for a review, see Cruttenden, 1997; Cutler, Dahan, & van Donselaar, 1997)[1]. At the pragmatic level, many cues to the meaning intended by the speaker are conveyed through subtle prosodic changes (e.g., clues for turn-taking, speaker’s attitude, speaker’s emotional state, focus and contrast, etc.; for a review, see Cutler, Dahan, & van Donselaar, 1997; House, 2006).

Although the consensus view is that one ought to find intact grammatical competence and a lack of pragmatic competence among HFA and AS participants, existing results are not so straightforward. As we describe in detail later, some studies comparing HFA and AS participants to matched controls reported deficits in the ability to perceive *grammatical* prosodic cues, while others did not identify such group differences. Given this lack of clarity, one reasonable first step in establishing a grammar/pragmatics distinction with respect to prosody is to rule out the existence of a potential deficit in grammatical prosody (which could then impact performance at the pragmatic level). In the present work, we address this point and assess the grammatical function of prosody in AS. In what follows, we briefly review the literature on the production and perception of prosody in HFA and AS. While keeping the grammar/pragmatics distinction in mind, we first concentrate on the production side and then turn to the reception of prosody. To anticipate, the literature on production generally conforms to the consensus view that grammar is spared in HFA and AS while pragmatics is not. However, the literature on reception is less clear-cut and requires more investigation. We thus put forward a novel experimental paradigm testing the understanding of grammatical prosody in AS and addressing some of the methodological issues found in previous experiments.

Producing prosody

Given the abundance of evidence highlighting prosodic abnormalities in ASDs, this area of research has been greatly under-explored (for a review, see McCann & Peppé, 2003). Yet, both Kanner's (1943) and Asperger's (1944) initial descriptions of the disorder mentioned abnormal prosody, using adjectives such as “odd”, “monotonous”, “singsong”, “unmodulated”, etc. Asperger, whom we quote at length below, clearly emphasises these issues in his seminal paper (translated in U. Frith, 1991):

> The abnormalities differ, of course, from case to case. Sometimes the voice is soft and far way, sometimes it sounds refined and nasal but sometimes it is too shrill and ear-splitting. In yet other cases, the voice drones on in a sing-song and does not even go down at the end of a sentence. However many possibilities there are, they all have one thing in common: the language feels unnatural… (p. 70)

Autobiographies and parental reports also frequently mention aberrant prosody as well as difficulties in interpreting prosodic cues. For instance, Temple Grandin (Grandin, 1986) reports on the strangeness of her voice and on the way this impacted her social life:

> Years later, I was shocked to learn that my speech still had minor abnormalities. I was not aware of the persistence of the hesitation and occasional flat tone of my speech. (...) I was aware that sometimes people didn’t want to talk to me, but I didn’t know why. (p. 87)

Recent reports in the literature further confirm that a) deficits in producing prosody are present in many patients (almost 50% of ASD individuals according to Paul, Augustyn, Klin, & Volkmar, 2005); b) they persist even when other areas of language improve (Paul, Augustyn, Klin, & Volkmar, 2005; Shriberg et al., 2001); and c) they “constitute one of the most significant obstacles to [the person's] social integration and vocational acceptance” (Paul et al., 2005, p. 205). In line with results in other areas of language, it appears that the pragmatic function of prosody is more affected than its linguistic function. For instance, AS individuals are able to place grammatical pauses correctly but use non grammatical pauses less often (Fine, Bartolucci, Ginsberg, & Szatmari, 1991; Shriberg et al., 2001; Thurber & Tager-Flusberg, 1993). Similarly, word stress is produced properly but contrastive stress is not (Fine, Bartolucci, Ginsberg, & Szatmari, 1991; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007; Shriberg et al., 2001). Furthermore, they tend to stress old and new information equally often (McCaleb & Prizant, 1985), thereby giving evidence of pragmatic difficulties. Finally, results regarding the production of the contour associated to questions and declaratives are less clear-cut. Peppé et al. (2007) found that HFA children sounded as if they were asking a question when a statement was required. However, using a similar design, Paul et al. (2005) found no difference between HFA / AS teenagers and matched controls.

Perceiving prosody

Autobiographical reports such as Donna Williams’ (1994) also mention serious difficulties, on the perception side, in dealing with prosodic cues:

> 'Speak to me through my words,’ I asked Dr. Marek. I wanted to cut down the struggle in putting mental pictures into words.
>
> ‘Can you take the dancing out of your voice and not pull faces so you don’t distract me from what you’re saying?’ (p. 95)

However, the perception of prosody has received less attention than the production side, except in the case of affect, where deficits have been identified (Golan, Baron-Cohen, Hill, & Rutherford, 2007; Rutherford, Baron-Cohen, & Wheelwright, 2002). Although limited in the range of pragmatic functions they assess, studies on the perception of pragmatic prosodic cues match the results reported for production. Indeed, both Paul et al. (2005) and Peppé et al. (2007) found that contrastive stress was understood inappropriately in HFA and AS participants and studies on the perception of irony in a speaker’s voice also reported issues (Rutherford, Baron-Cohen, & Wheelwright, 2002; Wang, Lee, Sigman, & Dapretto, 2006). As far as grammatical aspects are concerned, the literature is rather scarce and existing results are contradictory. Some studies find no obvious difference in the interpretation of grammatical prosody between participants with AS / HFA and controls, while others do reveal problems.

More specifically, in a study assessing a group of 27 teenagers with ASD (including 24 participants with AS or HFA and 3 with Pervasive Developmental Disorder; Mean age: 16.8), Paul et al. (2005) found no difference in syntactic chunking. In this experiment, the presence or absence of a pause affected syntactic parsing, and, consequently, the overall interpretation of the utterance. For instance, “Ellen, the dentist is here” differs from “Ellen, the dentist, is here”. In the former case, Ellen is the addressee of the utterance, she is the person who is told that the dentist is here; and in the latter case, “Ellen” actually refers to the dentist. Paul et al’s results thus indicate that AS / HFA participants are able to use pauses to interpret an utterance. Similarly, Peppé and her colleagues (2007) found that word chunking (i.e., “Chocolate, Cake and Bun” vs. “Chocolate-cake and Bun”) was unimpaired in a study involving 31 children with HFA (age range: 6.1 -13.6).

While these findings point to a common competence, results from the two above studies differ with respect to the ability to distinguish between interrogatives and declaratives. Indeed, while Paul and her colleagues found no difference between the group of HFA / AS adolescents and the controls, Peppé et al. (2007) reported that HFA children tended to wrongly judge questions as statements. Here it is worth mentioning another experiment (Erwin et al., 1991), testing 11 HFA adults, in which Event Related Potentials were recorded in response to the rare presentation of declaratives (appearing in 20% of the trials) amongst frequent interrogatives (appearing in 80% of the trials). One specific ERP component, the P3 potential, is known to be elicited in response to rare stimuli. It was therefore postulated that, provided that participants differentiate the two sentence types, the P3 should be greater in the rare trials involving declaratives. To the authors’ surprise, the data indicated remarkably normal P3 and behavioural processing of prosodic stimuli, a result which goes in the same direction as Paul et al.’s. In the case of grammatical stress, Paul et al.’s study, the only one (to the best of our knowledge), reports poorer performance in the HFA /AS group. However, this finding also requires some further examination since it is based on a relatively high p value ($p = .12$).

