J Autism Dev Disord (2010) 40:1104–1117
DOI 10.1007/s10803-010-0960-8

ORIGINAL PAPER

Scalar Inferences in Autism Spectrum Disorders

Coralie Chevallier · Deirdre Wilson · Francesca Happé · Ira Noveck

Published online: 9 February 2010

Abstract On being told "John or Mary will come", one might infer that *not both* of them will come. Yet the semantics of "or" is compatible with a situation where both John and Mary come. Inferences of this type, which enrich the semantics of "or" from an 'inclusive' to an 'exclusive' interpretation, have been extensively studied in linguistic pragmatics. However, the phenomenon has not been much explored in Autism Spectrum Disorders (ASDs), where pragmatic deficits are commonly reported. Here, we present an experiment investigating these inferences. We predicted that, as a result of the reported pragmatic deficits, participants with ASD would produce fewer inferential enrichments of "or" than matched controls. However, contrary to expectations, but in line with recent findings by Pijnacker et al. (Journal of Autism and Developmental Disorders, 39, 607–618, 2009), performances did not differ across groups. This unexpected finding is discussed in light of the literature on pragmatic abilities in autism.

Keywords Autism Spectrum Disorders · Scalar inference · Language · Prosody · Pragmatics

These studies were conducted through a collaboration between the Laboratoire Langage, Cerveau et Cognition, the Institute of Psychiatry and UCL's Department of Phonetics and Linguistics.

C. Chevallier · I. Noveck
Laboratoire Langage, Cerveau et Cognition, CNRS—Université de Lyon, Lyon, France

C. Chevallier · F. Happé
Institute of Psychiatry, King's College London, London, UK

D. Wilson
Department of Linguistics, University College London, London, UK

D. Wilson
CSMN, University of Oslo, Oslo, Norway

C. Chevallier (✉)
L2C2, Institut des Sciences Cognitives, 67 Boulevard Pinel, 69675 Bron cedex, France
e-mail: cchevallier@isc.cnrs.fr

Introduction

One of the challenges linked to everyday utterance interpretation is that there is always more to be understood than what is linguistically encoded (Burton-Roberts 2007; Horn and Ward 2004; Sperber and Wilson 1986/1995). To illustrate, consider the following three utterances (taken from Carston 2002; Wilson in press):

(1) A: Did you enjoy the evening at Bob and Mary's?
B: I'm not much of a party person.
(2) Susan is a wild rose.
(3) *(said in a downpour)* It's lovely weather.

In each case, the hearer has to go beyond the linguistic meaning in order to recognise *what the speaker intended to convey*. For instance, in (1), the answer to A's question is not explicitly stated, but the speaker clearly implied that B did not enjoy the evening. In the case of figurative utterances, the linguistically encoded meaning is merely a vehicle for communicating something else. Thus, in (2), the speaker might be understood as conveying that Susan resembles a wild rose in some respects, and in (3), the speaker would generally be understood as meaning that the weather is not so nice. In other words, there is a considerable gap between *sentence* meaning (i.e., the semantics of the sentence uttered) and *speaker's* meaning (i.e., what the speaker intended to convey). Bridging this gap requires a great deal of inference-making which involves a subtle interaction between linguistic properties and contextual factors. A central goal of pragmatics is to describe the principles governing these inferences, and thus explain how the gap between sentence meaning and speaker's meaning is bridged.

In typically developing children, there is some early evidence of the ability to recognise a communicator's meaning. For instance, 17 month-olds are able to track the communicator's epistemic state in order to infer what she intends to refer to when pointing (Southgate et al. accepted) and infants use pointing themselves to influence an adult's mental states even before this age—at around 12 months (e.g., Liszkowski et al. 2007; Tomasello et al. 2007). In contrast, children with an Autism Spectrum Disorder (ASD) generally fail to engage in spontaneous non-verbal communicative behaviours such as showing, waving, pointing (Mitchell et al. 2006) and show diminished abilities to take account of the speaker's communicative intentions in vocabulary acquisition and other pragmatic tasks (Bloom 2000; Surian and Siegal 2008). These difficulties have been related to their well-known impairments in attributing a variety of mental states (Baron-Cohen 2000), of which communicative intentions are a special kind (Sperber and Wilson 1986/1995, 2002). There is experimental evidence that metaphor, irony, sarcasm and jokes are often poorly understood by individuals on the autism spectrum (Dennis et al. 2001; Happé 1993, 1994; Martin and McDonald 2004), that they are less sensitive to violations of Gricean maxims (be truthful, informative, relevant, and clear) (Surian et al. 1996) and that they have difficulties dealing with bridging inferences and global inferences required for text interpretation (Jolliffe and Baron-Cohen 1999, 2000). This pragmatic deficit is usually described as universal across the autism spectrum, and even the subtlest forms of the disorder, High Functioning Autism (HFA) and Asperger Syndrome (AS), are associated with pragmatic difficulties (Frith 1998; Frith and Happé 1994; Kelley et al. 2006).

As summarized, ASDs are characterised by a generalised deficit in the ability to go beyond the linguistically encoded meaning to infer the communicator's intentions. This suggests that any kind of pragmatic inference should be problematic for this population. In this respect, the so-called 'scalar' inferences associated with logical terms such as "or" and "some"—often regarded as paradigm cases for experimental investigation in the linguistic-pragmatic literature (for a review, see Noveck and Reboul 2008; Noveck and Sperber 2004, 2007)—provide an interesting test case. To illustrate, consider (4) and (5):

(4) John or Mary will come.
(5) Some students passed.

In many circumstances, someone hearing (4) or (5) will understand that John or Mary, *but not both,* will come, or that some students, *but not all*, passed. This is despite the fact that from a strictly semantic point of view, "A or B" is logically equivalent to "A or B or both", and "Some Xs" is logically equivalent to "Some, and possibly all, Xs." "Or" and "some" are examples of a class of 'scalar' terms which have received considerable attention in both the linguistic and the experimental literature. According to Horn (1972), a scale is a set of alternate linguistic expressions ranked by order of informativeness (e.g. *<or, and>*, *<some, all>*, *<possible, certain>*, *<might, must>*, *<often, always>*, etc.). Scalar inferences occur when the speaker's choice of a less informative term is taken to imply that the more informative term is not applicable. In these cases, the use of "often" can be taken to imply *not always*, the use of "some" can be taken to imply *not all*, the use of "possibly" can be taken to imply *not certainly*, and so on. Thus, the hearer has to go beyond the encoded linguistic meaning in order to infer the speaker's meaning.

There has been a considerable debate about the exact principles or mechanisms underlying the production of these inferences (for survey and discussion, see Horn 2004). However, it is widely accepted in both theoretical work (Carston 1998; Chierchia 2004; Geurts 2009; Horn 2004; Levinson 2000; Noveck and Sperber 2007; Recanati 2003) and experimental studies (e.g., Bott and Noveck 2004; Breheny et al. 2006; Chevallier et al. 2008; De Neys and Schaeken 2007) that scalar inferences exhibit the main hallmarks of pragmatic inference (they are context dependent, cancellable without contradiction and calculable by reference to the standards that rational speakers are expected to meet), and should therefore be seen as genuinely pragmatic.[1]

Previous experimental studies have shown that typically developing children between the ages of about 4–10 are less likely than adults to make pragmatic inferences in interpreting scalar terms (Noveck 2001; Papafragou and Musolino 2003; Pouscoulous et al. 2007),[2] and that inference making in adults is generally associated with deeper, or more effortful processing (Bott and Noveck 2004; Breheny et al. 2006; Chevallier et al. 2008; Noveck and Posada 2003). In a typical experiment, participants are presented with a range of (spoken or written) utterances

[1] Alternative, more semantic, treatments are suggested in recent work by Chiercha and his colleagues (Chierchia, 2004, to appear).

[2] These studies have also shown that children are more likely to draw implicatures when task demands are reduced, which suggests that limitations in cognitive resources—rather than diminished pragmatic competence—may play an important role in children's pragmatic difficulties (Siegal and Surian 2004).

and asked to agree or disagree with the speaker, or to judge the utterance as 'right' or 'wrong', or 'true' or 'false'. In crucial cases, a single sentence can be judged true or false depending on whether or not a pragmatic inference is produced. It is then possible to compare conditions where participants produce scalar inferences with those where they do not. In Chevallier et al. (2008), for instance, after reading a word such as TABLE, adult participants were asked to judge whether the description "There is a T or a B" is true or false. Participants who go beyond the encoded linguistic meaning and produce a so-called 'exclusive' interpretation (in which the speaker is understood as meaning *not both T and B*) will judge the description false. By contrast, those who stick with the encoded meaning and produce a so-called 'inclusive' interpretation (in which the speaker is understood as meaning *T or B or both*) will judge the description true. The results showed that increasing the depth to which utterances were processed (by compelling participants to wait before responding, or by adding contrastive stress to "or") yielded more exclusive interpretations, suggesting that more pragmatic inferences had been drawn. Other authors have shown that participants take longer to read a disjunctive expression (e.g., "the class notes or the summary") when it occurs in a discourse context favouring an exclusive interpretation, as in (6), as compared to one favouring an inclusive interpretation, as in (7) (Breheny et al. 2006).

(6) John was taking a university course and working at the same time. For the exams he had to study from short and comprehensive sources. Depending on the course, he decided to read / the class notes or the summary.

(7) John heard that the textbook for Geophysics was very advanced. Nobody understood it properly. He heard that if he wanted to pass the course, he should read / the class notes or the summary.

Taken together, these findings suggest that enriched interpretations of scalar terms do not result from a default (or automatic) process, but from genuinely context-dependent pragmatic inferences.

In light of this general consensus, and of the well-known pragmatic deficits found in ASDs, it seems reasonable to predict that in interpreting ordinary connectives (e.g. "or") and quantifiers (e.g. "some"), participants on the autism spectrum will draw fewer scalar inferences than typically developing participants. The aim of this paper is to test this prediction. This also provides a follow-up to a study (undertaken independently of ours) which was published as we were writing this paper (Pijnacker et al. Geurts 2009). In that study, adult participants were presented with a range of written sentences containing the scalar terms "some" (e.g., "Some birds are sparrows") or "or" (e.g., "Snakes have paws or wings") and were asked to judge whether these were true or false (*True* in the first case; *False* in the second). In the *underinformative* condition, the very same sentence could be judged as true or false depending on whether or not the participant produced the scalar inference. For instance, "Zebras have black or white stripes" can be justifiably evaluated as false or true depending on whether or not the inference is made. If the participant interprets "black or white" exclusively (as *black stripes or white stripes but not both*), she will judge the sentence as *False*; but if she interprets the disjunction inclusively (as *black stripes or white stripes or both*), she will judge the sentence as *True*. In the first case, the participant has produced a pragmatic inference (*not both*), while in the latter she has not enriched the initial semantics of "or" (which is compatible with *both*). Contrary to the authors' initial hypothesis, participants with ASD were as likely as the controls to draw pragmatic inferences in both scalar tasks ("some" and "or").

Pijnacker et al. were then able to separate participants with HFA and participants with AS according to standard criteria (i.e., absence or presence of early language delay, see DSM-IV, APA 1994; and ICD-10, WHO 1992). The diagnostic distinction between HFA and AS is very controversial, with some authors suggesting that *current* verbal ability is more relevant and predictive of social cognitive abilities than *early* language skills (Williams et al. 2008; Witwer and Lecavalier 2008). In this respect, it is interesting that Pijnacker et al. found not only VIQ differences between their HFA and AS groups, but also differences in the way these two groups dealt with scalar terms. More specifically, they found that participants in the HFA group produced fewer pragmatic inferences than those in the AS group (who had control-like performances). In fact, in the case of "some", participants in the AS group produced even more pragmatic enrichments than the controls. These unexpected results cast doubt on the assumption that pragmatic processes are *universally* impaired across the autism spectrum, and are therefore worth exploring further.

The present study provides a follow-up to Pijnacker et al.'s (2009), from which it differs in two significant ways: (1) it involves younger participants, and (2) it is based on a spoken language paradigm. The sampling difference is critical, since it is often argued that good performances in adulthood may be the result of well-developed compensatory strategies, rather than intact pragmatic competence. One could then hypothesise that testing at earlier stages in development will lead to different patterns of performance. It is also important to check whether Pijnacker et al.'s findings can be extended to the processing of spoken language. Understanding spoken language presents particular challenges, such as the ability to produce inferences online, as the utterance unfolds, or to take rapidly into account a host of cues, including elements of the context, information about the speaker's mental states, prosodic cues, etc.

In particular, the prosodic form of an utterance makes a significant and systematic contribution to conveying the speaker's meaning. To illustrate, consider the utterances in (8):

(8)

a Paul cooked a ROAST.
b Paul COOKED a roast.

Although these two utterances are identical at the segmental level, the slight difference in intonation patterns is likely to yield different interpretations. Emphasis on "roast", in (8a), might be taken to imply that Paul cooked a roast as opposed to some other dish; whereas emphasis on "cooked", in (8b), might be taken to imply that he cooked it *himself* (as opposed to buying it, for example). Recent experimental work confirms that contrastive stress plays an early role in the on-line determination of reference (Weber et al. 2006), and in facilitating access to presupposed contrastive alternatives to the accented word (e.g. 'buying' rather than 'cooking', in example 8b) (Braun and Tagliapietra in press). Similarly, contrastive stress plays an important role in the interpretation of scalar terms. To illustrate, consider (4) again: if a speaker wants to emphasize her intention to communicate that *not both John and Mary will come*, she can produce the very same utterance, but with added stress on the disjunction, as in (9):

(9) John OR Mary will come.

Thus, contrastive stress provides a natural means of orienting the hearer towards an inferentially enriched interpretation (Wilson and Wharton 2006).

As noted above, the importance of prosody in the interpretation of disjunctive utterances has recently been tested experimentally in ordinary adults (Chevallier et al. 2008). The rationale for the experiment was that pitch accents increase the salience of the material they are associated with, and that speakers can exploit this fact to guide the interpretation process in one direction or another (Pierrehumbert and Hirschberg 1990; Sperber and Wilson 1986/1995). Given that this type of accent is optional and requires the hearer to invest more processing effort, it is likely to be interpreted as an attempt by the speaker to increase the salience of this particular part of the utterance and to encourage the hearer to construct an interpretation on which this extra salience makes sense (House 2006). This suggests that placing an accent on "or" should encourage the hearer to look for a different interpretation from the one she would have arrived at had the salience of the word not been increased. Chevallier et al.'s data are in line with this prediction. Indeed, most participants (77%) arrived at an *inclusive* reading when "or" was not accented, whereas most (73%) arrived at an *exclusive* reading when "or" was accented[3]: that is, the addition of an accent increased pragmatic inferences from 23 to 73%. This result confirms that prosodic cues play a major role in the interpretation process and often guide the hearer towards an interpretation that might otherwise not have come to mind.