Overall, the literature on the perception of grammatical prosody in HFA and AS is not conclusive. Arguably, the inconsistency in the results reported above could be due to the fact that methodologies differ greatly, making cross-study comparisons difficult (e.g., Erwin et al. tested adults, Paul et al. tested teenagers and Peppé et al. tested children). Moreover, the results presented by Paul et al. (2005) are difficult to interpret due to the presence of ceiling effects on most tasks. Although very rigorous in its methods and analysis, a close look at Peppé et al.'s study reveals potential interpretational difficulties for the participants, thus casting doubt on the reliability of the results. For instance, children were asked to distinguish interrogatives from declaratives by saying whether the speaker was "telling something" or "asking something". But since asking is a case of telling (strictly speaking someone *asking* is also conveying some information and uttering a sentence, thereby *telling* something), the two choices are not mutually incompatible under a literal interpretation. This may have been an issue for the HFA children, who tend to be quite literal minded (Happé, 1993). Furthermore, the studies mentioned so far concentrate on the analysis of accuracy rates and do not explore reaction times, which can be potentially revealing in identifying differences in strategies between AS participants and controls. Finally, the studies presented above might have been more convincing if an extra control involving physical measures had been added; that is, none of the above studies takes steps to ensure that primary perceptual abilities, on which the perception of prosody is based, are intact.

As noted above, prosodic cues are conveyed through basic acoustic changes such as variations in pitch, intensity and duration. The ability to detect these changes, together with normal hearing capacities, is of course crucial to the deployment of higher order prosodic skills. In individuals with an ASD it appears that peripheral hearing sensitivity is generally normal (Gravel, Dunn, Lee, & Ellis, 2003; Tharpe et al., 2006); however, auditory processing disorders (Bruneau, Bonnet-Brilhault, Gomot, Adrien, & Barthélémy, 2003; Siegal & Blades, 2003) and sound sensitivities (Kern et al., 2006) are commonly reported. Similarly, difficulties in appropriately allocating attention to some auditory stimuli have been reported (Ceponiene et al., 2003; Lepistö et al., 2005; Lepistö et al., 2006) and recent fMRI and PET findings have identified abnormalities in the network involved in processing voice stimuli (Boddaert et al., 2004; Gervais et al., 2004; Gomot et al., 2006). Given the possibility that some individuals with ASD suffer from auditory processing issues, it seems imperative that basic auditory skills be fully checked before assessing higher level prosodic abilities.

In the current study, the above issues are addressed. In each experiment, several control conditions were included, participants were exposed to repeated trials in order to ensure that the effects were robust across varied material types, and the experiments were devised so that both accuracy rates and reaction times could be measured. Participants were also chosen in one age group only (i.e. teenagers); furthermore, in order to minimise problems due to the heterogeneity of the Autism Spectrum, we concentrated on individuals with Asperger Syndrome. AS participants were then carefully matched to typically developing participants both according to their chronological age and to their verbal mental age. Importantly, we first assess the ability of adolescents with AS to perceive pitch, duration, intensity and prosodic contours with acoustic tasks. We then go on to investigate their ability to understand grammatical prosody using three novel experiments – one applied to an isolated word, a second to a sequence of words and a third to a whole utterance. More precisely, Experiment 1 assesses the interpretation of word stress (e.g., "PREsent" vs. "preSENT"), Experiment 2 assesses the ability to take rhythm into account in order to chunk a sequence of several words (e.g., "Dragonfly and Carrot" vs. "Dragon, fly, and Carrot") and finally, Experiment 3 assesses the discrimination of declaratives and interrogatives (e.g., "This animal is a monkey." vs. "This animal is a monkey?").

General method

This section presents a) the battery of tests that were used in the study, b) details of the recording procedure for our auditory stimuli and c) information about the general testing procedure common to all three experiments. Finally, we provide details of the participants who were tested in the study.

Materials

The British Picture Vocabulary Scales-II (Dunn, Dunn, & Whetton, 1997). Assesses the participant's verbal mental age by measuring receptive vocabulary. In the test, the child is asked to point to one picture (out of four), which corresponds to the word spoken by the experimenter.

Memory for contours. In this novel task, we present the participant with six pairs of five note melodies in a random order (Mel1: CGEGC, Mel2: CFDFC, Mel3: CDEAA, Mel4: CGGDC, Mel5: CCEEG, Mel6: GEEEC). For each pair, the participant has to decide whether the two melodies sound the same or different. Three pairs are identical (1-1, 4-4, 6-6) and three sound different (1-2, 5–6, 3-4).

Dinos. The three "Dino" tasks, designed and programmed by Dorothy Bishop, assess the participant's ability to discriminate duration, intensity and frequency (for previous studies using the Dinos task, see, e.g. Sutcliffe & Bishop, 2005). Two dinosaurs each make a sound separated by a 500ms interval in the intensity and duration tasks and by a 480ms interval in the frequency task. The child then has to decide which dinosaur is making the longest, loudest or highest sound (depending on whether she is completing the duration, intensity or frequency task respectively). Correct responses are reinforced with a small icon on the screen and a cheerful noise, and wrong answers with a cross and a sigh noise. The next trial starts after a

500ms interval. All three tasks are based on a “more virulent” PEST procedure (Findlay, 1978), which adaptively alters the gap separating the two sounds. Initially, the participant has to make very easy discriminations, and larger step sizes are subsequently used to increase difficulty level until an error is made. When an error is made, the discrimination is made easier. The task is stopped after 6 reversals have occurred or a maximum of 40 trials has been completed. The PEST procedure is set to converge on the 75% correct point and the threshold is taken as the average target across the last four reversals in the track. Note that low thresholds are indicative of optimal performance.

Auditory stimuli for Experiments 1, 2 and 3. The stimuli were recorded in an anechoic chamber at University College London with the help of a professional acoustician. The speaker was a native male speaker of Southern standard British English, trained to record auditory stimuli. He sat in an armchair equipped with a headrest ensuring that the distance between his mouth and the microphone remained constant. The microphone (Bruel & Kjaer 2231 Sound Level Meter fitted with a Type 4165 Microphone) was linked up to a Sony DAT reader connected to a PC. The recordings were made in a mono format, using a 44.1 kHz sampling rate. The items to be read were presented on a suspended computer screen using ProRec version 1.0© (Huckvale, 2003). The wave files were then segmented using the Speech Filing System© (Huckvale, 2004) and a 100 ms silence was inserted immediately before and after the sound signal.