However, individuals on the autism spectrum struggle with prosody. Abnormal prosody has been reported since the initial descriptions of the condition (Asperger 1944; Kanner 1943) and is often mentioned by clinicians, parents and individuals on the spectrum themselves. Prosodic deficits are also an important cue for diagnosis (Lord et al. 2000); they often persist even when other areas of language improve (Shriberg et al. 2001), and they are found in almost half of people on the autism spectrum (Paul et al. 2005). In fact, Peppé and collaborators (2007) recently demonstrated that *all* the children with HFA included in their sample had difficulty with *at least one aspect* of prosody. Yet this area has remained relatively under researched (for a review, see McCann and Peppé 2003). So far, research indicates that differences can be found in both expressive and receptive prosody (Paul et al. 2005; Peppé et al. 2007). On the expressive side, some of the most commonly reported atypical features include monotonous, sing song, pedantic, robotic, or even foreign sounding intonation. The voice is often described as having an odd quality, often with increased fundamental frequency variation (Diehl et al. 2009). On the receptive side, recent research suggests that individuals with autism have difficulties reading emotions in the voice (Golan et al. 2007; Rutherford et al. 2002) or understanding irony and sarcasm (Wang et al. 2006, 2007). Peppé et al.'s work (2007) also confirms that children with HFA especially struggle with affective and pragmatic functions (see also Baltaxe and Guthrie 1987; Shriberg et al. 2001).

In the present paper, we follow up on Pijnacker et al.'s (2009) work and adapt Chevallier et al.'s (2008) paradigm to further investigate scalar inferences among teenagers with an ASD. Given the well-document deficit with the pragmatic function of prosody found in people with autism, we predict that participants in the ASD group will be less likely than matched controls to use contrastive stress as a cue to go beyond the linguistic meaning and arrive at a pragmatically enriched exclusive interpretation of "or." Our paradigm involves the presentation of two pictures, followed by a disjunctive or conjunctive sentence

[3] This rate of pragmatic enrichment (leading to an exclusive interpretation) is among the highest reported in laboratory tasks dealing with scalar terms.

(recorded and presented through headphones). For example, participants are shown the picture of a horse and a goat, followed by a sentence such as "There is a horse or a dog" (which is True), "There is a bucket and a goat" (False), etc. Six sentence types were used, based on two connectives—"and" and "or"—and three truth conditions (see Table 1).

The word "or" in sentences mentioning two items which were indeed shown in the pictures (i.e. the True-True, Or TT, condition) carried contrastive stress (e.g. the picture of a horse and a goat was followed by the sentence "There is a horse OR a goat"). In this Or TT condition, it is acceptable to answer either "true" or "false"; "true" responses indicate that the participant has accessed an interpretation corresponding to the semantics of the disjunction (compatible with situations where both disjuncts are true) and "false" responses indicate that a scalar inference has been derived (leading the participant to reject the description in situations where both disjuncts are true). As mentioned above, applying prosodic stress to "or" strongly encourages participants to produce pragmatic inferences. Thus, if participants in the ASD group do not draw scalar inferences as readily as matched controls in this task, it will provide further evidence that they have difficulties with pragmatic processing. On the other hand, if performance is equivalent in both groups, Pijnacker et al.'s (2009) unexpected finding will be supported and extended to the processing of spoken language (which would suggest that the pragmatic skills underlying the production of scalar inferences are spared in High Functioning Autism). The other conditions (TF/FT and FF) act as control conditions and should not give rise to any difference between the two groups.

Methods

Participants

Forty-four male adolescents (22 with ASD and 22 Typically Developing, henceforth TD) took part in the experiment. The participants with ASD were seen at two special education schools in England which require formal diagnosis of an Autism Spectrum Disorder according to standard clinical criteria (APA 1994; WHO 1992). The diagnostic information was gathered from school files of documented medical diagnoses made by a clinical psychologist and/or psychiatrist. Six participants were diagnosed as having HFA and 16 as having Asperger Syndrome. As we were unable to collect ADI-R (Autism Diagnostic Interview-Revised) data in order to confirm that the HFA/AS distinction was based on accurate information regarding early language and cognitive delay, we decided to group all participants together. This grouping is justifiable given the current lack of clear evidence that HFA and AS are distinct conditions and recent findings that the AS diagnosis is used extremely variably (Williams et al. 2008; Witwer and Lecavalier 2008). The controls were seen in two regular English schools. TD participants and participants with ASD all spoke English at home, and none had any significant hearing loss, visual impairment, or major physical disability. The control participants were closely matched on chronological age and verbal mental age using the British Picture Vocabulary Scales-II (Dunn et al. 1997) (see Table 2 for detailed information). We also ensured that participants in both groups reached similar levels in their ability to perceive pitch, duration and intensity, using simple auditory tasks designed and programmed by Dorothy Bishop (for previous studies using the Dinos task, see, e.g. Sutcliffe and Bishop 2005). These tasks are based on a "more virulent" PEST procedure (Findlay 1978) which makes it possible to determine the smallest difference in frequency, duration and intensity that the participant can detect.

Stimuli

All the stimuli are presented in Appendix 1. Sixty pictures were selected from the Snodgrass database (Snodgrass and Vanderwart 1980), randomly paired and assigned to one of six conditions (following a 2 (connective: And vs. Or), by 3 (truth-condition: TT, TF/FT, FF) design). The auditory

Table 1 An illustration of each of the six experimental conditions

Connective	Truth condition	Example of pictures	Corresponding sentence	Logically correct response
And	TT	Frog + Flower	There is a frog and a flower.	True
	TF/FT	Button + Brush	There is a banana and a brush.	False
	FF	Basket + Bear	There is a donkey and a cake.	False
Or	TT	Sun + Train	There is a sun OR a train.	True[a]
	TF/FT	Shirt + Rope	There is a shirt or a bear.	True
	FF	House + Hand	There is a bee or a chair.	False

[a] A true response is justified because the sentence contains a true disjunct and a false response is justified because "or" can be taken to pragmatically imply "not both"

Table 2 Participants' age and BPVS score and discrimination threshold values for intensity, duration and frequency

	TD participants $N = 22$		Participants with ASD $N = 22$		*t* test
	Mean (SD)	Range	Mean (SD)	Range	*t*(df); *p*
Age	13;10 (1;5)	10;10–16;03	13;04 (1;4)	11;01–15;11	$t = -1.25$; $p = 0.22$
BPVS score	107.1 (18.9)	76–146	111.0 (22.7)	72–145	$t = 0.62$; $p = 0.54$
Intensity (dB)	2.43 (1.62)	0.54–5.4	2.43 (1.35)	0.81–5.67	$t = -0.29$; $p = 0.77$
Duration (ms)	32 (24)	8–80	56 (48)	8–224	$t = -1.78$; $p = 0.09$
Frequency (Hz)	82 (62)	2–252	92(102)	2–342	$t = -0.64$; $p = 0.52$

Note that low thresholds are indicative of optimal performance

stimuli were recorded in an anechoic chamber 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 (Bruel and Kjaer 2231 Sound Level Meter fitted with a Type 4165 Microphone) remained constant. 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.

Design and Procedure

Written parental consent and oral child assent were obtained prior to the testing phase. Pupils were seen individually at school during two 35-min sessions (results for the other experiments are reported in Chevallier et al. 2009). The experiment lasted about 10 min and was presented using a laptop equipped with headphones. The instructions—presented on the screen—were read out to the participants:

> You are going to hear someone describing objects that he is looking at. First, you will see the pictures of the objects which he was looking at. Then, you will have to listen carefully to the description. When the description is finished, your job is to say whether the description was right or wrong.

This was followed by a three-trial training phase. Each trial started with a 2,000 ms screen showing two pictures. The test sentence was then uttered (in English) while the pictures remained on the screen. The participant was then required to decide whether the speaker's description was right or wrong, using the 'right' or 'wrong' response key ('E' and 'P' counterbalanced), and the next trial started 1,000 ms later. The experimental session was divided into two blocks of 15 trials, separated by a break screen displaying the message *"You're half way through!"*. The 30 trials were equally divided between the two connectives (And, Or) and the three conditions (FF, FT/TF, TT), yielding 5 stimuli per condition. Half the questions required a "Right" answer.

Results

Prior to data analysis, individual cutoff values were calculated for each participant as the mean ± 2 standard deviations over all items. Reaction times above or below this cutoff were considered outliers and were excluded from both the choice proportion analysis and the reaction time analysis (these amounted to 5.2% in the ASD group and 5.8% in the TD group). Response proportions were not normally distributed and were thus analysed using non-parametric statistics. Given the small sample sizes typically used in ASD research, it is important to establish whether negative findings reflect lack of power. We therefore report effect sizes as well as *p* values. Though effect sizes are often not calculated for Mann–Whitney tests, Green and Salkind (2008) suggest that differences in mean ranks between the two groups can be used as an effect size index. For reaction times, partial eta-squared are reported. Following Cohen (1988), effect sizes above 0.20 reflect a small effect, effect sizes above 0.50 reflect a medium effect, and effect sizes above 0.80 reflect a large effect.

Analysis of Response Proportions

Percentages of correct answers for all conditions and for both groups are shown in Table 3. By convention, inclusive interpretations of "or" are coded as "correct", and exclusive answers—resulting from a scalar inference—are coded as "incorrect". For example, in response to "There is a sun OR a train" associated to the presentation of a picture of a sun and a picture of a train, answering "True" will be taken

Table 3 Percentage of correct answers as a function of connective type (And, Or), condition (FF, FT/TF, TT) and group (TD, ASD); results of the Mann–Whitney *U* tests comparing the groups

Connective	Truth condition	TD ($N = 22$) Mean (SD)	ASD ($N = 22$) Mean (SD)	Mann–Whitney *U* test z; *p*	Difference in mean ranks (Effect size index)
And	FF	100 (0)	99 (4)	$z = 0.00$; $p = 1.00$	0
	TF-FT	99 (4)	99 (4)	$z = -0.77$; $p = 0.44$	−0.07
	TT	100 (0)	97 (7)	$z = -0.25$; $p = 0.80$	−0.02
Or	FF	100 (0)	99 (5)	$z = 0.83$; $p = 0.40$	0.07
	TF-FT	42 (45)	57 (47)	$z = 0.35$; $p = 0.72$	0.03
	TT	43 (44)	48 (48)	$z = 0.01$; $p = 0.99$	0.001

as an inclusive interpretation and coded as correct, while answering "False" will be taken as an exclusive interpretation and coded as incorrect.

The three And control conditions and the Or FF condition are associated with ceiling scores which do not differ between the TD group and the ASD group (see *p* values in Table 3). The Or TF/FT condition leads to lower scores than the FF condition for both groups (42% in the TD group and 57% in the ASD group). These low scores do not reflect chance performance but rather a bimodal distribution, with some children consistently accepting situations where only one of the disjuncts is true and others consistently (and wrongly) rejecting these trials. Moreover, the distribution does not differ across the two groups, $\chi^2(2) = 3.94$, $p = 0.14$ (9 ASD and 10 TD participants score below 20%, 13 ASD and 8 TD participants score above 75% and 4 TD scored at chance levels, i.e., between 40% and 60%). Some possible explanations for these unexpected results are considered in the discussion section.

In the critical Or TT condition, 43% of responses among TD participants and 48% among participants with ASD interpret the disjunction inclusively. Contrary to our initial hypothesis, the rate of pragmatic inference did not differ across the two groups (see Table 3). It must also be noted that the scores obtained in the Or TT condition do not reflect chance performances but, again, a bimodal distribution, with some participants consistently deriving the pragmatic inference (25% or below of "Right" responses in the Or TT condition: 12 TD and 11 ASD) and others consistently not doing so (75% or above "Right" responses in the Or TT condition: 7 TD and 10 ASD). There are only a few inconsistent participants (40% to 60% "Right" responses in the Or TT condition: 3 TD and 1 ASD). Again, the distribution does not differ between groups, $\chi^2(2) = 1.57$, $p = 0.46$.

We also divided participants into two groups depending on their response patterns: participants who produced a pragmatic inference (i.e., "Wrong" answer in the Or TT condition) more than 50% of the time were classified as "Pragmatic responders" and the others were classified as "Literal responders" (i.e., mostly "Right" answer in the Or TT condition). Eleven ASD and 13 TD participants were classified as Pragmatic responders; 11 ASD and 9 TD participants were classified as Literal responders. Again, the distribution of pragmatic and literal responders does not differ between groups, $\chi^2(1) = 0.37$, $p = 0.55$.

So far, the results suggest that there is no difference between ASD and TD participants. However, there is one concern with the present set of data, linked to the low proportion of correct responses in the TF/FT condition. This raises the question of whether participants who fail in this condition are resorting to an alternative strategy to answer: for instance, using a simple matching strategy without really considering the meaning of "or". In that case, they might (wrongly) reject utterances in the TF/FT condition because the items mentioned in the utterance do not match the pictures presented on the screen. In the TT condition, by contrast, where both items are presented, a matching strategy would lead participants to regard the description as true. This would of course undermine confidence in our results for the TT condition, since it would imply that some participants were classified as literal responders when they were in fact merely looking for a match between the words and the pictures, rather than correctly applying the semantics of inclusive "or".

There are several reasons for thinking that this is not what is happening. In the first place, rates of correct responses in the TF/FT and TT conditions do not correlate, $r = -0.18$; $p = 0.26$, even when the two groups are considered separately, TD group: $r = -0.03$; $p = 0.89$; ASD group: $r = -0.31$; $p = 0.16$. This implies that failing (or passing) in the TF/FT condition does not covary with the response pattern found in the TT condition. In the second place, we reanalysed our data including only the 21 participants who had scored above 75% in the TF/FT condition, and found the same pattern across all conditions (And FF: TD mean = 100%, ASD mean = 100%, $z = 0.00$; $p = 1.00$; And TF/FT: TD mean = 98%, ASD mean = 100%, $z = 0.00$; $p = 1.00$; And TT: TD mean = 100%, ASD mean = 98%, $z = -0.30$; $p = 0.76$; Or FF: TD mean = 100%, ASD mean = 98%, $z = -0.30$; $p = 0.76$; Or TF/FT: TD mean = 97%, ASD mean = 95%, $z = 0.00$; $p = 1.00$; Or TT: TD mean = 30%, ASD mean = 38%, $z = -0.23$; $p = 0.82$). Finally, it is worth stressing that in both our analyses, ASD and TD participants provided comparable rates of correct responses in all the conditions (including the TF/FT condition). Possible explanations for these low performances are addressed in the discussion.