General testing procedure

Written parental consent was obtained prior to the testing phase and children were then also asked whether they agreed to take part in the study. Pupils were seen individually at school during two 35-minute sessions. The experiments testing perception and grammatical prosody were presented using a laptop and the sounds were played through Sennheiser headphones (which were calibrated for consistency of dB before use). None of the children had problems agreeing to wear the headphones and all of them were comfortable with computers (this computerized environment is predictable and less socially demanding than direct one-to-one interactions, and is thus less stressful for AS children,). Experiments 1, 2 and 3 were presented using E-prime (which was also used to measure accuracy rates and reaction times) and all the trials were presented in a random order. Each trial started with a 1000ms “Listen carefully” screen followed by an auditory stimulus. The participant then had to answer using one of two response keys (E and P counterbalanced) and the next trial started 1000ms later. In each experiment, the instructions – presented on the screen – were read out to the participants and a training phase followed. When participants were halfway through each task, the message *"You're half way through!"* was displayed so that they had the opportunity to take a break. More specific details of the procedure will be provided in the description of each experiment.

Participants

Thirty-four male adolescents (17 with AS and 17 Typically Developing, henceforth TD) took part in the experiments presented in this paper. The pupils with AS were at a special education school in Somerset (England) which requires formal diagnosis of AS according to standard clinical criteria (APA, 1994). The diagnostic information was gathered from school files of documented medical diagnoses made by a clinical psychologist and/or psychiatrist. The controls were seen in a regular school in the Sheffield area. TD and AS participants all spoke English at home, and none had any significant hearing loss, visual impairment, or major physical disability. The control participants were matched on chronological age (AS-Mean = 13;8 , TD-Mean = 14;2, *t*(32) = .77; *p* = .41) and verbal mental age (Standardised BPVS score: AS-Mean = 106, TD-Mean = 99; *t*(32) = 1.20, *p* = .16, see Table 1 for detailed information).

	TD participants		Participants with AS		
	Mean (SD)	Range	Mean (SD)	Range	t(df); *p*
Age	14;2 (1;7)	11;6-16;8	13;8 (1;11)	11;1-17;10	t(32) = .77 ; *p* = .41
BPVS score	99 (14)	76-128	106 (20)	78-145	t(32) = 1.20 ; *p* = .16

Table 1. Participants' age and BPVS score.

Finally, the Dinos tasks revealed no group difference in any of the conditions, indicating that the discrimination thresholds for duration, intensity and frequency gaps were similar in both groups (see Table 2).

	TD	AS	t (dl) ; *p*
Intensity – Mean (SD)	2.70dB (1.62dB)	2.43dB (1.62dB)	t(32) = .35 ; *p* = .73
Duration – Mean (SD)	72 ms (64ms)	40ms (24ms)	t(32) = 1.86 ; *p* = .07
Frequency – Mean (SD)	142Hz (112Hz)	122Hz (92Hz)	t(32) = .64 ; *p* = .52

Table 2. Mean threshold values (and standard deviations) for the Intensity, Duration and Frequency tasks.

The existing literature on pitch discrimination in ASDs points in the same direction and provides quite unanimous evidence that individuals on the spectrum generally perform as well as–or better than - controls. This is true for tasks involving pitch discrimination (e.g., Bonnel et al., 2003; Heaton, 2003, , 2005; O'Riordan & Passetti, 2006), pitch naming (Heaton, 2003), and pitch memory (Heaton, 2003). The high prevalence of absolute pitch (Happé, 1999) and musical savants (Hermelin, 2001) within the autistic population also supports this view. As far as loudness and duration are concerned, the results are scarce and contradictory, making it hard to draw any firm conclusion.

Having checked that both groups have similar perceptual abilities, we now turn to the three novel experiments designed to test the understanding of grammatical prosody. Experiments 1, 2 and 3 assess the ability to deal with lexical stress, to chunk compounds, and to distinguish questions from declaratives, respectively.

Data analysis for Experiments 1-3

The data was analysed using SPSS 13 for Mac OSX. By convention, we refer to F-values obtained with participants as the random factor as F_1 (or t_1), while F-values obtained with items as the random factor are referred to as F_2 (or t_2) for all the analyses presented in this paper. All p-values assume a two-tailed test. For all reaction time analysis, a log transformation was carried out beforehand to improve the conformity of the data to the standard assumptions of ANOVA (e.g., Howell, 1997). Reaction times of more than three standard deviations from the mean were considered outliers and were excluded from both the reaction time and the choice proportion analysis. Moreover, only correct responses were retained in the reaction time analysis.

Experiment 1 – Lexical Stress

In this experiment, we assess the participant's ability to select the most appropriate pronunciation of an utterance on the basis of the stress pattern assigned to a disyllabic noun or verb. The target items used in the experiment belong to pairs of Noun-Verb homographs with different stress patterns (e.g., "He got the best PREsent he could dream of." - "I preSENT the late-night news."). An equal number of control items, which do not belong to such pairs of homographs, was added (e.g., "He got the best PUzzle he could dream of." - "I Edit the late-night news."). Participants heard the same sentence pronounced twice – once with a correct stress placement and once with an incorrect stress placement[2] – and were instructed to decide whether the "first" utterance or the "second" one was pronounced best, using the corresponding response keys (counterbalanced). The experiment was preceded by a two-trial training phase.

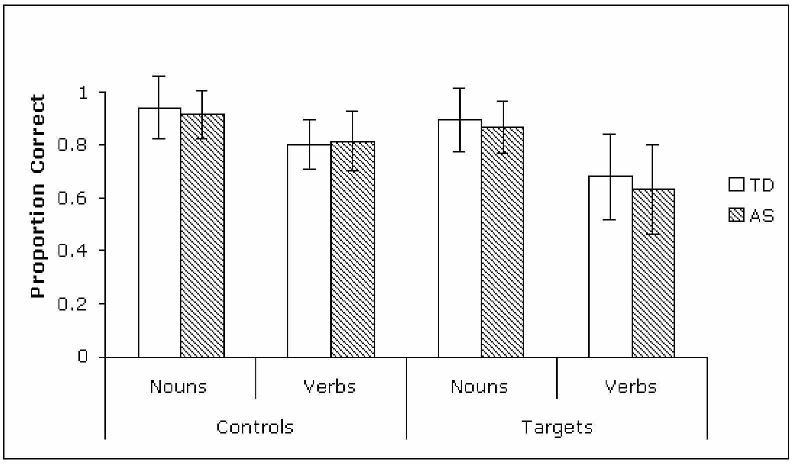
Method

Participants. Two AS participants (chronological age: 16;10 and 12;01, VMA: 91 and 94) did not take part in this experiment because they asked to go back to class before having gone through all three experiments. We thus tested 32 pupils (17 TD, 15 AS).

Material and design. This experiment is based on a 2 (Item-Type: Target, Control) by 2 (Grammatical-Category: Noun, Verb) design. The 20 most frequent Noun-Verb homonyms in British English were selected. These were presented to 15 adults who had to decide whether the pronunciation was strange or fine. Four items were subsequently removed so that only those pairs for which both the noun and the verb triggered rates of correct responses above chance were included in the experiment. Each item could appear in one of four conditions (Target-Noun, Target-Verb, Control-Noun, Control-Verb). Four lists were made up in this way; they included 16 target items (8 Nouns, 8 Verbs) and 16 control items (8 Nouns, 8 Verbs), with each item appearing twice: once pronounced correctly and once incorrectly.