Overall, we thus replicate Pijnacker et al.'s (2009) unexpected result and find similar rates of pragmatic inferences in ASD and TD participants. However, Pijnacker et al. did detect differences once HFA and AS were considered separately. As noted in the introduction, participants in the AS group were more prone to deriving scalar inferences than those in the HFA group. Furthermore, Pijnacker et al. found a significant VIQ difference between the two groups and a negative correlation between VIQ and number of logical interpretations in the HFA group.[4] In what follows, we focus on the role of VIQ in the interpretation of the Or TT condition. We divide the ASD and TD group into two: those participants with a verbal IQ above the average of the group and those with a verbal IQ below the average of the group. Twelve ASD participants and 11 TD were included in the higher IQ group; 10 ASD participants and 11 TD were included in the lower IQ group.

A Mann–Whitney U test revealed that the IQ factor had a significant effect in the ASD group, $z = -1.98$, $p < 0.05$, difference in mean ranks = −0.22, but not in the TD group, $z = 0.65$, $p = 0.66$, difference in mean ranks = 0.08 (see Fig. 1). No differences were observed in the other conditions, all $zs > |1.18|$, all ps = ns. A correlation analysis confirmed that VIQ scores correlated with performance in the Or TT condition in the ASD group only, ASD: $r = -0.44$, $p < 0.05$; TD: $r = 0.25$, $p = 0.27$.

Reaction Time Analysis

For all reaction time analyses, a log transformation was carried out beforehand to improve the conformity of the data to the standard assumptions of ANOVA (e.g., Howell 1997). Shapiro–Wilk tests were carried out and show that the data were normally distributed in every condition, all $Ws > 0.98$, all $ps > 0.58$. An ANOVA using repeated measures with the within-subject factors "connective" (And vs. Or), "truth-condition" (TT, TF/FT, FF) and the between-subject factor "Group" (TD, HFA/AS) reveals a main effect of connective

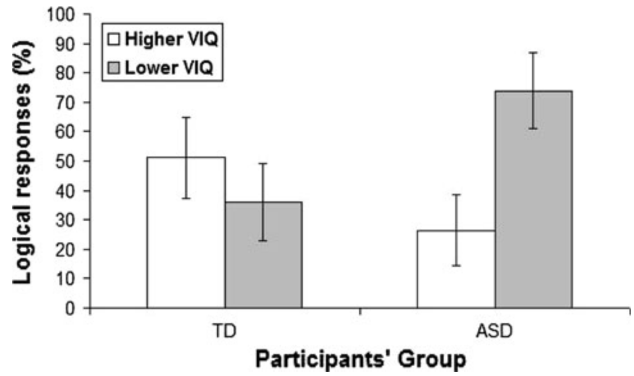

Fig. 1 Percentage of inclusive answers in the Or TT condition as a function of VIQ category (higher VIQ, lower VIQ) and group (TD, ASD)

on reaction times, $F(1,42) = 16.55$, $p < 0.001$, *Partial* $\eta^2 = 0.28$, with longer processing times for "or" (mean = 1,415 ms) than for "and" (mean = 1,102 ms) (see Fig. 2). This result is unsurprising given that the truth conditions of "and" are more straightforward than those of "or".[5] A main effect of Truth-Conditions is also revealed, $F(2,84) = 19.28$, $p < 0.0001$, *Partial* $\eta^2 = 0.32$, due to longer processing times associated with the TF/FT condition (mean = 1,397 ms) than with the other two (mean FF = 1,009 ms, Tukey HSD post-hoc test, $p < 0.001$; mean TT = 1,207 ms, $p < 0.005$). This last effect is in line with the pattern of accuracy rates reported above (i.e. difficulties associated with the TF/FT condition until late in development). There is no significant main effect of Group, $F(1,42) = 0.16$, $p = 0.69$, *Partial* $\eta^2 = 0.004$, no connective × group interaction, $F(1,42) = 0.002$, $p = 0.97$, *Partial* $\eta^2 = 0.00$, and no truth-condition × group interaction, $F(2,84) = 1.93$, $p = 0.15$, *Partial* $\eta^2 = 0.04$. Furthermore, Tukey HSD tests reveal that in all six conditions, reaction times do not differ across groups, And FF: $p = 0.88$, And TF/FT: $p = 0.88$, And TT: $p = 0.60$, Or FF: $p = 0.16$, Or TF/FT: $p = 0.60$, Or TT: $p = 0.24$. Finally, a connective × truth-condition interaction, $F(2,84) = 12.74$, $p < 0.0001$, *Partial* $\eta^2 = 0.23$, was found. Post hoc tests reveal slower reaction times in the Or TT condition (mean = 1,757 ms) compared to the And TT condition (mean = 793 ms), $p < 0.001$. In contrast, the two connectives do not differ in the other truth conditions, TF/FT: $p = 1.00$; FF: $p = 0.96$.

Since previous findings have demonstrated that the production of pragmatic inferences is associated with slower reaction times (Bott and Noveck 2004; Breheny et al. 2006; Noveck and Posada 2003; Pijnacker et al. 2009), we decided to compare reaction times among

[4] The correlation between VIQ and scalar inference rate is not reported for the ASD group as a whole. This correlation is not found in the AS group or in the TD group, both $ps > .1$.

[5] "A and B" is only true if both A and B are true, whereas "A or B" is true if A alone is true, or if B alone is true, or if both A and B are true when an inclusive reading of the disjunction is chosen.

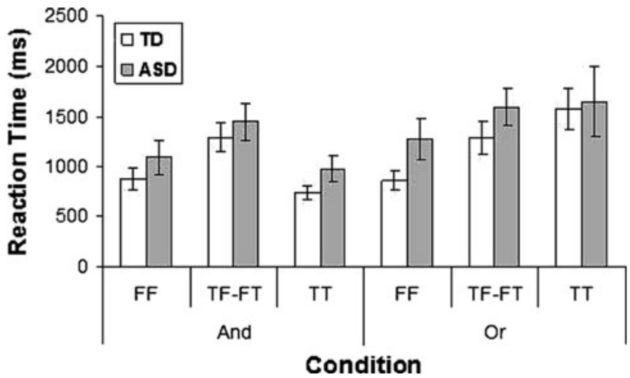

Fig. 2 Mean reaction times in ms as a function of connective type (And, Or), condition (FF, FT/TF, TT) and group (TD, ASD)

"Pragmatic responders" and "Literal responders" in the Or TT condition. To do this, a factorial ANOVA with the between-subject factor "Group" (TD, HFA/AS), "Responder type" (Pragmatic, Literal) and "VIQ Group" (High, Low) was conducted. The analysis revealed no main effect of Responder Type, $F(1,36) = 0.40$, $p = 0.53$, *Partial* $\eta^2 = 0.010$, Group (HFA/AS, TD), $F(1,36) = 0.23$, $p = 0.63$, *Partial* $\eta^2 = 0.006$, or VIQ Group (High, Low), $F(1,36) = 1.21$, $p = 0.28$, *Partial* $\eta^2 = 0.033$. There were no interactions between these factors, but a slight trend for the responder type × IQ group interaction is to be noted, $F(1,36) = 1.96$, $p = 0.16$, *Partial* $\eta^2 = 0.052$ though post hoc Tukey HSD tests indicate that simple comparisons are not significant (all $ps > 0.20$). To summarise, we do not find a lag time difference wherein "Literal responders" are faster than "Pragmatic responders" on the critical (Or TT) items, contrary to what is often reported in these studies. Possible reasons for this unexpected finding are put forward in the discussion.