Results and discussion

Results in all the control conditions were well above chance (the mean rates of correct answers ranging from 80.1% to 94.1%, see Figure 1, top). In the Target condition, Nouns also elicited good scores (89.5% in the TD group and 86.7% in the AS group) but Verbs were associated to comparatively lower performances (67.9% in the TD group and 63.3% in the AS group). However, scores were above chance in all the conditions and for both groups (all *t*s > 3.4, all *p*s <.005).



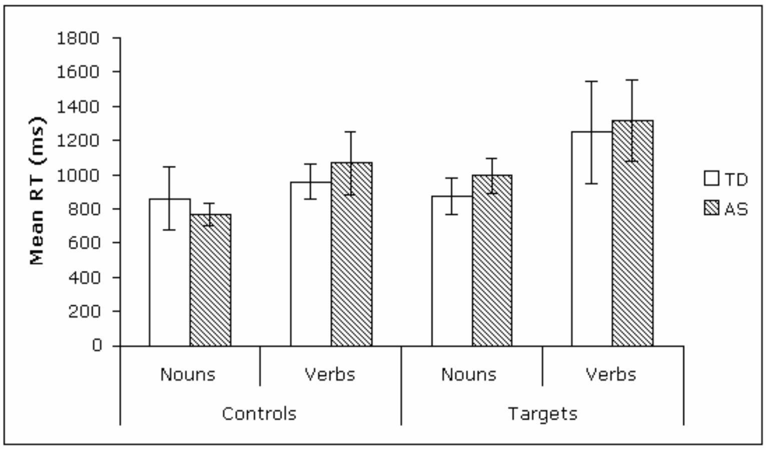


Figure 1. Proportion of correct answers (top) and reaction times (bottom) as a function of stimulus type (control, targets), grammatical category (nouns, verbs), and group (TD, AS).

We conducted a mixed repeated measures ANOVA with three factors: the within-subject factors “Item-Type” (2: Target, Control) and “Grammatical-Category” (2: Noun, Verb) and the between subject factor “Group” (2: TD, AS). Here, the dependent variable was the rate of correct responses. The ANOVA revealed a main effect of Item-Type ($F_1(1,30) = 13.4$, $p_1 < .001$; $F_2(1,15) = 7.5$, $p_2 < .02$) and a main effect of Grammatical-Category ($F_1(1,30) = 37.1$, $p_1 < .001$; $F_2(1,15) = 13.9$, $p_2 < .005$). However, there was no significant main effect of the group ($F_1(1,30) = 0.5$, $p_1 = .47$; $F_2(1,15) = 2.7$, $p_2 = .12$), no Item-Type X Group interaction ($F_1(1,30) = 0.3$, $p_1 = .56$; $F_2(1,15) = .29$, $p_2 = .60$), and no Grammatical-Category X Group interaction ($F_1(1,30) =.04$, $p_1 = .85$; $F_2(1,15) = .04$, $p_2 = .84$). This indicates that AS participants are no less accurate than the controls. We then ran *Post hoc* Bonferroni tests to check the origin of the main effects we reported above. These indicate that the main effect of Item-Type is due to the fact that targets are harder than controls (Target-Mean = 79.9, Control-Mean = 86.8; $p < .001$) and the main effect of the Grammatical-Category is due to verbs being harder than nouns (Noun-Mean = 90.5, Verb-Mean = 73.3; $p < .0001$). Finally, the Item-Type X Grammatical-Category interaction ($F_1(1,30) = 7.7$, $p_1 < .01$; $F_2(1,15) = 2.37$, $p_2 = .14$) is explained by the fact that target verbs are harder than all the other types of items (all *ps* < .001). A possible explanation as to why target nouns were easier than target verbs is that 94% of Nouns have a trochaic pattern, while 69% of verbs have an iambic pattern (Kelly & Bock, 1988), which implies that the stress pattern of nouns is much more predictable.

A repeated measures ANOVA including the same factors as those used in the analysis of accuracy rates with participants' reaction times as the dependent measure revealed the same effects (see Figure 1, bottom). There was a main effect of Item-Type ($F_1(1,30) = 10.0$, $p_1 < .004$; $F_2(1,15) = 14.8$, $p_2 < .005$), and a main effect of Grammatical-Category ($F_1(1,30) = 12.8$, $p_1 < .001$; $F_2(1,15) = 8.9$, $p_2 < .01$), but no significant main effect of the group ($F_1(1,30) = 0.2$, $p_1 = .66$; $F_2(1,15) = 1.8$, $p_2 = .20$), no Item-Type X Group interaction ($F_1(1,30) =.02$, $p_1 = .88$; F_2 $(1,15) = .06$, $p_2 = .81$), and no Grammatical-Category X Group interaction ($F_1(1,30) =.03$, $p_1 = .86$; $F_2(1,15) = .34$, $p_2 = .57$). Our results thus indicate that AS participants are able to use intonation to detect grammatical categories and that they do so in ways that are very similar to control participants.

Experiment 2 – Chunking compounds

In this Experiment, we assess the participant's ability to take rhythm into account in chunking sequences of two or three words appropriately and to associate the sequence to the right set of pictures. Three word-types were used: Compounds ("Dragonfly and carrot"), Split-compounds ("Dragon, fly and carrot") and Controls ("Fly, apple and carrot"). These stimuli appeared in three experimental conditions (see Figure 2).

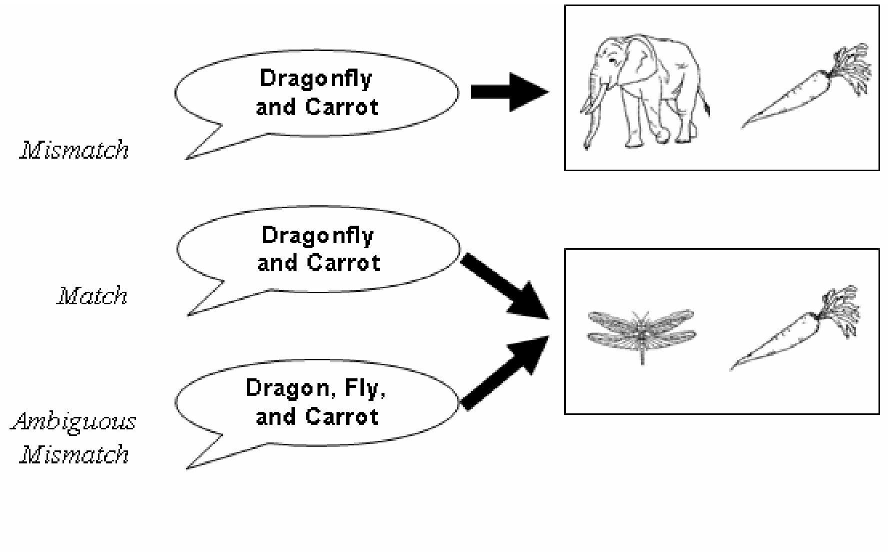


Figure 2. Experimental design for Experiment 2.

In the Ambiguous Mismatch condition, the participants heard a compound and saw the picture corresponding to the split-compound (or vice versa), e.g., "Dragonfly and carrot" was played out and was followed by pictures of a dragon, a fly and a carrot. In the Match condition, the right set of pictures was associated to the sequence, e.g., "Dragonfly and carrot" was followed by the picture of a dragonfly and a carrot. In the Mismatch condition, the sequence was followed by a set of pictures which clearly did not match, e.g., "Dragonfly and carrot" was followed by the picture of an elephant and a carrot. The participant first heard the description and then saw the set of pictures. They were instructed to decide whether the description was "right" or "wrong", using the appropriate response key (counterbalanced). The experiment was preceded by a three-trial training phase.

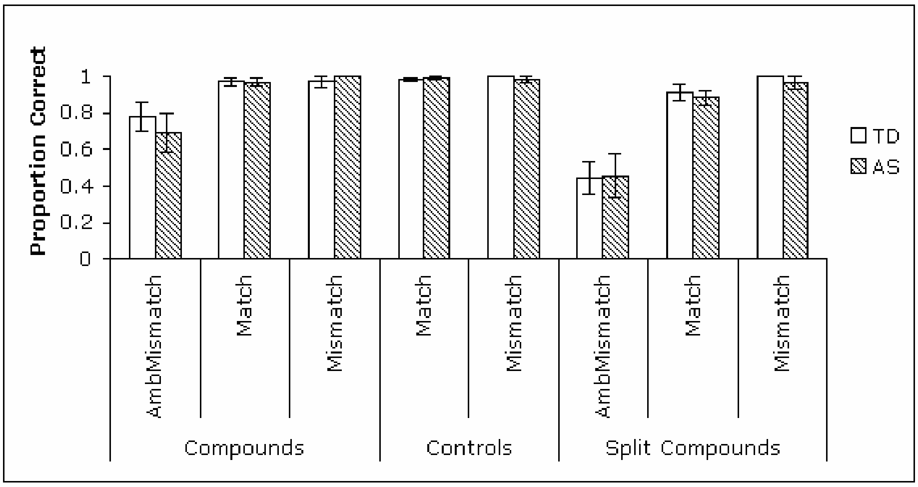
Methods

Participants. Two AS participants (chronological age: 16;10 and 17;10, VMA: 91 and 78) did not take part in this experiment because they asked to go back to class before having gone through all three experiments. We thus tested 32 pupils (17 TD, 15 AS).

Material and design. This experiment is based on a 3 (Word-Type: Compound, Split-Compound, Control) by 3 (Matching: Ambiguous Mismatch, Match, Mismatch) design. 40 Noun-Noun compounds were selected making sure that both constituent nouns and the resulting compounds could be represented on a picture. The Snodgrass database (Snodgrass & Vanderwart, 1980) and pictures drawn by a professional artist were used in this experiment. The artist was asked to use a similar drawing style to the one used in Snodgrass and we subsequently made sure that both sets were indistinguishable. Each item could be presented in two forms: as a compound (“Dragonfly”) or as a split compound (“Dragon”, “fly”). It was then followed by another word preceded by the “and” connective (e.g., “Dragonfly and carrot” or “Dragon, fly and carrot”). This last word was chosen from among the most frequent mono- bi- and tri- syllabic words in English, and so that all the items had the same total number of syllables. Forty control items were then created using the various constituents of the experimental items arranged in a random order, so that none of the combinations corresponded to a possible compound (e.g., “Headphones” and “Dragonfly” would be mixed into “Head, fly, and microscope”). Four lists were then made up, each including 10 compounds, 10 split-compounds and 12 control items, each item appearing in only one of its possible forms in each list. Half the items were included in the Match condition (expected answer: “Right”) the other half being equally divided between the Mismatch and the Ambiguous Mismatch conditions (expected answer: “Wrong”) so that the participants were supposed to answer “Right” and “Wrong” equally often.

Results and discussion

Means and standard deviations for targets judged to sound correct, for the three word types (Compound, Split-Compound and controls) and for the three Matching conditions (Match, Ambiguous-Mismatch, Mismatch) are shown in Figure 3 (top).



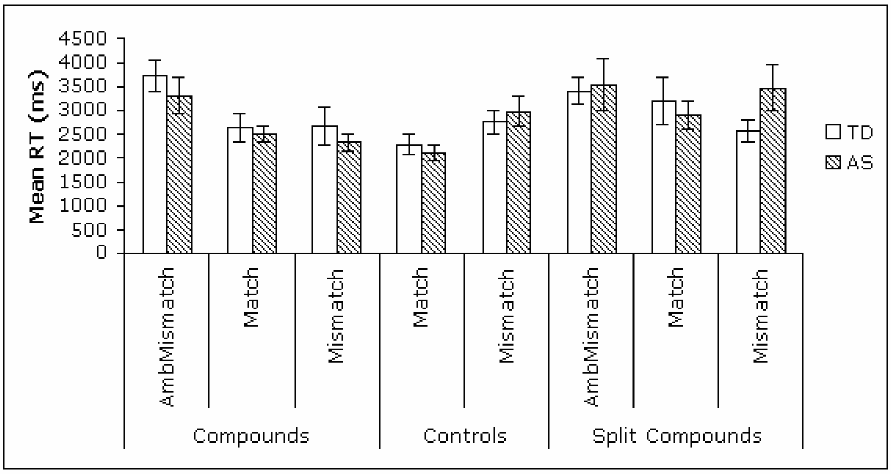


Figure 3. Proportion of correct answers (top) and reaction times (bottom) as a function of word types (Compound, Split-Compound and controls) and for the three Matching conditions (Match, Ambiguous-Mismatch, Mismatch), and group (TD, AS).

Since both groups obtained ceiling scores for the control items, we concentrate on the target word types (Compounds and Split-Compounds) in the analyses which follow. In the Match and Mismatch conditions, mean rates of correct responses are extremely high, ranging from 90.1% to 100% and thus appear not to cause any particular problems for either group of participants. In contrast, the Ambiguous-Mismatch condition gives rise to lower rates of correct responses and to extremely high standard deviations, especially in the Split-Compound condition. The rates observed in this condition (72.7% on average for Compounds and 42.2% for Split-Compounds) do not, however, reflect chance levels but rather the existence of two groups of participants: one consistently (and wrongly) accepting ambiguous mismatches and the other consistently (and rightly) rejecting them[3]. We conducted an ANOVA using repeated measures with three factors: the within-subject factors "Wordtype" (2: Compound, Split-Compound), and "Matching" (3: Match, Ambiguous-Mismatch, Mismatch) and the between-subject factor "Group" (2: TD, AS). Here, the dependent variable was the rate of correct responses. The ANOVA revealed a main effect of Wordtype ($F_1(1,30) = 31.4$, $p_1 < .0001$; $F_2(1,54) = 22.2$, $p_2 < .0001$), and of Matching ($F_1(2,60) = 27.3$, $p_1 < .0001$; $F_2(2,54) = 92.6$, $p_2 < .0001$), and a significant Wordtype X Matching interaction ($F_1(2,60) = 13.4$, $p_1 < .0001$; $F_2(2,54) = 17.7$, $p_2 < .0001$). However, there was no main effect of the group ($F_1 (1,30) = 0.25$; $p_1 = .62$; $F_2 (1,54) = .003$; $p_2 = .95$) and this factor did not interact with either of the other two (Wordtype: $F_1(1,30) = 0.11$, $p_1 = .73$; $F_2(1,54) =.47$, $p_2 = .49$; Match: $F_1(2,60) = 0.08$, $p_1 = .76$; $F_2(2,54) = .42$, $p_2 = .73$). It thus appears that the AS group was indistinguishable from the controls. *Post-hoc* Bonferroni tests reveal that the main effect of