Discussion

To summarise, our experiment aimed to test whether children with ASD spontaneously produce pragmatic inferences in response to prosodic cues used in disjunctive utterances. Our data indicate, first, that ASD and TD participants produce similar rates of exclusive interpretation of stressed "or". Second, the proportion of children who consistently make the scalar inference does not differ between the two groups. Finally, ASD and TD participants respond at comparable speeds in all the conditions. Overall, these data converge to indicate that participants with ASD are capable of drawing scalar inferences, thereby confirming Pijnacker et al.'s (2009) conclusions and extending them to the processing of spoken language and to a younger sample. The fact that even younger individuals with ASD can derive scalar inferences is remarkable, and shows that the adults included in Pijnacker et al.'s study did not simply grow out of their pragmatic difficulties.

Our further analyses also revealed that VIQ had a different impact on participants' performances depending on which group they belonged to. Higher VIQs were associated with higher rates of scalar inferences in the ASD group, but not in the TD group. This suggests that it is crucial to take VIQ into account when comparing performances in individuals with HFA and AS. In Pijnacker et al.'s study, for instance, some of the difference observed between the two ASD groups might be explained by their differences in mean VIQ scores. The fact that VIQ correlates with scalar inference production in ASD but not in typically developing participants is nonetheless important. It suggests that participants with ASD need higher verbal skills to produce pragmatic inferences, or that they resort to verbal intelligence in a different way to deal with pragmatic tasks. This is in line with findings showing that performance in Theory of Mind related tasks correlate with VIQ in ASDs more than in Down Syndrome or typical development (Fisher et al. 2005; Happé 1995). However, the exact nature of the mechanisms used to produce scalar inferences in ASD requires further investigation, and might benefit from recent work on typically developing children which demonstrates that some transient difficulties in young children are due to limitations in cognitive resources (for a review, see e.g. Siegal and Surian 2004). In relation to this, it is worth mentioning that stressed "or" yielded fewer inferences in our sample of children than in Chevallier et al.'s adult participants (2008) (55% vs. 73%, respectively). Here too, limitations in cognitive resources may account for the discrepant findings between children and adults, and would be worth investigating. Further work *directly* comparing the effect of stress in adults' and children's production of pragmatic inference would thus be useful.

Another area which requires further investigation is the unexpectedly low performance by both TD and ASD children in the TF/FT condition. As noted in the results section, one possible explanation is that the children misinterpreted "or" as "and", or simply used a matching strategy, checking whether the items mentioned in the utterance were indeed present in the picture. However, as we also noted, this interpretation does not seem to square with the lack of correlation between performance in the TF/FT condition and the TT condition, or with the fact that the pattern of results did not change when children who scored at or below chance in this condition were excluded. This finding clearly deserves further discussion, especially given that similarly low rates of correct responses in various tasks are occasionally reported in the classic developmental literature. For instance, Paris (1973) presented children and teenagers with disjunctive sentences whose content was concrete but arbitrary (e.g. “The bird is in the nest or the shoe is on the foot”), and found poor performance in the TF and FT conditions, with 33.5% of correct responses in 10–11 year olds and 51% in 13–14 year olds. Using another paradigm where a puppet made disjunctive statements about the content of a box (e.g. “There is either a horse or a duck in the box”), Braine and Rumain (1981) also found low scores in the TF and FT conditions (see also, Sternberg 1979). This suggests that our result pattern is not an isolated finding which can be simply explained in terms of our specific task’s demands. One possible explanation is that the children are actually not making mistakes, but are rather interpreting unstressed “or” in these types of case as a ‘free choice’ disjunction (Chierchia et al. to appear). On a free choice interpretation, the disjunction “There is a horse $\text{or}_{[\text{free choice}]}$ goat” is understood as authorising the hearer to treat *both* disjuncts as true and to act on whichever of them he chooses. Note that such an interpretation is often encountered in daily interactions. Suppose for instance that you are invited to a dinner party and the host tells you (10):

(10) There’s beer or wine in the fridge.

Now, if you open the fridge and see only beer, you will probably think that the speaker was wrong in uttering (10). It is possible that children are exposed to these free choice uses of unstressed disjunctions quite often, and start off favouring this interpretation. One could then hypothesise that the use of stressed “or” in the Or TT condition would encourage a move away from the ‘free choice’ interpretation to an inferentially enriched interpretation equivalent to exclusive “or”. It should follow that if one re-tested using a stressed “or” in the TF/FT case, results would dramatically improve.

Another unexpected finding in this paper was the absence of a reaction time difference between pragmatic and literal responders in the Or TT condition. As noted in the introduction, a number of studies have provided convincing evidence that the inference making process is a costly one, which requires extra time and effort. In line with this, Pijnacker et al. (2009) found an effect of responder type, with slower responses from pragmatic responders than from logical responders. One important difference between those studies and the present work is that they involved only adult participants; to the best of our knowledge, no scalar study has measured reaction times in children or teenagers. It is possible that variability in reaction times is greater in children than in adults, and that the age range would need to be drastically narrowed in order to reduce variability enough to detect a lag difference. Another possibility is that some children start off by treating stressed “or” as either linguistically encoding an exclusive interpretation, or else as triggering a ‘default’ scalar inference. In that case, they should be just as fast as literal responders in producing their answers, with fully context-dependent pragmatic inferences appearing later in development. A further potentially important difference between our study and Pijnacker et al.’s is that theirs involved the use of unstressed ‘or’, which leaves it much more up to the hearer to spot the need for the inference. By providing a very helpful cue that such an inference is needed, the use of stressed ‘or’ might indeed be enough to neutralise the difference between pragmatic and literal responders.

Finally, we would like to turn to the main finding of this paper: the unexpected relative ease with which participants with ASD were able to draw scalar inferences. Contrary not only to our own expectations, but also to the assumption that there is a universal pragmatic deficit in ASDs, participants with ASD turned out to produce as many pragmatic enrichments as the controls. When null effects are observed, there is always a concern that the sample size was too small for an existing difference to be detected. Although we cannot exclude this possibility, we would like to stress that our sample size is standard in autism research, that none of the reported effects size indexes exceed Cohen’s value even for a small effect, and that our findings replicate results which have also been obtained elsewhere, in a study conducted independently from ours. Still, the fact that children with ASD should derive pragmatic inferences as often as typically developing children, and at similar speeds, calls for some further discussion.