Matching is due to the Ambiguous-Mismatch condition which leads to fewer correct responses (Mean = 59.5 %) than the Match condition (Mean = 96.1 %) and the Mismatch (Mean = 98.8 %) condition (both *p*s < .0001). Furthermore, the Wordtype X Matching interaction is also due to the Ambiguous-Mismatch condition which differs from all the others, both for compounds and for split-compounds (in all comparisons including the Ambiguous-Mismatch condition: all *p*s < .0001, no other comparison reaches significance). Results reveal that the Ambiguous-Mismatch condition is harder than all the others, especially when Split-Compounds are involved (Mean = 44.5 % for Split-Compounds vs. 72.7 % for Compounds). This indicates that participants are more likely to take a split-compound (e.g., the words "Dragon" and "Fly") as compatible with the picture corresponding to its compound counterpart (e.g., the picture of a dragonfly) than they are to accept the reverse association (e.g., the word "Dragonfly" associated to the picture of a dragon and of a fly).

A repeated measures ANOVA including the same factors as those used in the analysis of accuracy rates with participants' reaction times as the dependent measure revealed similar effects (see Figure 3, bottom). There was a main effect of Matching ($F_1(2,18) = 5.5$, $p_1 < .05$; $F_2(2,51) = 4.9$, $p_2 < .05$), no main effect of Wordtype ($F_1(1,19) = 3.0$, $p_1 = .09$; $F_2(1,51) = 1.9$, $p_2 = .17$) and group ($F_1(1,19) = 0.04$, $p_1 = .85$; $F_2(1,51) = .008$, $p_2 = .93$) and no interaction between the various factors (all F_1s < 1.7, all p_1s > .21; all F_2s < 2.12, all p_2s > .13). Again, our results reveal that adolescents with AS appear to be as able as TD adolescents to take prosody into account in chunking two- or three-word sequences.

Experiment 3 – Question contour

In this Experiment, we assess the participant's ability to distinguish questions from declaratives on the basis of prosodic and syntactic cues. All participants took part in this experiment. In the "Syntax" condition, both intonation and word order indicate that the utterance is a question (e.g., "Is this a dog?"), in the "Prosody" condition, the word order is identical to that of a declarative (e.g., "This is a dog?") and the only clue that the utterance as a question comes from intonation; in the "Declarative" condition, both intonation and word order indicate that the utterance is a declarative (e.g., "This is a dog."). The participant then has to decide whether the speaker sounded "sure" or "unsure" of what he said, and choose the corresponding response key (counterbalanced). The experiment was preceded by a two-trial training phase. Three lists each containing 15 items, were made up, with the constraint that each item appeared in only one of its possible forms ("Syntax", "Prosody" or "Declarative").

Results and discussion

Means and standard deviations of correct responses for the three sentence types (Syntax, Prosody and Declarative) are shown in Figure 4 (top).

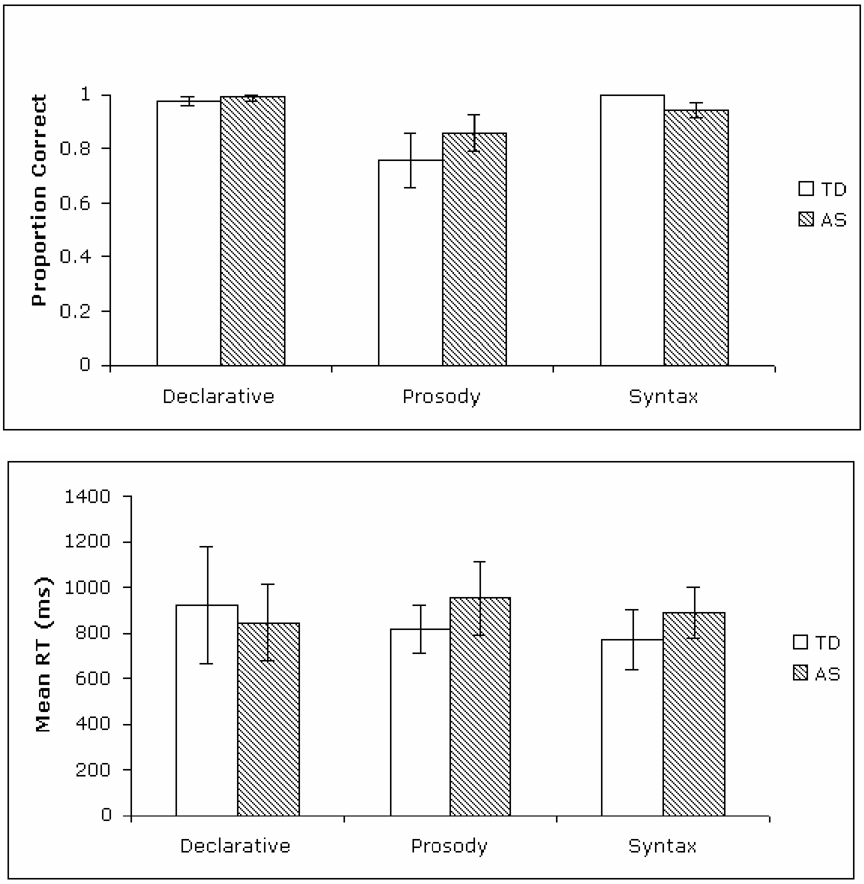


Figure 4. Proportion of correct answers (top) and reaction times (bottom) as a function of Sentence-type (Declarative, Prosody, Syntax) and group (TD, AS).