These data are indeed in sharp contrast with much of the literature on pragmatics in ASDs. As noted in the introduction, much research on verbal individuals with autism indicates that their pragmatic abilities are disrupted (Tager-Flusberg et al. 2005). However, it is also evident that the higher end of the spectrum is not associated with a complete absence of pragmatic competence. For one thing, many individuals with HFA have functional language: they can hold a conversation (albeit in ways that some will consider awkward), they read books and newspapers, write autobiographies, and so on. Moreover, some individuals reach this level of communicative competence without experiencing any sort of language acquisition delay—and research in developmental psychology has provided strong evidence that there is an intimate relationship between vocabulary acquisition and pragmatic skills (Bloom 2000). Although these observations are compatible with *some* degree of pragmatic impairment, they clearly do not square with the idea that there is a *complete* lack of pragmatic competence in individuals on the high end of the autism spectrum. In fact, the writings of such individuals provide evidence that areas traditionally thought of as massively impaired can be functional. Consider, for example, how Donna Williams, an adult with HFA, refers to her collection of poems (2004a) using beautiful metaphorical language: "My most natural language is actually poetry, that world between words and music where the two dance together on the page" (Williams 2004b).

This anecdotal evidence is also backed by experimental evidence from case and group studies. For instance, in a case study on a teenage boy with AS, Smith et al. found flawless performance "on judgements of irony and sarcasm, metaphor, jokes, the use/mention distinction, and other examples involving metarepresentational ability" (Smith et al. 2003). Recent work on irony also points in this direction. For instance, Wang and his collaborators (Wang et al. 2006) report a significant difference in performance between the AS group and the controls, but a dazzling 85% of participants with Asperger Syndrome did pass the irony task, and reaction times indicate that they did so at similar speeds. It is also worth noting that such observations are not merely recent phenomena. For instance, in Happé's (1993) seminal paper, a subgroup of children with autism passed second-order ToM tests and were able to understand both metaphorical and ironical utterances.

Taken together, these data suggest that there is at least a subpopulation of individuals on the autism spectrum with some degree of functional pragmatic competence. However, in analysing pragmatic performance in ASDs, it is necessary to pay attention to the specificities of the pragmatic phenomenon at stake. Although at the most general level pragmatic processes form a natural kind and all have something in common (i.e., they are geared to identifying the communicator's meaning using contextual information combined with verbal or non-verbal cues), there are important differences between pragmatic phenomena of different sub-types. Among figurative utterances, for instance, some, like irony, have been claimed to require higher order ToM skills, while others, like metaphor, only require lower order ToM (Happé 1993; Wilson in press). Moreover, recognition of a communicator's meaning can vary in the amount of verbal skill and contextual or encyclopaedic knowledge required: understanding a point or an ostensive sigh requires no specific lexical knowledge, while the interpretation of an utterance or a piece of poetry generally requires not only lexical knowledge but considerable background or encyclopaedic information. More recently, pragmatic theorists have begun to distinguish between pragmatic processes geared to inferring the speaker's explicit meaning (what is stated, or asserted) and those geared to inferring the speaker's implicit meaning (or implicatures) (Carston 2002). As noted in the introduction, while it is widely accepted that scalar inferences are genuinely pragmatic, there is still considerable debate about whether they contribute to the speaker's explicit meaning (as argued by Noveck and Sperber 2007) or to implicatures (as argued by, e.g., Geurts 2009; Horn 2004; Levinson 2000), and exactly what type of inferential mechanism is involved. The phenomenon investigated here thus reflects only one of a wide variety of pragmatic processes used in communication. However, the relative ease with which children with ASD manage to derive scalar inferences in our task confirms that some aspects of pragmatics are spared in at least a subgroup of individuals on the autism spectrum. Future work will need to carefully characterise this subpopulation, the nature of their social deficit, and the scope of the pragmatic processes they can readily deal with. In particular, one limitation of the present study is the absence of confirmatory ASD diagnoses using the gold standard ADI-R and ADOS. Apart from ensuring the validity of the diagnosis, including such clinical measures in the future would allow for a better characterisation of the subgroups who pass pragmatic tests.

Another important issue for future research will be to determine to what extent these findings apply in more naturalistic contexts, especially given that some authors have demonstrated that individuals with ASD perform differently in explicit experimental settings and in more naturalistic situations (Klin et al. 2000). In particular, it would be worth exploring whether people with autism *spontaneously* produce scalar inferences in conversational contexts. If a dissociation between pragmatic performances in conversational and experimental settings were to be identified, this would have to be taken into consideration during intervention. Indeed, strategies aimed at improving pragmatic skills in ASDs would have to find some way of dealing with the fact that underlying competence may be present, but may not be readily put to use for other, perhaps more *socio*pragmatic, reasons and to tackle these directly.

Acknowledgments Many thanks to the children and staff in North Hill House (Frome, UK), Southlands (Lymington, UK), Henry Fanshaw School (Dronfield, UK), Chelmer Valley High School (Chelmsford, UK) and Haberdasher's Aske's Boys School (Herts, UK).

Appendix 1

See (Table 4).

Table 4 Complete list of stimuli

Connective	Truth condition	Picture 1	Picture 2	Sentence
And	FF	Basket	Bear	There is a donkey and a cake.
		Balloon	Arm	There is a cow and a broom.
		Bee	Bed	There is a candle and a dress.
		Plane	Apple	There is a bottle and a comb.
		Banana	Ball	There is a button and a dog.
	TF/FT	Button	Brush	There is a banana and a brush.
		Broom	Box	There is a balloon and a box.
		Duck	Donkey	There is a duck and an egg.
		Fish	Dustpan	There is a fish and a lamp.
		Ear	Dress	There is an ear and a key.
	TT	Frog	Flower	There is a frog and a flower.
		Fork	Eye	There is a fork and an eye.
		Fly	Litterbin	There is a fly and a litterbin.
		Foot	Finger	There is a foot and a finger.
		Spoon	Table	There is a spoon and table.
Or	FF	House	Hand	There is a bee or a chair.
		Necklace	Lamp	There is a strawberry or a rabbit.
		Lemon	Key	There is a pig or a skirt.
		Horse	Goat	There is a pear or a pen.
		Knife	Heart	There is a rope or a piano.
	TF/FT	Pen	Lion	There is a house or a lion.
		Rabbit	Nose	There is a horse or a nose.
		Orange	Leaf	There is a hat or a leaf.
		Shirt	Rope	There is a shirt or a bear.
		Shoe	Skirt	There is a shoe or a bed.
	TT	Sun	Train	There is a sun or a train.
		Lorry	Wheel	There is a lorry or a wheel.
		Turtle	Wolf	There is a turtle or a wolf.
		Window	Monkey	There is a window or a monkey.
		Star	Rooster	There is a star or a rooster.

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.

Baltaxe, C. A. M., & Guthrie, D. (1987). The use of primary sentence stress by normal, aphasic, and autistic children. *Journal of Autism and Developmental Disorders, 17*(2), 255–271.

Baron-Cohen, S. (2000). Theory of mind and autism: A 15-year review. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from developmental cognitive neuroscience* (pp. 3–21). Oxford: Oxford University Press.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge: The MIT Press.

Bott, L., & Noveck, I. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51*(3), 437–457.

Braine, M., & Rumain, B. (1981). Development of comprehension of 'Or': Evidence for a sequence of competencies. *Journal of Experimental Child Psychology, 31*(1), 46–70.

Braun, B., & Tagliapietra, L. (in press). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100*(3), 434–463.

Burton-Roberts, N. (2007). *Pragmatics*. Basingstoke: Palgrave-Macmillan.

Carston, R. (1998). Informativeness, relevance and scalar implicature. *Relevance Theory: Applications and Implications, 1*, 79–236.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2009). From acoustics to grammar: Perceiving and interpreting grammatical prosody in adolescents with Asperger Syndrome. *Research in Autism Spectrum Disorders, 3*, 502–516.

Chevallier, C., Noveck, I., Nazir, T., Bott, L., Lanzetti, V., & Sperber, D. (2008). Making disjunctions exclusive. *The Quarterly Journal of Experimental Psychology, 61*(11), 1741–1760.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond* (pp. 39–103). Oxford: Oxford University Press.

Chierchia, G., Fox, D., & Spector, B. (to appear). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics *The handbook of semantics*.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology, 54*(2), 128–133.

Dennis, M., Lazenby, A. L., & Lockyer, L. (2001). inferential language in high-function children with autism. *Journal of Autism and Developmental Disorders, 31*(1), 47–54.

Diehl, J., Watson, D., Bennetto, L., McDonough, J., & Gunlogson, C. (2009). An acoustic analysis of prosody in high-functioning autism. *Applied Psycholinguistics, 30*(03), 385–404.

Dunn, L. M., Dunn, L. M., & Whetton, K. (1997). *The British picture vocabulary scale* (2nd ed.). Windsor, UK: National Foundation for Education Research—Nelson.

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

Fisher, N., Happé, F., & Dunn, J. (2005). The relationship between vocabulary, grammar, and false belief task performance in children with autistic spectrum disorders and children with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 46*(4), 409–419.

Frith, U. (1998). What autism teaches us about communication. *Logopedics Phoniatrics Vocology, 23*(2), 51–58.

Frith, U., & Happé, F. (1994). Language and communication in autistic disorders. *Philosophical Transactions: Biological Sciences, 346*(1315), 97–104.

Geurts, B. (2009). Scalar implicature and local pragmatics. *Mind & Language, 24*, 51–79.

Golan, O., Baron-Cohen, S., Hill, J., & Rutherford, M. (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.

Green, S. B., & Salkind, N. J. (2008). *Using SPSS for Windows and Macintosh: Analysing and understanding data* (5th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

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

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24*(2), 129–154.

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66*(3), 843–855.

Horn, L. (1972). *On the semantic properties of logical operators in English*. Los Angeles: University of California Los Angeles.

Horn, L. (2004). Implicature. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 4–28). Oxford: Blackwell.

Horn, L., & Ward, G. (2004). *The handbook of pragmatics*. Oxford: Blackwell.

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: 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/.

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

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30*(5), 1169.

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.

Klin, A., Schultz, R., & Cohen, D. (2000). *Theory of mind in action: Developmental perspectives on social neuroscience. Understanding other minds: Perspectives from developmental neuroscience* (2nd ed., pp. 357–388). Oxford: Oxford University Press.

Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*: Mit Pr.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2007). Reference and attitude in infant pointing. *Journal of Child Langauge, 34*(1), 1–20.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule—generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*(3), 205–223.

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.

McCann, J., & Peppé, S. (2003). Prosody in Autism Spectrum Disorders: A critical review. *International Journal of Language & Communication Disorders, 38*(4), 325–350.

Mitchell, S., Jessica Brian, P., Zwaigenbaum, L., Roberts, W., Szatmari, P., Isabel Smith, P., et al. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics, 27*(2), S69.

Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition, 78*(2), 165–188.

Noveck, I., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language, 85*(2), 203–210.

Noveck, I., & Reboul, A. (2008). Experimental pragmatics: A Gricean turn in the study of language. *Trends in Cognitive Sciences, 12*(11), 425–431.

Noveck, I., & Sperber, D. (2004). *Experimental pragmatics*. New York: Palgrave Macmillan.

Noveck, I., & Sperber, D. (2007). The why and how of experimental pragmatics: The case of 'scalar inferences'. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 184–212). Basingstoke: Palgrave.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition, 86*(3), 253–282.

Paris, S. G. (1973). Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology, 16*(2), 278–291.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. (2005). Perception and production of prosody by speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220.

Peppé, 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.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Pijnacker, J., Hagoort, P., Buitelaar, J., Teunisse, J.-P., & Geurts, B. (2009). Pragmatic inferences in high-functioning adults with autism and Asperger Syndrome. *Journal of Autism and Developmental Disorders, 39*, 607–618.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). Processing costs and their impact on the development of scalar implicature. *Language Acquisition, 14*(4), 347–375.

Recanati, F. (2003). Embedded implicatures. *Philosophical Perspectives, 17*, 299–332.

Rutherford, M., 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. (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., & Surian, L. (2004). Conceptual development and conversational understanding. *Trends in Cognitive Sciences, 8*(12), 534–538.

Smith, N., Hermelin, B., & Tsimpli, I. (2003). Dissociation of social affect and theory of mind in a case of Asperger Syndrome. *UCL Working Papers in Linguistics, 15*, 303–322.

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.

Southgate, V., Chevallier, C., & Csibra, G. (accepted). 17-month-olds appeal to false beliefs to interpret others' communication. *Developmental Science*.

Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and cognition*. Oxford: Blackwell.

Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mindreading. *Mind and Language, 17*(1), 3–23.

Sternberg, R. (1979). Developmental patterns in the encoding and combination of logical connectives. *Journal of Experimental Child Psychology, 28*(3), 469–498.

Surian, L., Baron-Cohen, S., & Van der Lely, H. (1996). Are children with autism deaf to Gricean Maxims? *Cognitive Neuropsychiatry, 1*(1), 55–71.

Surian, L., & Siegal, M. (2008). Language and communication in autism and Asperger Syndrome. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of neuroscience of language*. Amsterdam: Elsevier.

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

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of Autism and Pervasive Developmental Disorders, 1*, 335–364.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development, 78*(3), 705–722.

Wang, A., Lee, 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.

Wang, A., Lee, S., Sigman, M., & Dapretto, M. (2007). Reading affect in the face and voice: Neural correlates of interpreting communicative intent in children and adolescents with Autism Spectrum Disorders. *Archives of General Psychiatry, 64*(6), 698.

Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech, 49*(3), 367.

WHO. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: World Health Organisation.

Williams, D. (2004a). *Not just anything: A collection of thoughts on paper*. London: Jessica Kingsley.

Williams, D. (2004b). Poetry and Prose Retrieved 8th January 2008, from http://www.donnawilliams.net/poetryprose.0.html.

Williams, K., Tuck, M., Helmer, M., Bartak, L., Mellis, C., & Peat, J. (2008). Diagnostic labelling of Autism Spectrum Disorders in NSW. *Journal of Paediatrics and Child Health, 44*(3), 108–113.

Wilson, D. (in press). Pragmatic processes and metarepresentational abilities: The case of verbal irony. In T. Matsui (Ed.), *Pragmatics and theory of mind*. Amsterdam: John Benjamins.

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

Witwer, A., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders, 38*(9), 1611–1624.