All the conditions are associated to very good performances ranging from 75.6 to 100 %. Unsurprisingly, the Declarative condition (Global Mean = 98.2 %; TD-Mean = 97.6 %; AS-Mean = 98.8 %) and the Syntax condition (Global Mean = 97.1 %; TD-Mean = 100 %; AS-Mean = 94.1 %) trigger ceiling performances. In the Prosody condition, the rates of correct responses are very high although not at ceiling (Global Mean = 80.6 %; TD-Mean = 75.3 %, $t(16) = -2.8$; AS-Mean = 85.9 %, $t(16) = -2.1$; all means are different from ceiling scores: all *p*s < .05). We conducted a mixed ANOVA using repeated measures with two factors: the within-subject factor “Sentence-type” (3: Declarative, Prosody, Syntax) and the between subject factor “Group” (2: TD, AS). Here, the dependent variable was the rate of correct responses. The ANOVA reveals a main effect of Sentence-type ($F_1(2,64) = 7.6$, $p_1 < .001$; $F_2(2,42) = 35.4$, $p_2 < .0001$), but no significant main effect of the group ($F_1(1,32) = 0.22$, $p_1 = .63$; $F_2(1,42) = 0.87$, $p_2 = .36$). The Sentence-type X Group interaction was significant in the item analysis only ($F_1(2,64) = 1.33$, $p_1 = .27$; $F_2(2,42) = 3.99$, $p_2 < .05$). *Post hoc* Bonferroni tests indicate that the Prosody condition differs from the other two (Declaratives: $p < .005$; Syntax: $p < .005$). This is evidence that the Prosody condition is harder, possibly because the information provided by word order and intonation conflict in this condition whereas they match in the other two conditions. Again, our results indicate that performance in the AS group does not differ from that of the controls.

A repeated measures ANOVA including the same factors as the ones used in the analysis of accuracy rates with participants' reaction times (see Figure 4, bottom) as the dependent measure revealed no main effect and no interaction between the various factors (all F_1s < 1.4, all p_1s > .80; all F_2s < .52, all p_2s > .60). Overall, our results demonstrate that AS participants are able to use intonation to decide whether the speaker sounds sure or unsure of what he says, and hence that they easily detect the question contour.

General discussion

The literature on the perception of grammatical prosody in HFA and AS has not yet provided a fully conclusive picture. Conflicting evidence has been put forward, some suggesting that grammatical abilities are intact, and some suggesting possible impairments. In this paper, we have argued that the inconsistency of the results obtained in the literature might have arisen for methodological reasons (for a similar argument, see McCann and Peppé's (2003) recent review mentioning recurring methodological flaws in the literature: small samples, absence of control population, poorly defined prosodic categories, small numbers of trials, etc.). In particular, we have stressed the importance of checking that there is no disorder at the perceptual level (e.g., inability to perceive differences in pitch, intensity or duration) before assessing the perception of prosody and we have highlighted the need to measure not only accuracy rates but also reaction times. Indeed, the systematic use of reaction time measures provide an extra tool to compare AS and TD performances, and allows one to explore potential differences in processing strategies. Particular care was devoted to addressing methodological worries mentioned in recent work focusing on prosody in autism. With respect to the construction of the tasks, control conditions were systematically included, the number of trials was either equal to or above the standards mentioned in Peppé et al.'s latest work (2007), and the stimuli were recorded in optimal acoustic conditions. As far as participants are concerned, our group included subjects with a narrow diagnosis (AS only) and from a narrow age range. Finally, performance in our tasks indicates that the control participants had no problems with the design and no ceiling effect was observed in the target conditions.

To summarise, our results indicate that adolescents with AS are able to decide on the most appropriate stress pattern for dissyllabic words (Experiment 1), to correctly chunk compounds on the basis of rhythmic cues (Experiment 2), and to take into account the intonational contour of a whole sentence in determining whether it is a statement or a question (Experiment 3). Strikingly, response time data concurred with the pattern observed for accuracy rates as AS participants responded at speeds similar to those of carefully matched control participants. Finally, identical patterns of accuracy and reaction times were observed when item analyses were performed, which gave us further confidence in the data. The data we obtained thus seem to be consistent across our various tasks (Experiments 1-3), measures (accuracy, reaction times) and analyses (by subjects or by items), and are consistent with previous findings indicating that grammar is generally spared in ASDs. Indeed, it is important to stress that explanatory theories of autism predict that teenagers with HFA and AS will have no difficulty dealing with the grammatical function of prosody. Overall, a reliable picture of AS children's ability to deal with grammatical prosody seems to emerge.

Nonetheless, it is important to stress that null effects are always difficult to interpret and should not be overstated. Furthermore, there are limitations to the present study which need to be acknowledged here and addressed in future work. For instance, the sample size, though larger than in many studies on language skills in ASDs, may not have been large enough to identify small sized effects. Similarly, the number of trials could be increased in future work (although our paradigms involved comparable numbers of trials to those used in previous work, such as Paul et al.'s (2005) - 24 trials per task - or Peppé et al.'s (2007) - 16 trials per task). A further issue is that this study focused on AS, and it is conceivable that our results would not generalise to the rest of the autism spectrum. Adapting our paradigms to lower functioning children might thus be an interesting route for future investigations, and differences between HFA and AS might also be worth exploring. Replication with more complex tasks would also be useful insofar as our measures of prosody may have been too simplistic and not sensitive enough to pick up prosodic deficits in older (and more able) individuals. For instance, the interplay between grammatical category and word-stress could be extended beyond homographs. Furthermore, the material employed here is clearly quite different from natural situations where prosody has to be processed in real time. Finally, although reaction times provide a more fine-grained tool than accuracy rates to detect potential processing differences, online measures (e.g. of brain activity) would be useful in order to completely rule out the possibility that AS participants performed like the controls while relying on compensatory strategies.

More generally, a possible line of research would be to further investigate *pragmatic* prosody in Asperger Syndrome. It has been noted that even amongst those individuals with an ASD who have fluent language (Hale & Tager-Flusberg, 2005), and even when the outcome is optimal (Kelley, Paul, Fein, & Naigles, 2006), a residual pragmatic deficit is always found. However, this claim is brought into question by recent findings reporting few group differences in pragmatic skills or ToM skills (e.g., Back, Ropar, & Mitchell, 2007; Ponnet, Roeyers, Buysse, De Clercq, & Van Der Heyden, 2004; e.g., Wang, Lee, Sigman, & Dapretto, 2006). Here again, measuring reaction times could be especially informative, since it may be that AS participants often use compensatory strategies which mask important group differences.

References

APA. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Asperger, H. (1944). Die „Autistischen Psychopathen" im Kindesalter. *European Archives of Psychiatry and Clinical Neuroscience, 117*(1), 76-136.

Attwood, T. (2006). *The Complete Guide to Asperger's Syndrome*: Jessica Kingsley Publishers.

Back, E., Ropar, D., & Mitchell, P. (2007). Do the Eyes Have It? Inferring Mental States From Animated Faces in Autism. *Child Development, 78*(2), 397-411.

Bailey, A., Phillips, W., & Rutter, M. (1996). Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry, 37*(1), 89-126.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced Pitch Sensitivity in Individuals with Autism: A Signal Detection Analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.

Ceponiene, R., Lepisto, T., Shestakova, A., Vanhala, R., Alku, P., Naatanen, R., et al. (2003). Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proceedings of the National Academy of Sciences, 100*(9), 5567-5572.

Cruttenden, A. (1997). *Intonation* Cambridge: Cambridge University Press.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141-201.

Dunn, L. M., Dunn, L. M., & Whetton, K. (1997). *The British Picture Vocabulary Scale* (2nd ed.). Windsor, UK: National Foundation for Education Research–Nelson.

Erwin, R., Van Lancker, D., Guthrie, D., Schwafel, J., Tanguay, P., & Buchwald, J. S. (1991). P3 responses to prosodic stimuli in adult autistic subjects. *Electroencephalogr Clin Neurophysiol, 80*(6), 561-571.

Findlay, J. M. (1978). Estimates on probability functions: A more virulent PEST. *Perception and Psychophysics, 23*(2), 181–185.

Fine, J., Bartolucci, G., Ginsberg, G., & Szatmari, P. (1991). The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 32*(5), 771-782.

Frith, U. (1991). *Autism and Asperger Syndrome*. Cambridge: Cambridge University Press.

Frith, U., & Happé, F. (1994). Language and Communication in Autistic Disorders. *Philosophical Transactions: Biological Sciences, 346*(1315), 97-104.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., et al. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience, 7*(8), 801-802.

Ghaziuddin, M., Thomas, P., Napier, E., Kearney, G., Tsai, L., Welch, K., et al. (2000). Brief Report: Brief Syntactic Analysis in Asperger Syndrome: A Preliminary Study. *Journal of Autism and Developmental Disorders, 30*(1), 67-70.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice'Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders, 37*(6), 1096-1106.

Gomot, M., Bernard, F. A., Davis, M. H., Belmonte, M. K., Ashwin, C., Bullmore, E. T., et al. (2006). Change detection in children with autism: An auditory event-related fMRI study. *Neuroimage, 29*(2), 475-484.

Grandin, T., & Scariano, M. (1986). *Emergence, Labeled Autistic*. New York: Warner Books.

Gravel, J. S., Dunn, M., Lee, W. W., & Ellis, M. A. (2003). Peripheral audition of children on the autistic spectrum. *Ear Hear, 2006*, 299-312.

Hale, C. M., & Tager-Flusberg, H. (2005). Brief Report: The Relationship between Discourse Deficits and Autism Symptomatology. *Journal of Autism and Developmental Disorders, 35*(4), 519-524.

Happé, F. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition, 48*(2), 101-119.

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3*(6), 216-222.

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry, 44*(4), 543-551.

Heaton, P. (2005). Interval and Contour Processing in Autism. *Journal of Autism and Developmental Disorders, 35*(6), 787-793.

Hermelin, B. (2001). *Bright splinters of the mind: A personal story of research with autistic savants*. London: Jessica Kingsley.

House, J. (2006). Constructing a context with intonation. *Journal of Pragmatics, 38*(10), 1542-1558.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA. 4th edition.: Wadsworth.

Huckvale, M. (2003). Prorec (version 1.0). *University College London*, Downloaded from http://www.phon.ucl.ac.uk/resource/prorec/#download.

Huckvale, M. (2004). Speech Filing System suite (version 4.6). *University College London*, Downloaded from http://www.phon.ucl.ac.uk/resource/sfs/.

Kanner, L. (1943). Autistic Disturbances of Affective Contact. *Nervous Child, 2*, 217-250.

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual Language Deficits in Optimal Outcome Children with a History of Autism. *Journal of Autism and Developmental Disorders, 36*(6), 807-828.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 389 - 403.

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

Landa, R. (2000). Social language use in Asperger syndrome and high-functioning autism. *Asperger syndrome*, 125-155.

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research, 1066*(1-2), 147-157.

Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., & Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. *Clinical Neurophysiology, 117*(10), 2161-2171.

Lyons, V., & Fitzgerald, M. (2004). Humor in Autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders, 34*(5), 521-531.

Martin, I., & McDonald, S. (2004). An Exploration of Causes of Non-Literal Language Problems in Individuals with Asperger Syndrome. *Journal of Autism and Developmental Disorders, 34*(3), 311-328.

McCaleb, P., & Prizant, B. M. (1985). Encoding of New versus Old Information by Autistic Children (Vol. 50, pp. 230-240): ASHA.

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

O'Riordan, M., & Passetti, F. (2006). Discrimination in Autism Within Different Sensory Modalities. *Journal of Autism and Developmental Disorders, 36*(5), 665-675.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205-220.

Peppe, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language and Hearing Research, 50*(4), 1015.

Ponnet, K. S., Roeyers, H., Buysse, A., De Clercq, A., & Van Der Heyden, E. (2004). Advanced Mind-Reading in Adults with Asperger Syndrome. *Autism, 8*(3), 249.

Pry, R., Petersen, A., & Baghdadli, A. (2005). The relationship between expressive language level and psychological development in children with autism 5 years of age. *Autism, 9*(2), 179.

Rapin, I. (1991). Autistic children: diagnosis and clinical features. *Pediatrics, 87*(5), 751-760.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the Mind in the Voice: A Study with Normal Adults and Adults with Asperger Syndrome and High Functioning Autism. *Journal of Autism and Developmental Disorders, 32*(3), 189-194.

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), 1097-1115.

Siegal, M., & Blades, M. (2003). Language and auditory processing in autism. *Trends in Cognitive Sciences, 7*(9), 378-380.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.

Sutcliffe, P., & Bishop, D. (2005). Psychophysical design influences frequency discrimination performance in young children. *Journal of Experimental Child Psychology, 91*(3), 249-270.

Tharpe, A. M., Bess, F. H., Sladen, D. P., Schissel, H., Couch, S., & Schery, T. (2006). Auditory Characteristics of Children with Autism. *EAR AND HEARING, 27*(4), 430.

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.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain, 129*(4), 932.

Werth, A., Perkins, M., & Boucher, J. M. (2001). 'Here's the weavery looming up': verbal humour in a woman with high-functioning autism. *Autism, 5*(2), 111.

Williams, D. (1994). *Somebody Somewhere*. London: Jessica Kingsley.

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics, 38*(10), 1559-1579.

Footnotes

[1] Contrastive stress and word stress differ in English. For instance, consider the following utterance: “The uniVERsity MUST be rebuilt” (with capital letters indicating stressed syllables). Here, the word stress on “university” is obligatory whereas the contrastive stress on “must” is optional. In other words, word stress is obligatory and is encoded in the linguistic representation of each word whereas contrastive stress is optional and allows the speaker to express a certain pragmatic attitude towards the proposition she is expressing (Rossi, 1999).

[2] Note that incorrectly stressing the control items resulted in neologisms (e.g., “I like puZZLEs”) whereas incorrectly stressing the target items resulted in a change in the grammatical category of the item (e.g., “I gave him a preSENT”).

[3] Note that the same distribution is found in both groups for split compounds (χ^2 (2, 31) = 0.02, *p* = n.s.): 7 AS and 8 TD participants score below 25%, 6 AS and 7 TD participants score above 75%, and 2 score at chance in each group; and for Compounds: (χ^2 (2, 31) = 1.04, *p* = n.s.): 4 AS and 2 TD participants score below 25%, 9 AS and 11 TD participants score above 75%, and 2 AS and 3 TD participants score at chance.